



HAL
open science

The impact of incorrect social information on collective wisdom in human groups

Bertrand Jayles, Ramón Escobedo, Stéphane Cezera, Adrien Blanchet, Tatsuya Kameda, Clément Sire, Guy Theraulaz

► **To cite this version:**

Bertrand Jayles, Ramón Escobedo, Stéphane Cezera, Adrien Blanchet, Tatsuya Kameda, et al.. The impact of incorrect social information on collective wisdom in human groups. *Journal of the Royal Society Interface*, 2020, 17 (170), pp.20200496. <10.1098/rsif.2020.0496>. <hal-02991100>

HAL Id: hal-02991100

<https://hal.science/hal-02991100v1>

Submitted on 17 Feb 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

The impact of incorrect social information on collective wisdom in human groups

Bertrand Jayles^{1,2,3,*}, Ramón Escobedo³, Stéphane Cezera⁴, Adrien Blanchet^{4,5},
Tatsuya Kameda⁶, Clément Sire², Guy Theraulaz^{3,5}

¹Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

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

³Centre de Recherches sur la Cognition Animale, Centre de Biologie Intégrative, Centre National
de la Recherche Scientifique (CNRS), Université de Toulouse - Paul Sabatier (UPS), France

⁴Toulouse School of Economics, INRA, Université de Toulouse (Capitole), 31000 Toulouse, France

⁵Institute for Advanced Study in Toulouse, 31015 Toulouse, France

⁶Department of Social Psychology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo
113-0033, Japan

A major problem that results from the massive use of social media networks is the possible spread of incorrect information. However, few studies have investigated the impact of incorrect information on individual and collective decisions. We performed experiments in which participants had to estimate a series of quantities before and after receiving social information. Unbeknownst to them, we controlled the degree of inaccuracy of the social information through “virtual influencers”. We find that a large proportion of individuals partially follow the social information, thus resisting incorrect information. Moreover, incorrect social information can help a group perform better when it overestimates the true value, by partly compensating a human underestimation bias. Our results suggest that incorrect information does not necessarily impair the collective wisdom of groups.

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

26 The digital revolution has changed the way people access and share information. In
27 particular, the past few decades have seen an exponential increase of media sources and
28 amount of available information (1). Moreover, a growing distrust in traditional media
29 has given an increasing share of news consumption to social networks and other pathways
30 to relay information. This facilitated and more diverse access to information may arguably
31 enhance people’s ability to make informed decisions, but at the same time such an information
32 overload dramatically increases the difficulty to verify information, understand an issue,
33 or make efficient decisions (2, 3). In certain cases, it has also disrupted the relationship
34 between citizens and the truth (4, 5), leading to polarized communities unable to listen to
35 each other (6). Recently, the effects of large scale diffusion of incorrect information and fake
36 news on the behavior of crowds have gained increasing interest, because of their major social
37 and political impact (7). The propagation of false information is also reinforced by the use of
38 social bots simulating the behavior of Internet users (8). In particular, there has been recent
39 evidence that fake news can propagate faster and affect people deeper than true information
40 on Twitter, especially when they carry political content (9). In this context, there is a strong
41 need to understand how the diffusion of incorrect information among group members affects
42 individual and collective decisions.

43 To address this issue, we use the experimental framework of estimation tasks, which is
44 highly suitable for quantitative studies on social influenceability (10–16). We performed
45 experiments in which subjects had to estimate a series of quantities with varying levels of
46 demonstrability, before and after having received social information. The demonstrability of
47 a quantity can be interpreted as the amount of prior information a group has about it. To put
48 it in simple terms, it represents the “difficulty” to determining the actual value of a quantity,
49 a notion which will be made explicit and quantitative hereafter. Knowing the individuals’
50 estimates before and after receiving the social information, as well as the value of the social
51 information, we can deduce their sensitivity to social influence. Moreover, by introducing
52 “*virtual influencers*” providing either the true value or some incorrect information – without

53 the subjects being aware of it – we control the *quality* of the information provided to the
54 subjects, and quantify its resulting impact on individual and collective accuracy.

55 To compare estimates of different quantities, it is usually necessary to normalize them.
56 A natural and commonly used normalization consists in dividing estimates by the true value
57 of the quantity of interest (17, 18). Here, we show that this normalization is insufficient for
58 comparing quantities of a very different nature, and that the dispersion of estimates must be
59 included in the normalization process, which has hitherto largely been neglected (11, 12, 16–
60 19). We provide an adequate normalization procedure, and discuss its implications in terms
61 of distributions of estimates. Moreover, we demonstrate that providing a moderate amount of
62 incorrect information to individuals can counterbalance a human tendency to underestimate
63 quantities (20–23), and thereby improve estimation accuracy. We also find that when social
64 information contradicts the underestimation bias, a strategy which consists in compromising
65 between one’s opinion and the group’s opinion increases the performance of groups, even when
66 the social information is substantially inaccurate. However, compromising with inaccurate
67 social information that does not contradict the bias may amplify its deleterious effects, and
68 hamper collective performances.

69 Finally, we use a modified version of an agent-based model developed in (18) to better
70 understand the present results, and to analyze the collective response of human groups to
71 information of which levels of inaccuracy go beyond the values tested in our experiments. The
72 model quantitatively reproduces the experimental results, and confirms the counter-intuitive
73 observation that incorrect information can improve a group’s performance, in particular when
74 the group underestimates the true value and when the information compensates this bias.

75 **Experimental design**

76 180 subjects participated in our experiment. 20 sessions were organized, in each of which 9
77 subjects were asked to estimate 32 quantities. Each quantity was estimated twice: subjects

78 first provided their personal/prior estimate E_p . Next, they received as social information the
79 geometric mean G of the τ previous estimate(s) in the sequence ($\tau = 1$ or 3), and were then
80 asked to provide a second/final estimate E_s . The value of τ was unknown to the subjects and
81 so was the exact nature of the mean provided. Moreover, this second estimate E_s was used
82 to update the social information for the corresponding subject in the next session. Hence,
83 our experiment produced $9 \times 32 \times 20 = 5760$ personal and second estimates, adding up to a
84 total of 11520 estimates.

85 We controlled the quality of the social information provided to the subjects, without them
86 being aware of it. To that end, we inserted in the sequence of 20 final estimates given by
87 the subjects – unbeknownst to them – $n = 0, 5,$ or 15 artificial estimates. These additional
88 estimates correspond to a fraction $\rho = \frac{n}{20+n} = 0\%, 20\%,$ or 43% of “*virtual influencers*”.
89 Each sequence thus consisted of $N = 20 + n = 20, 25,$ or 35 estimates overall, among which
90 20 were estimates given by 20 actual participants, one per session. The influencers’ estimates
91 were introduced at random locations in the sequences. The value T_I of the influencers’
92 estimates provided to the participants was controlled through a parameter α , which represents
93 a distance to the true value quantifying the (in)correctness of the influencers’ estimates
94 ($\alpha = -2, -1, -0.5, 0.5, 1, 1.5, 2, 3$). α was normalized so as to make it independent of
95 the questions and its computation will be described in the Results section. In each session
96 and for each question, a subject was thus assigned a value of ρ, τ and α , and his/her second
97 estimate was a single step in a sequence of 20, 25, or 35 estimates (Supplementary Fig. S1
98 provides a graphic representation of the protocol). Note that the estimates of the *virtual*
99 *influencers* are also used to update the social information which is then provided to the next
100 subjects in the sequence.

101 The quantities to estimate were grouped into four categories: visual perception (number
102 or length of objects in an image), population of large cities in the world, daily life facts, and
103 extreme (astronomical or biological/geological) events. As we will see, the separation into
104 these loosely defined categories is reflected in the collected data. Three additional questions

105 were asked, which cannot be assigned to any of these categories (see the list of questions
106 in Supplementary Information). All experimental details are given in the Materials and
107 Methods section in the Supplementary Information.

108 **Results**

109 **Comparing quantities of very different nature**

110 Because humans perceive numbers roughly as their order of magnitude (24, 25), the logarithm
111 of estimates is the natural quantity to consider in estimation tasks, especially for large quan-
112 tities, rather than the actual estimates themselves. Distributions of estimates have indeed
113 often been found highly right-skewed, while the distribution of their common logarithm is
114 generally much more symmetric (11, 13, 17). An important issue in estimation tasks is to find
115 a proper way to normalize and aggregate estimates arising from questions with very different
116 quantitative answers. Within studies, how can one aggregate estimates of quantities that
117 differ by several orders of magnitude? Between studies, how can we compare findings coming
118 from different sets of quantities?

119 In line with other works (26, 27), we find that the median log-estimate scales linearly with
120 the logarithm $\log(T)$ of the true value (Fig. 1a), which leads to the natural normalization:
121 $X_p = \log\left(\frac{E_p}{T}\right)$. X_p represents the deviation of an estimate from the true value in orders of
122 magnitude, and is often used as the quantity of interest in estimation tasks (11, 12, 17, 18).
123 However, this normalization does not take into account the dispersion of the log-estimates
124 $\langle |\log(E_p) - \text{Median}(\log(E_p))| \rangle$ (where $\langle x \rangle$ refers to the mean of x) which can vary considerably
125 for different questions (Fig. 1b). In the following, we simply refer to X as the “estimates”,
126 dropping the “log-” prefix.

127 Fig. 1c presents the median m_p and Fig. 1d the dispersion $\sigma_p = \langle |X_p - m_p| \rangle$ of the personal
128 estimates X_p , for all questions asked in this experiment (sorted by category of questions).
129 One can notice the extreme variation of both quantities depending on the question, suggesting

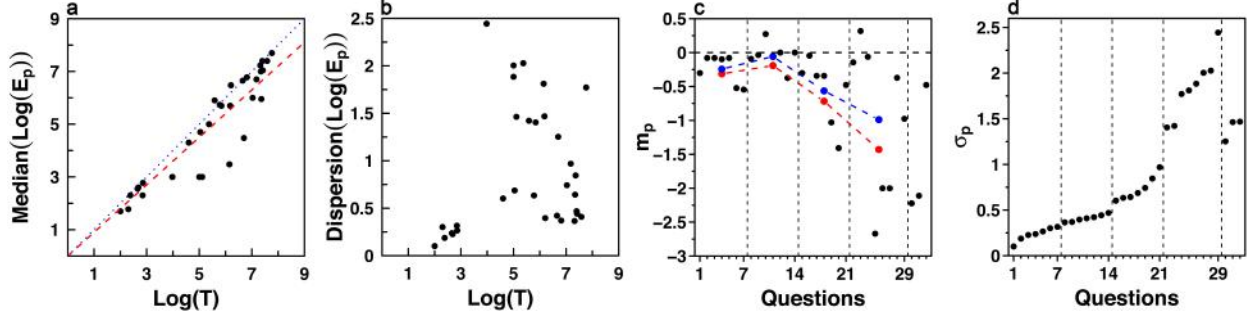


Fig. 1: (a) Median and (b) mean dispersion $\langle |\log(E_p) - \text{Median}(\log(E_p))| \rangle$ of the logarithms of estimates E_p , for the 32 questions asked in the experiment. $\text{Median}(\log(E_p))$ scales linearly with the log of the true value T . The red dashed line in (a) is the linear regression, which slope is lower than 1 (blue dotted line), revealing the human tendency to underestimate quantities. (c) Median m_p and (d) dispersion $\sigma_p = \langle |X_p - m_p| \rangle$ of (log-)estimates $X_p = \log(\frac{E_p}{T})$, for the 32 questions asked in the experiment, whose ID are ranked according to their σ_p which also reflects their demonstrability. The 4 categories of questions (from left to right: visual perception, population of large cities in the world, daily life facts, extreme events), plus the three additional questions, are separated by dashed lines. The categories are well distinct in panel (d), indicating that σ_p is characteristic of the type of quantity to estimate, and more precisely of a question’s demonstrability. In panel (c), the correlation with demonstrability is much less clear, although m_p tends to grow on average when the demonstrability decreases (i.e., when the question ID and σ_p increase). The blue and red dashed lines in (c) are respectively the average value of m_p , and the quantity $-\sqrt{\langle m_p^2 \rangle}$, in each category.

130 that including m_p and σ_p in the normalization process is crucial to compare quantities of a
 131 different nature. Fig. 1d shows that the category of a question is clearly identifiable by the
 132 dispersion of estimates σ_p (but not by the median m_p , see Fig. 1c). The natural classification
 133 that we have chosen a priori is thus reflected in the experimental data. Moreover, we see
 134 that the less demonstrable a question is, the higher the dispersion of estimates. This is
 135 further supported by the three unclassified questions (30 to 32): one could have predicted
 136 that they had a low demonstrability (i.e., that people have little prior information about
 137 them), and that they would therefore be closer to the “extreme events” category than to the
 138 other categories, as observed.

139 Full normalization of estimates

140 In a previous study, we found and justified that the estimates X_p for low demonstrability ques-
 141 tions have a probability distribution function (PDF) close to the Cauchy distribution (18).

142 This property can be explained by a simple probabilistic argument: if two people provide
 143 estimates X_1 and X_2 of a quantity about which they have *no information at all*, then the
 144 average $\frac{X_1+X_2}{2}$ of both estimates cannot be a statistically better estimation of the correct
 145 answer T . Hence, this average has necessarily the same probability distribution as X_1 and
 146 X_2 , and the only distribution that satisfies such a property is the Cauchy distribution (see
 147 also Material and Methods). Our model based on Cauchy distributions convincingly repro-
 148 duced the experimental data, and in particular fitted well with the experimental distribution
 149 of estimates X_p (18).

150 However, as we pointed out above, m_p and σ_p have to be considered to compare esti-
 151 mates for questions with answers spanning several orders of magnitude. Hence, for each
 152 question characterized by its intrinsic median m_p and width σ_p , we normalize the estimate
 153 as $Z_p = \frac{X_p - m_p}{\sigma_p}$. Fig. 2 shows that the normalized estimates Z_p follow the standard Laplace
 154 distribution (i.e., with center 0 and width 1), $f(Z) = \exp(-|Z|)/2$, implying that the X_p
 155 are also Laplace distributed for *individual questions*. It is only when different questions with
 156 arbitrary widths σ_p are aggregated without our normalization that an overall Cauchy-like
 157 distribution for the X_p emerges. Similarly, note that after social influence (red dots), the
 158 $Z_s = \frac{X_s - m_s}{\sigma_s}$, with $m_s = \text{Median}(X_s)$ and $\sigma_s = \langle |X_s - m_s| \rangle$ also follow the standard Laplace
 159 distribution, implying that the X_s also follow a Laplace distribution for each question. We
 160 will therefore slightly modify the model developed previously (18), to replace Cauchy distri-
 161 butions by Laplace distributions (see Model section).

162 By measuring m_p and σ_p and using them in the normalization process, we fix the quantity
 163 $\langle |Z_p| \rangle = 1$, and therefore have *some* information about the distribution, instead of *none* for
 164 the Cauchy distributions argument presented above. As shown in the Material and Methods
 165 by exploiting the principle of maximum entropy, the most likely distribution satisfying such
 166 a constraint is the Laplace distribution.

167 This constraint on the dispersion of estimates can be understood as an intrinsic property of
 168 the system {group of individuals, question}: the dispersion is characteristic of a given group

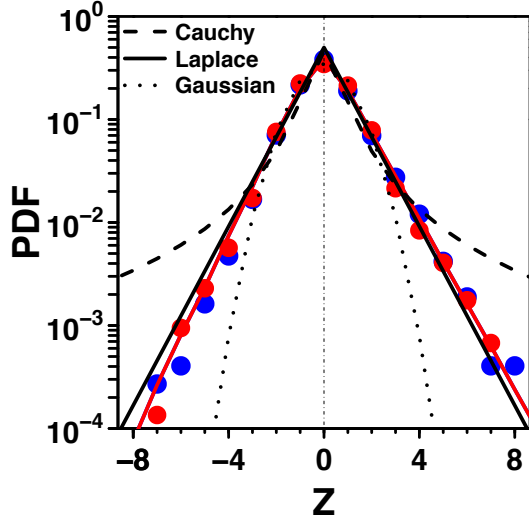


Fig. 2: Distribution of fully normalized estimates $Z = \frac{X-m}{\sigma}$, before (blue) and after (red) social influence. m and σ are respectively the median of the estimates $X = \log(\frac{E}{T})$ and their dispersion, for each corresponding question. E are the actual estimates and T the true value for each corresponding quantity. The black lines are the standard (center 0 and width 1) Laplace distribution (full line), the Cauchy distribution (dashed line) and the Gaussian distribution (dotted line) of same width. The Laplace distribution fits the experimental data the best. Red lines (overlapping blue lines) are model simulations.

169 of individuals estimating a given quantity, and gives the typical range of answers that would
 170 seem reasonable to most people in the group for that question. The lower the demonstrability
 171 of a question (i.e., the lower the amount of prior information held by individuals in a group
 172 about that question), the larger this range. This is intuitive when considering the following
 173 example: an estimate three orders of magnitude from the true value would seem absurd if
 174 one considers the age of death of a celebrity, while it would seem perfectly plausible if one
 175 considers the number of stars in the universe. While the normalization by m_p is somewhat
 176 trivial (it simply shifts the center of the distribution of X to 0 for every question), the
 177 normalization by σ_p is crucial in order to be able to properly compare and aggregate estimates
 178 from different questions (and possibly, from different studies). We wish to insist on the fact
 179 that this prescription is not a mere methodological detail and that it should be adopted by
 180 future works in the field.

181 In Fig. S2, we show the distribution of Z for the four categories of questions. One can
182 notice that for very large quantities (Fig. S2c and d), the left side of the distribution collapses
183 faster than the right side, suggesting that people have an intuition that such quantities must
184 be large, even though they know little about them, such that very small estimates are less
185 frequent. Such asymmetric Laplace distributions can also be derived from the principle of
186 maximum entropy, by adding a constraint that penalizes small or large estimates (see the
187 Materials and Methods section in the Supplementary Information).

188 Model

189 In (18), we have introduced an agent-based model to better understand the effects of in-
190 dividual sensitivity to social influence, and of the quantity of information delivered to the
191 individuals, on collective performance and accuracy observed at the group level in estimation
192 tasks. Estimates X_p are drawn from Laplace distributions, the center and width of which are
193 respectively the median m_p and dispersion $\sigma_p = \langle |X_p - m_p| \rangle$ of the experimental personal
194 estimates X_p for each question. Fig. 3a presents the distribution of estimates X for all ques-
195 tions combined, before (blue) and after social influence (red), as well as the corresponding
196 distributions generated by our model, when the X_p are generated from Cauchy distributions
197 (as in our previous research (18), dashed lines) and Laplace distributions (full lines).

198 The Laplace distribution is able to capture the estimates far from the truth ($X_{p,s} >$
199 5) better than the Cauchy distribution. It is important to mention that in our previous
200 study (18), the range of possible answers were limited to plus or minus 3, 5, or 7 orders of
201 magnitude from the true value, depending on the question. By not allowing extreme answers,
202 we probably increased artificially the probability of estimates in the interval [5,7], making
203 the distribution even closer to a Cauchy distribution.

204 After providing its personal (log-)estimate X_p , each agent receives as social information
205 the arithmetic mean M of the τ previous final estimates in the sequence, among which
206 some information V (provided by the *virtual influencers*) is introduced with probability ρ .

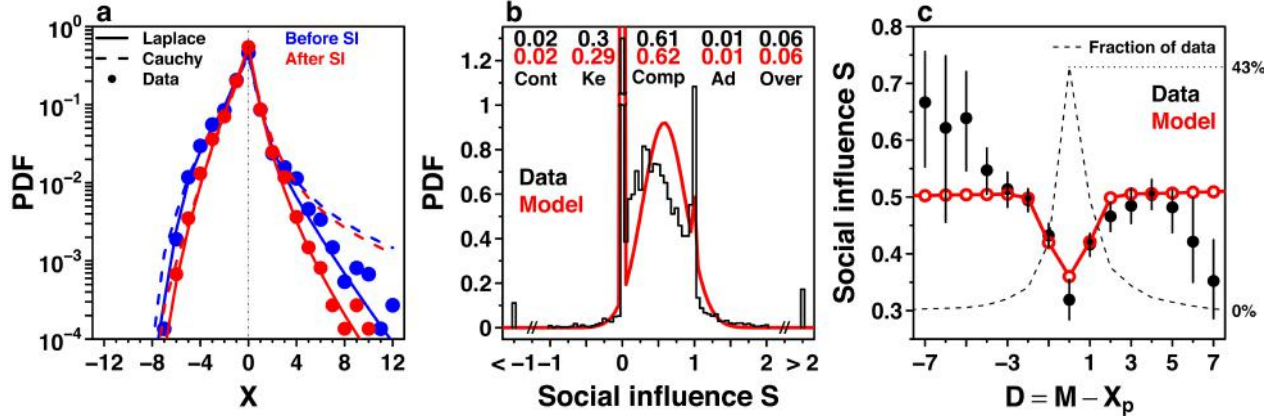


Fig. 3: (a) Probability density function (PDF) of individual estimates, before (X_p , blue) and after social influence (X_s , red). Dots show experimental data, dashed lines are model simulations based on Cauchy distributions and full lines model simulations based on Laplace distributions. Note the sharp decay on the left side of the distribution, well reproduced by the model. The questions asked in our experiment imposed answers higher than one, which translates into $X > -\log(T)$. (b) PDF of the sensitivity to social influence S . The fractions of the five behavioral categories are shown, from left to right: contradicters (“Cont”, $S < 0$), keepers (“Ke”, $S = 0$), compromisers (“Comp”, $0 < S < 1$), adopters (“Ad”, $S = 1$) and overreacters (“Over”, $S > 1$). Experimental data are shown in black, and model simulations in red. (c) Average sensitivity to social influence S against the distance $D = M - X_p$ between the social information M and the personal estimate X_p . Because the average is sensitive to extreme values, we excluded the values such that $|S| > 100$, which represent less than 1% of the data. Black dots correspond to the experimental data, and red empty circles to the model simulations. The dashed line shows the fraction of data for each dot. Beyond three orders of magnitude ($|D| > 3$) are only about 7.6 % of the data, such that we neglect, in the model, the asymmetric weighting of social information observed in this range of values.

207 Note that the actual participants were provided the geometrical mean G of the τ previous
208 estimates. In terms of log-estimates, the social information $M = \log(G)$ indeed transforms
209 into the standard arithmetic mean. The agent then provides a second estimate X_s , defined
210 as the weighted average of its personal estimate X_p and the social information M : $X_s = (1 -$
211 $S) X_p + S M$, where S is the weight given to the social information, that we call *sensitivity to*
212 *social influence*. S can thus be expressed as $S = \frac{X_s - X_p}{M - X_p}$. In Fig. 3b, we show the distribution of
213 S from which five natural behavioral categories can be identified: subjects keep their opinion
214 (“keepers”, $S = 0$), compromise with the social information (“compromisers”, $0 < S < 1$),
215 adopt the social information (“adopters”, $S = 1$), contradict it (“contradicters”, $S < 0$) or
216 overreact to it (“overreacters”, $S > 1$). In the model, after receiving the social information, an
217 agent keeps its personal estimate ($S = 0$) with probability P_0 , adopts the social information

218 ($S = 1$) with probability P_1 , or draws an S in a Gaussian distribution of center m_g and width
219 σ_g with probability $P_g = 1 - P_0 - P_1$.

220 Fig. 3c shows that the average sensitivity to social influence S increases linearly with the
221 distance $D = M - X_p$ between the average social information M and the personal estimate
222 X_p . This is implemented in the model by making the probability P_g increase linearly with
223 D , according to the equation: $\langle S \rangle = P_1 + P_g m_g = a + b|D|$, where the intercept a and
224 the slope b characterize the linear cusp observed in Fig. 3c. More details can be found in
225 the Materials and Methods section in the Supplementary Information. Notice the subjects'
226 tendency to give more weight to social information that is much lower than their personal
227 estimate ($D < -3$), than to social information that is much higher ($D > 3$). Since this
228 concerns only about 7.6 % of the data, we neglect this effect in the model.

229 The model, originally developed in (18), predicted that by providing subjects with in-
230 formation that overestimates the truth, it was possible to improve individual and collective
231 accuracy more than by providing them with the truth itself, by partly compensating the
232 underestimation bias. We provide empirical evidence for this prediction in the next section.

233 Note that the distribution of X narrows after social influence (red dots and lines in
234 Fig. 3a), implying that estimates have overall gotten closer to the truth, all conditions mixed.
235 This may seem counter-intuitive, since in most conditions, incorrect information was provided
236 into the sequence of estimates. To understand this result, we next investigate the impact of
237 incorrect information on estimation accuracy for each condition separately.

238 **Impact of incorrect information on estimation accuracy**

239 As explained above, we controlled the quality of the social information received by the in-
240 dividuals, by introducing $n = 0, 5$, or 15 *virtual influencers* providing artificial estimates of
241 value T_1 randomly inserted in the sequences of 20 estimates provided by the participants,
242 and hence affecting the social information delivered to them. Since we are looking for an
243 information parameter that is independent of the questions, we define, consistently with the

244 previous discussion on the normalization procedure, the normalized (log) deviation from the
 245 truth $\alpha = \frac{\log\left(\frac{T_I}{T}\right)}{\sigma_{\text{pexp}}} = \frac{V}{\sigma_{\text{pexp}}}$ as an indicator of information quality, where σ_{pexp} is an expected
 246 value of the dispersion of personal estimates X_p (the values of the σ_{pexp} are given in Sup-
 247 plementary Table S1), and V the (log) deviation from the truth of the *virtual influencers*
 248 estimates T_I . We obviously did not know the dispersion of estimates before running the
 249 experiment. Yet, since the questions were similar to others used in a previous study (18), we
 250 could formulate reasonable expectations. Indeed, Supplementary Fig. S3 shows that σ_{pexp}
 251 scale more or less linearly with the actual dispersion of estimates σ_p , although it tends to
 252 underestimate it.

253 α represents the deviation of T_I from the truth T in the (expected) natural scale of
 254 each question, which is the dispersion σ_{pexp} of individual errors for that question (X_p are
 255 deviations from the truth or errors). The value T_I introduced in the sequence of estimates is
 256 hence $T_I = T \cdot 10^{\alpha \sigma_{\text{pexp}}}$, and equals the true value T when $\alpha = 0$. Subsequently, to study the
 257 impact of information quality on the group performance, we introduce the variable $Y = \frac{X}{\sigma_p}$,
 258 where σ_p is the dispersion of X_p for a given question, and define:

- 259 • **individual accuracy** as the median of the absolute values of the Y of all individuals
 260 i , averaged over all questions q : $\langle \text{Median}_i(|Y_{i,q}|) \rangle_q$, and
- 261 • **collective accuracy** as the absolute value of the median of the Y of all individuals i ,
 262 averaged over all questions q : $\langle |\text{Median}_i(Y_{i,q})| \rangle_q$.

263 The individual accuracy measures how close individual estimates are to the truth (i.e., close
 264 to 0 in terms of log variables X) on average, while the collective accuracy measures how
 265 close the median estimate is to the truth. Fig. 4 shows that both measures improve after
 266 social influence (i.e., red dots are closer to 0 than blue dots), *over almost the whole range of*
 267 *the considered values of α* , suggesting that incorrect information can, counterintuitively, be
 268 beneficial to the performance of groups. Our results also suggest that individual accuracy
 269 slightly improves after social influence when $\rho = 0\%$ (i.e., no *virtual influencers*), but not
 270 collective accuracy, confirming previous findings (18).

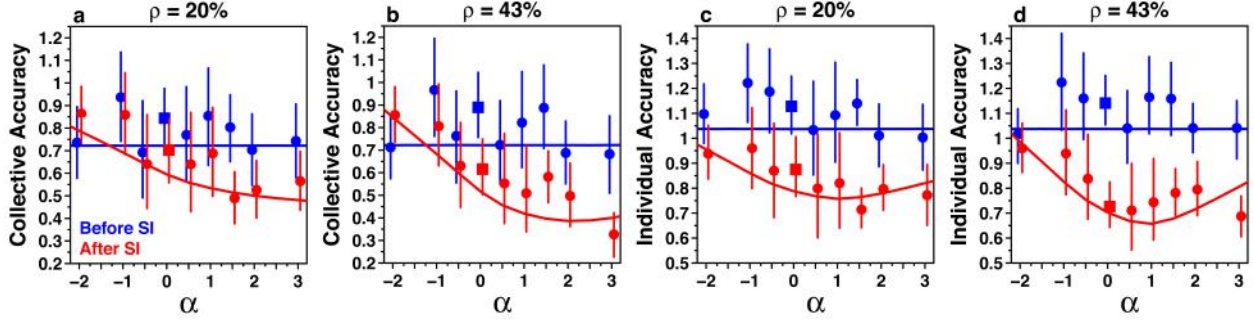


Fig. 4: Collective (a and b) and individual (c and d) accuracy, as a function of the quantifier of information quality α , before (blue) and after (red) social influence, for $\rho = 20\%$ (a and c) and $\rho = 43\%$ (b and d) of influencers in the sequence of estimates. Dots are experimental data from the experiment presented here, while squares at $\alpha = 0$ are experimental data from a previous study, in which the same percentage of *virtual influencers* provided some perfectly accurate information (18). Full lines are model simulations. Surprisingly, incorrect information can be beneficial to collective and individual accuracy, which optima are reached for positive values of α , i.e., for incorrect information that overestimates the truth. The computation of the error bars is explained in the Materials and Methods section in the Supplementary Information.

271 Moreover, the optimum value α_{opt} of α at which collective or individual accuracy im-
 272 proves the most is strictly positive, confirming the model prediction that such improvement
 273 is maximized not by providing perfectly accurate information to individuals, but informa-
 274 tion that overestimates the true value. Such incorrect information partly compensates the
 275 underestimation bias, thus bringing second estimates closer to the truth.

276 Collective accuracy before social influence (blue dots and lines) represents the absolute
 277 value of the collective bias of the group, averaged over all questions, i.e., the distance between
 278 the median estimate and the truth. If the value of the collective bias is $\alpha_0 \approx -0.72$, one may
 279 naively expect that $\alpha_{\text{opt}} = -\alpha_0$ in order to compensate the collective bias and thus optimize
 280 collective accuracy. But, since not all subjects follow the social information fully, one should
 281 rather expect $\alpha_{\text{opt}} > -\alpha_0$, as supported by the data and model.

282 The fraction ρ of *virtual influencers* has no significant effect on collective accuracy in
 283 the data in Fig. 4. However, the simulations of the model predict that collective accuracy
 284 degrades after social influence either when $\alpha < \alpha_{\text{min}} \approx -1.2$ (for both $\rho = 20\%$ and $\rho = 43\%$)
 285 or when $\alpha > \alpha_{\text{max}} \approx 13.4$ for $\rho = 20\%$ and $\alpha_{\text{max}} \approx 7.2$ for $\rho = 43\%$, which corresponds
 286 respectively to $\langle \sigma_{\text{pexp } q} \rangle_q \times \alpha_{\text{max}} \approx 0.46 \times 13.4 \approx 6.2$ and $\langle \sigma_{\text{pexp } q} \rangle_q \times \alpha_{\text{max}} \approx 0.46 \times 7.2 \approx 3.3$

287 orders of magnitude beyond the true value (see Fig. S4). The impact of α is therefore not
288 symmetric with respect to its optimum α_{opt} : incorrect information that largely overestimates
289 the truth can still be beneficial to collective accuracy, while incorrect information that only
290 moderately underestimate the true value is enough to damage collective accuracy. The same
291 analysis remains true for individual accuracy, only with different values of α_{opt} , α_{min} , α_{max} ,
292 and α_0 .

293 **Incorrect information and sensitivity to social influence**

294 It has been shown that estimation accuracy strongly depends on the sensitivity to social
295 influence of individuals in groups (18). Analyzing the above results in the light of the five
296 behavioral categories of sensitivity to social influence (Fig. 3b) helps us to understand the
297 mechanisms underlying them. They cannot be explained by contradicters ($S < 0$), adopters
298 ($S = 1$) or overreacters ($S > 1$), who only represent a small percentage of the population.

299 The top row in Fig. 5 shows individual accuracy as a function of α , for the keepers
300 (Fig. 5a, b) and compromisers (Fig. 5c, d), which each represents a substantial fraction of
301 the population ($\sim 91\%$). Note that the effects are clearer when this separation into behavioral
302 categories is made. Since keepers disregard social information, we observe no improvement
303 in individual accuracy after social influence. However, compromisers, who partly follow the
304 social information, significantly improve their performance over the whole range of incorrect
305 information tested here (except for $\alpha = -2$ and $\rho = 43\%$ of *virtual influencers*). Indeed,
306 because subjects in general, and compromisers in particular, tend to substantially underes-
307 timate quantities, they can improve their estimates by following incorrect social information
308 that is closer to the true value than their own personal estimate. Moreover, partially follow-
309 ing the social information that overestimates the truth allows their second estimates to reach
310 more accurate values, even when the overestimation is quite pronounced. Contrariwise, indi-
311 vidual accuracy degrades quickly when subjects are given incorrect social information which
312 reinforces their natural cognitive bias by underestimating the true value. Compromising thus

313 allows group members to resist incorrect information, as long as this information goes against
 314 their cognitive bias. Similar conclusions can be drawn for collective accuracy (see Fig. S5),
 315 but the patterns are slightly less pronounced.

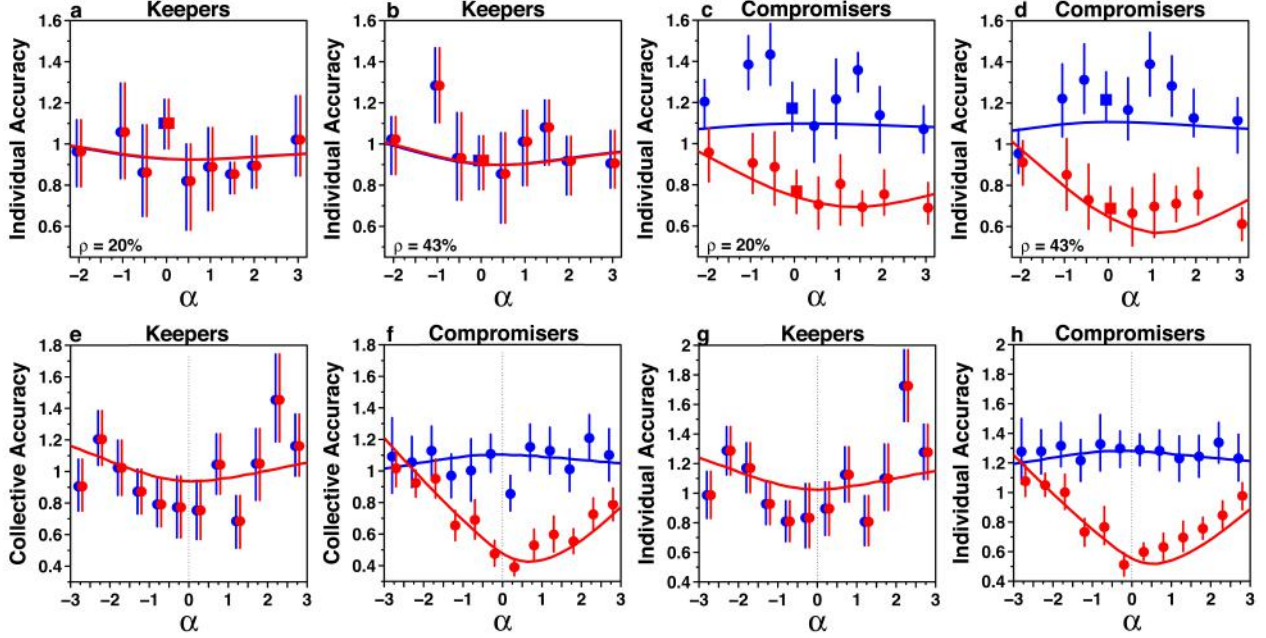


Fig. 5: Top row: individual accuracy as a function of the quantifier of information quality α , before (blue) and after (red) social influence, for $\rho = 20\%$ (a, c) and $\rho = 43\%$ (b, d) of influencers in the sequence of estimates. Keepers ($S = 0$) are shown in (a, b) and compromisers ($0 < S < 1$) in (c, d). Dots are experimental data from the experiment presented here, while squares at $\alpha = 0$ are experimental data from a previous study, in which the same percentage of influencers provided some perfectly accurate information (18). Full lines are model simulations. By disregarding social information, keepers are unable to improve in individual accuracy after social influence. Compromisers however, by partly following social information, improve in individual accuracy after social influence, even when influencers provide some incorrect information; Bottom row: Collective (e, f) and individual (g, h) accuracy against the quantifier of information quality α , before (blue) and after (red) social influence, in the separated experiment with the isolated subjects. (e and g) keepers ($S = 0$); (f and h) Compromisers ($0 < S < 1$). Dots are experimental data, and full lines are model simulations.

316 The bottom row in Fig. 5 shows the equivalent graphs for the “isolated” subjects of our
 317 experiment (see Materials and Methods section in the Supplementary Information). Isolated
 318 subjects received as social information for each question, an estimate T_1 generated from
 319 a *random* value of α uniformly distributed in the interval $[-3, 3]$. These graphs confirm
 320 the above conclusions, but display sharper patterns, due to a discretization effect: social
 321 information in the main experiment was generated from a discrete set of values of α , whereas

322 for isolated subjects, it was drawn from a continuous distribution.

323 Before social influence (blue), we find that keepers are slightly more accurate than com-
324 promisers (average collective accuracy: 0.98 versus 1.07; average individual accuracy: 1.08
325 versus 1.28). This was already observed in (18), and justified by the fact that a higher ten-
326 dency to disregard social information is usually associated with a higher average confidence of
327 the subjects in their own estimates, which often comes with a higher prior knowledge about
328 the quantity to estimate. Note the slight U-shaped curve for keepers in Fig. 5e and g. This
329 effect is a direct consequence of people’s tendency to stick to their personal estimate more
330 when the social information is closer to it (Fig. 3c): when participants receive inaccurate
331 information and retain their opinion, it is often because they were close to it and therefore
332 relatively inaccurate too. Conversely, when participants receive accurate information and
333 keep their opinion, it is often because they were close to it and therefore quite accurate too.
334 Both effects can be observed in Fig. 5 top row, but are less pronounced there.

335 **Influence of the fraction of virtual influencers on individual behavior**

336 We have seen that compromisers, by partially following social information, were able to
337 improve their accuracy over a wide range of incorrect social information. Fig. 6 shows the
338 fraction of keepers and compromisers as a function of α , when $\rho = 20\%$ (Fig. 6a) and $\rho = 43\%$
339 (Fig. 6b) of *virtual influencers* are introduced in the sequence of estimates.

340 When $\rho = 20\%$, both fractions of compromisers and keepers remain more or less indepen-
341 dent of α (Fig. 6a). However, when the proportion of *virtual influencers* providing incorrect
342 information is doubled ($\rho = 43\%$, Fig. 6b), the fraction of compromisers gradually increases
343 (from 0.5 to 0.68, orange line) with α (from -2 to 3), at the expense of the fraction of keepers
344 which decreases (from 0.38 to 0.25, brown line). For this to happen, this increasing tran-
345 sition from keeping to compromising behavior as α increases thus necessitates a significant
346 proportion of subjects to be provided with incorrect social information. Moreover, this result
347 also suggests that subjects not only adapt their behavior to the degree of incorrectness of the

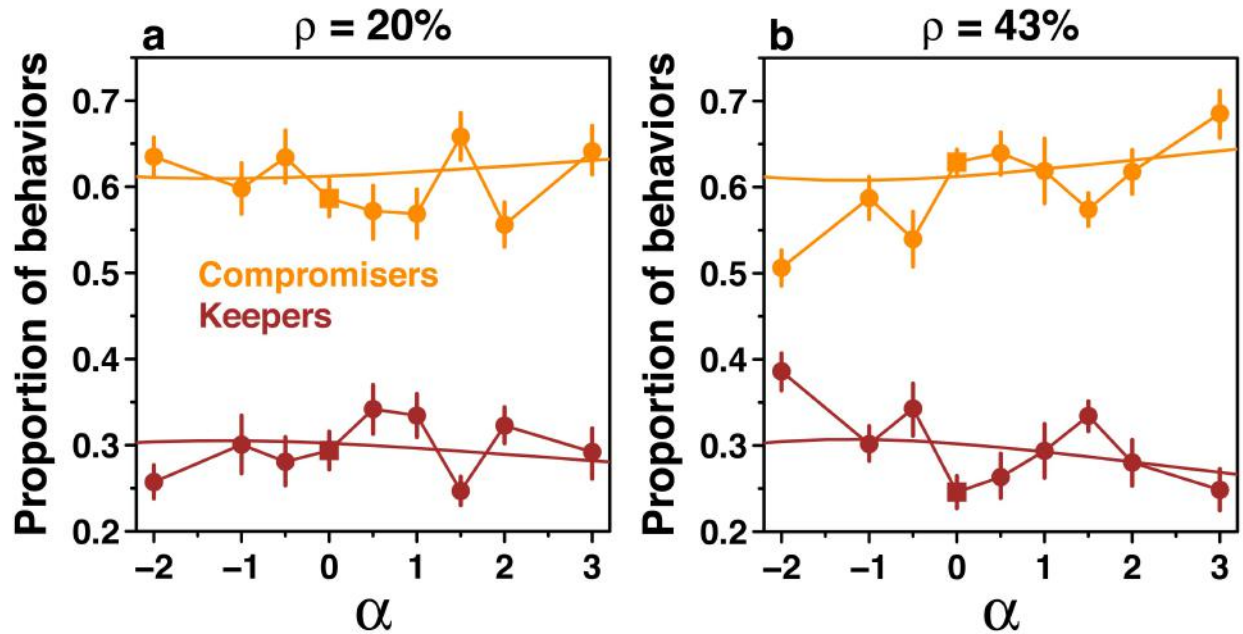


Fig. 6: Proportion of compromisers (orange) and keepers (brown) as a function of the quantifier of social information quality α , for $\rho = 20\%$ (a) and $\rho = 43\%$ (b) of influencers in the sequence of estimates. Dots are experimental data from the experiment presented here, while squares at $\alpha = 0$ are experimental data from a previous study, in which the same percentage of *virtual influencers* provided some perfectly accurate information (18). When $\rho = 43\%$, the fraction of compromisers increases with α at the expense of the fraction of keepers, which decreases with α .

348 social information they receive, but also tend to compromise more with some social infor-
 349 mation that overestimates the truth, than with some social information that underestimates
 350 it. The model predicts this increased behavioral transfer with α , even when $\rho = 20\%$. This
 351 is a direct consequence of the cusp relationship between the sensitivity to social influence S
 352 and the distance to the social information D (Fig. 3c): people tend to compromise with the
 353 social information more as it gets farther from their personal estimates (and from the truth)
 354 on average. However, this effect is significantly stronger in the data, suggesting that other
 355 mechanisms exist, that are not implemented in the model.

356 Supplementary Fig. S6 demonstrates that this increasing fraction of compromisers with
 357 α when $\rho = 43\%$ leads to an increased improvement in individual and collective accuracy
 358 after social influence (but not when $\rho = 20\%$).

Discussion

Understanding the effects of incorrect information on individual and collective decisions is crucial in modern digital societies, where social networks and other vectors of information transmission allow a fast and deep flow of information, the accuracy of which is increasingly hard to verify (28). Here, we rigorously controlled the quality of the information delivered to subjects in estimation tasks, by means of “*virtual influencers*”, i.e., virtual agents inserted into the sequence of estimations – unbeknownst to the subjects – and providing a value whose level of inaccuracy was controlled. We were thus able to precisely quantify the impact of information quality on individual and collective accuracy in those tasks.

We demonstrated that a proper normalization of estimates must take into account their dispersion, which gives the natural range of “reasonable” estimates of a given quantity for a given group. This normalization process led to the conclusion that estimates follow a Laplace distribution when subjects have little prior information about a quantity to estimate. Early research showed that in many data sets, estimates X (i.e., deviations from the truth) were often close either to Gaussian distributed or to Laplace distributed (29, 30). Later work have encompassed Laplace and Gaussian distributions into a broader family of exponential distributions, the Generalized Normal Distribution (GND) family (31, 32), of PDF:

$$f(X, m, \sigma, n) = \frac{1}{2\sigma\Gamma(1+1/n)} \exp\left\{-\left|\frac{X-m}{\sigma}\right|^n\right\}, \quad (1)$$

where m is the center of the distribution (often called *location* parameter), σ is the width of the distribution (often called *scale* parameter) and n is the tailedness (often called *shape* parameter), which controls the thickness of the tails. The fatter the tails of a distribution, the higher the probability to find outliers (i.e., estimates that are very far from the distribution center). More recent work has studied various data sets of estimates and forecasts in the light of Generalized Normal Distributions, and showed that the tailedness of distribu-

382 tions ranged from $n = 1$ (Laplace distribution) to $n = 1.6$, n being equal to 2 for Gaussian
383 distributions (33). They concluded that most distributions of estimates for usual quantities
384 are actually closer to Laplace distributions than to Gaussian distributions. This discussion
385 can be related to the amount of prior information held by a group about a certain quantity.
386 We found that when a quantity is “hard” to estimate (i.e., low demonstrability, correspond-
387 ing to a low amount of prior information about the quantity in the group), the expected
388 distribution of estimates is very close to a Laplace distribution. When a quantity is “easy”
389 to estimate (i.e., high demonstrability, corresponding to a high amount of prior information
390 about the quantity in the group), few outliers are expected, such that the distribution of es-
391 timates could be expected to be closer to a Gaussian distribution. However, our results show
392 that regardless of the questions demonstrability, distributions of estimates are significantly
393 closer to Laplace distributions than to Gaussian distributions when properly normalized, in
394 agreement with (33). In any case, we consider that future studies involving estimation tasks
395 should apply the normalization procedure presented here when comparing and aggregating
396 the estimates of different quantities, for which the width σ_p should be used to quantify their
397 demonstrability.

398 We then studied the impact of incorrect information on individual and collective accuracy,
399 and found that providing incorrect information that overestimates the true value can help
400 a group perform better than providing the correct value itself, by partly compensating for
401 the human underestimation bias. Moreover, collective and individual accuracy can improve
402 after social influence over a surprisingly wide range of incorrect information. This counter-
403 intuitive result is a consequence of a large proportion of individuals compromising with the
404 social information, i.e., partially following it. By doing so, subjects are able to benefit not
405 only from relatively accurate social information, but also from incorrect information that
406 goes against their cognitive bias. Indeed, because of the human tendency to underestimate
407 quantities, partially following an overestimation of the truth – even a large one – can bring
408 second estimates closer to the truth, thus improving accuracy. However, incorrect social

409 information can also harm accuracy if it amplifies the bias. This may be related to some
410 deleterious effects of social information observed at times, for instance, how the spread of
411 misinformation can deeply affect the behavior of crowds as well as public opinion (34, 35).

412 In a former study (18), we showed that adopting the social information was the best
413 strategy in order to improve accuracy, if *virtual influencers* provide perfectly accurate infor-
414 mation in the sequence of estimates. However, while adopting can lead to higher performance
415 than compromising in this particular case, our results show that compromising offers more
416 resilience when the information provided is potentially less accurate.

417 We also found that subjects were sensitive to the degree of incorrectness of the social
418 information they received. They adapted their behavior to the social information, by com-
419 promising more with the social information as it overestimated the truth more, and compro-
420 mising less as it underestimated the truth more. This asymmetric strategy is surprisingly
421 well adapted to counter the human underestimation bias. Indeed, as explained above, fol-
422 lowing (even partly) social information that underestimates the truth may increase the bias,
423 while following social information that overestimates the truth may decrease it. Following
424 less in the former case, and more in the latter is thus bound to increase the performance of
425 groups. Former studies have already observed this subjects' tendency to rely more on social
426 information that is higher than their personal estimate, than on social information that is
427 lower, and showed that it had valuable consequences for collective performance in estimation
428 tasks (26, 27). In (27), it is suggested that people can more easily assess the validity of small
429 numbers compared to large numbers, because they have no direct experience with events
430 related to those large numbers (36), and as a consequence reject more often small numbers
431 provided by the social information.

432 We then used a modified version of a model of collective estimation developed in (18). The
433 predictions of the model are in good agreement with the experimental data, and confirm that
434 to optimize collective accuracy, social information must do more than compensate the initial
435 collective bias, as most individuals only partly follow social information. In addition, the

436 impact of the quality of information is not symmetric with respect to its optimum: collective
437 accuracy can be improved by delivering incorrect information which overestimates the true
438 value by up to several orders of magnitude, whereas it decays fast if the information delivered
439 only slightly underestimates it. In other words, social information reinforcing the bias of the
440 group has a strong negative impact on its accuracy.

441 Overall, we found that incorrect social information does not necessarily impair the collec-
442 tive wisdom of groups, and can even be used to counter some deleterious effects of cognitive
443 biases. Individuals demonstrated an ability to discriminate the validity of social information,
444 depending on its distance from their personal estimates, and thus to benefit from accurate
445 social information, while at the same time resisting inaccurate social information. These
446 results suggest that people may be more resilient to malicious information than is often
447 thought, and at the same time that the negative effects of identified biases can be dampened
448 by exchanging relevant social information, thus improving collective decisions.

449 References

- 450 1. Castells, M. *The Rise of the Network Society*. (Oxford: Wiley-Blackwell, 2009).
- 451 2. Schick, A. G. & Gordon, L. A. Information overload: a temporal approach. *Accounting,*
452 *Organization and Society* **15**, 199–220 (1990).
- 453 3. Klingberg, T. *The Overflowing Brain: Information Overload and the Limits of Working*
454 *Memory* (Oxford University Press, 2009).
- 455 4. Viner, K. How technology disrupted the truth. *Guardian (London)*, July 12. Accessed
456 December 26, 2016. [https://www.theguardian.com/media/2016/jul/12/how-technology-](https://www.theguardian.com/media/2016/jul/12/how-technology-disrupted-the-truth)
457 *disrupted-the-truth*. (2016).

- 458 5. Lewandowsky, S., Ecker, U. K. & Cook, J. Beyond misinformation: Understanding and
459 coping with the “post-truth” era. *Journal of Applied Research in Memory and Cognition*
460 **6**, 353–369 (2017).
- 461 6. Bessi, A., Coletto, M., Davidescu, G. A., Scala, A., Caldarelli, G. & Quattrociocchi, W.
462 Science vs conspiracy: collective narratives in the age of misinformation. *PLoS ONE*
463 **10(2)**, e0118093 (2015).
- 464 7. Vicario, M. D. *et al.* The spreading of misinformation online. *Proceedings of the National*
465 *Academy of Sciences of the United States of America* **113**, 554–559 (2016).
- 466 8. Mønsted, B., Sapieżyński, P., Ferrara, E. & Lehmann, S. Evidence of complex contagion
467 of information in social media: An experiment using Twitter bots. *PLoS ONE* **12(9)**,
468 e0184148 (2017).
- 469 9. Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**,
470 1146–1151 (2018).
- 471 10. Moussaïd, M., Kämmer, J. E., Analytis, P. P. & Neth, H. Social influence and the
472 collective dynamics of opinion formation. *PLoS ONE* **8**, e78433 (2013).
- 473 11. Mavrodiev, P., Tessone, C. J. & Schweitzer, F. Quantifying the effects of social influence.
474 *Scientific Reports* **3**, 1360 (2013).
- 475 12. Madirolas, G. & de Polavieja, G. G. Improving collective estimations using resistance to
476 social influence. *PLOS Computational Biology* **11**, e1004594 (2015).
- 477 13. Chacoma, A. & Zanette, D. H. Opinion formation by social influence: from experiments
478 to modeling. *PLOS ONE* **10**, e0140406 (2015).
- 479 14. Yaniv, I. Receiving other people’s advice: influence and benefit. *Organizational Behavior*
480 *and Human Decision Processes* **93(1)**, 1–13 (2004).

- 481 15. Soll, J. B. & Larrick, R. P. Strategies for revising judgement: how (and how well) peo-
482 ple use others' opinions. *Journal of Experimental Psychology: Learning, Memory, and*
483 *Cognition* **35(3)**, 780–805 (2009).
- 484 16. Luo, Y., Iyengar, G. & Venkatasubramanian, V. Social influence makes self-interested
485 crowds smarter: an optimal control perspective. *IEEE Transactions on Computational*
486 *Social Systems* **5(1)**, 200–209 (2018).
- 487 17. Lorenz, J., Rauhut, H., Schweitzer, F. & Helbing, D. How social influence can undermine
488 the wisdom of crowd effect. *Proceedings of the National Academy of Sciences* **108**, 9020–
489 9025 (2011).
- 490 18. Jayles, B. *et al.* How social information can improve estimation accuracy in human
491 groups. *Proceedings of the National Academy of Sciences* **114**, 12620–12625 (2017).
- 492 19. Kerckhove, C. V. *et al.* Modelling influence and opinion evolution in online collective
493 behaviour. *PLoS ONE* **11(6)**, e0157685 (2016).
- 494 20. Krueger, L. E. Single judgements of numerosity. *Perception & Psychophysics* **31(2)**,
495 175–182 (1982).
- 496 21. Krueger, L. E. Reconciling Fechner and Stevens: toward a unified psychophysical law.
497 *Behavioral and Brain Sciences* **12**, 251–320 (1989).
- 498 22. Izard, V. & Dehaene, S. Calibrating the mental number line. *Cognition* **106(3)**, 1221–
499 1247 (2008).
- 500 23. Scheibehenne, B. The psychophysics of number integration: evidence from the lab and
501 from the field. *Decision* **6(1)**, 61–76 (2019).
- 502 24. Dehaene, S. The neural basis of the Weber–Fechner law: a logarithmic mental number
503 line. *TRENDS in Cognitive Sciences* **7(4)**, 145–147 (2003).

- 504 25. Dehaene, S., Izard, V., Spelke, E. & Pica, P. Log or linear? Distinct intuitions of the
505 number scale in western and amazonian indigene cultures. *Science* **320**, 1217–1220
506 (2008).
- 507 26. Kao, A. B. *et al.* Counteracting estimation bias and social influence to improve the
508 wisdom of crowds. *Journal of the Royal Society Interface* **15**, 20180130 (2018).
- 509 27. Jayles, B. & Kurvers, R. H. J. M. Debiasing the crowd: selectively ex-
510 changing social information improves collective decision-making. Preprint at
511 <https://doi.org/10.31234/osf.io/hn8rz> (2020).
- 512 28. Centola, D. *How behavior spreads: The science of complex contagions*. (Princeton:
513 Princeton University Press, 2018).
- 514 29. Laplace, P. S. Mémoire sur la probabilité des causes par les évènements. *Mémoires de*
515 *l'Academie Royale des Sciences Présentés par Divers Savan* **6**, 621–656 (1774).
- 516 30. Wilson, E. B. First and second laws of error. *Journal of the American Statistical Asso-*
517 *ciation* **18**, 841–851 (1923).
- 518 31. Rider, P. R. A generalized law of error. *Journal of the American Statistical Association*
519 **19**, 217–220 (1924).
- 520 32. Nadarajah, S. A generalized normal distribution. *Journal of Applied statistics* **32**, 685–
521 694 (2005).
- 522 33. Lobo, M. S. & Yao, D. Human judgement is heavy tailed: empirical evidence and im-
523 plications for the aggregation of estimates and forecasts. *INSEAD working paper No.*
524 *2010/48/DS* (2010).
- 525 34. Karlova, N. A. & Fisher, K. E. A social diffusion model of misinformation and disinfor-
526 mation for understanding human information behaviour. *Information Research* **18(1)**,
527 1–12 (2013).

528 35. Mocanu, D., Rossi, L., Zhang, Q., Karsai, M. & Quattrociocchi, W. Collective attention
529 in the age of (mis)information. *In: Computing for Human Learning, Behaviour and*
530 *Collaboration in the Social and Mobile Networks Era, Computers in Human Behavior* **51**
531 **(Part B,)** 1198–1204 (2015).

532 36. Resnick, I., Newcombe, N. S. & Shipley, T. F. Dealing with big numbers: Representation
533 and understanding of magnitudes outside of human experience. *Cognitive Science* **41**,
534 1020–1041 (2017).

535 **Data:** The data supporting the findings of this study are available at **figshare:**

536 <https://doi.org/10.6084/m9.figshare.11903421>

537 **Funding:** This work was supported by Agence Nationale de la Recherche project 11-IDEX-
538 0002-02–Transversalité–Multi-Disciplinary Study of Emergence Phenomena, a grant from the
539 CNRS Mission for Interdisciplinarity (project SmartCrowd, AMI S2C3), and by Program
540 Investissements d’Avenir under Agence Nationale de la Recherche program 11-IDEX-0002-
541 02, reference ANR-10-LABX-0037-NEXT. B.J. was supported by a doctoral fellowship from
542 the CNRS, and R.E. was supported by Marie Curie Core/Program Grant Funding Grant
543 655235–SmartMass. T.K. was supported by Japan Society for the Promotion of Science
544 Grant-in-Aid for Scientific Research JP16H06324 and JP25118004.

545 **Author Contributions:** B.J., C.S., and G.T. designed research; B.J., R.E., S.C., A.B.,
546 T.K., C.S., and G.T. performed research; B.J., C.S., T.K., and G.T. analyzed data; and
547 B.J., C.S., and G.T. wrote the article with critical input from all other authors.

548 **Competing Interests:** The authors declare that they have no competing interests.

16 Experimental design

17 180 subjects – mostly students from the University of Toulouse – participated in our exper-
18 iments. 20 sessions were organized, in each of which 9 subjects had to answer 32 questions
19 for which they had first to give a prior/personal estimate and then a second/final estimate,
20 the latter after being confronted to social information. Each final estimate in each session
21 constituted a step in a sequence of 20 estimates. Hence, our main experiment (see the ex-
22 periment with “isolated subjects” hereafter) produced a total of 9 subjects \times 32 questions \times
23 20 sessions = 5760 final estimates and the same number of prior estimates. For each ques-
24 tion, subjects first provided their prior/personal estimate E_p . Next, they received as social
25 information the geometric mean G of the τ previous estimate(s) in the sequence ($\tau = 1, 3$),
26 and were then asked to provide a second/final estimate E_s (see illustration in Supplementary
27 Figure S1). Subjects were oblivious to the actual value of τ , only being told that the social
28 information was the average of the final answers of *some* previous participants. The time
29 allowed to provide an answer was limited to 40s per estimate, after which a blinking text
30 urging the subjects to answer quickly would appear on their computer screen.

31 The specificity of our experiment lies in the design of a system aimed to control the
32 quantity and quality of the social information provided to the subjects, without them being
33 aware of it. To that end we inserted, at random locations in the sequence of estimates (and
34 unbeknownst to the subjects), $n = 0, 5$, or 15 artificial estimates, corresponding to a fraction
35 $\rho = n/(20 + n) = 0\%$, 20%, or 43% of *virtual influencers*. Each sequence thus consisted of
36 a total of $N = 20 + n = 20, 25$, or 35 estimates and the estimates of the *virtual influencers*
37 were also used to compute the updated social information for the following participants in
38 each sequence. We controlled the value T_I of the artificial estimates, through a parameter
39 α ($\alpha = -2, -1, -0.5, 0.5, 1, 1.5, 2, 3$) defined such that $T_I = T \cdot 10^{\alpha \sigma_{\text{pexp}}}$, where T is the true
40 value of a quantity to estimate, and σ_{pexp} an expected value of the dispersion of estimates
41 σ_p to be obtained. The closer T_I to T (i.e., the closer α to 0), the higher the information
42 quality. In particular, when $\alpha = 0$, artificial estimates are equal to the true value ($T_I = T$). A

43 proper justification of the choice of α as the quantifier of information quality, and a detailed
44 explanation of its physical interpretation, are provided in the main text. The range of values
45 of α used in the experiment was chosen by exploiting preliminary simulations of the model
46 developed in (18), that we describe in detail below.

47 Each participant in each session was associated a value of ρ ($\rho = 0\%$ for subject 1,
48 $\rho = 20\%$ for subjects 2 to 5 and $\rho = 43\%$ for subjects 6 to 9). For subjects 2 to 9 ($\rho \neq 0\%$;
49 no α for $\rho = 0\%$), one particular value of α was associated to each question. We distributed
50 the values of α such that at the end of each session (20 sessions), each combination of a
51 fraction ρ of influencers ($\rho \neq 0$) and information quality α was repeated 16 times, such
52 that the experimental points in Figures 5, 6 and 8 are all constructed from $16 \times 20 = 320$
53 values. For $\rho = 0\%$, there were 32 sequences per session (640 values). Although our previous
54 study did not show evidence for a significant effect of τ on the dynamics (?), we decided
55 to perform an additional check. A value of τ (1 or 3) was thus associated to each question,
56 independent of the values of ρ and α . Preliminary results confirmed that the effect of τ can
57 be neglected, and we thus combined the data from both values of τ in all graphs presented in
58 this paper. Supplementary Table S1 summarizes the conditions (ρ, α, τ) for each subject and
59 each question in a session. In all sessions, these conditions were repeated and the order of the
60 questions was randomized. In the first session (i.e., the first step of all sequences), subjects
61 received as social information (initial condition) the value T_1 provided by the influencers.
62 In the particular case $\rho = 0\%$ (no influencers), the initial social information provided was
63 the true answer to the question T . Anyway, note that our previous study showed that the
64 choice of the initial condition had little impact on the subsequent dynamics (18). At the
65 end of each session, subjects were rewarded according to their overall performance/accuracy:
66 20€ for the two first ones, 15€ for the next four ones and 10€ for the three last ones.

67 As a safeguard against subjects not showing up, 3 additional subjects were recruited in
68 each session. Hence, when more than nine subjects were present in a session, the additional
69 subjects (“isolated subjects”; 51 in total) were given a special treatment, unbeknownst to

70 them: they were not part of a sequence or associated to any specific condition (ρ, α, τ) ,
71 and received instead, as social information, a value $T_I = T \cdot 10^{\alpha \sigma_{\text{pexp}}}$ generated from a random
72 value of α uniformly distributed in the interval $[-3, 3]$. These “isolated subjects” were paid
73 10€, independently of their performance.

74 Cauchy and Laplace distributions

75 In the context of estimation tasks, the Cauchy and Laplace distributions emerge naturally.

76 Cauchy distribution

The Cauchy distribution centered at m (the median) and of width σ is given by

$$f(X, m, \sigma) = \frac{1}{\pi} \frac{\sigma}{(X - m)^2 + \sigma^2}. \quad (\text{S1})$$

77 The Cauchy distribution has fat tails and an infinite variance. One of the notable and unique
78 properties of the Cauchy distribution is that, if X_1, \dots, X_N are independently drawn from the
79 same Cauchy distribution, then the average $\bar{X}_n = \sum_i^n X_i/N$ (or any weighted average of
80 the X_i) has *exactly the same* Cauchy distribution as the X_i , i.e., with the same center and
81 width. In particular, the distribution of the average \bar{X}_n does not have a smaller width than
82 the distribution of the X_i , as one would obtain if the X_i were drawn from a distribution
83 with a finite variance. For instance, if the X_i are drawn in the same Gaussian distribution
84 with standard deviation σ , the average \bar{X}_n is also Gaussian distributed, but with a reduced
85 standard deviation $\sigma_n = \sigma/\sqrt{N}$, a basic property leading to the law of large numbers.

86 In the context of estimation tasks, if N subjects having absolutely no clue about the
87 actual answer to a question draw their personal estimates X_1, \dots, X_N from some random
88 distribution, one cannot expect the average of the estimates \bar{X}_n to provide more information
89 than the individual estimates, and \bar{X}_n should then have the same distribution as the X_i . The
90 unique probability distribution satisfying this property is the Cauchy distribution.

91 **Laplace distribution**

Now, let us assume that the distribution for a random variable X is known to be symmetric and centered in m with a known width given by $\sigma = \int_{-\infty}^{\infty} |X - m|f(X) dX$. What is the most likely distribution satisfying this property? To answer this question, one has to maximize the entropy $S(f) = - \int_{-\infty}^{\infty} f(X) \log[f(X)] dX$, subject to the above constraint for σ along with the normalization constraint $\int_{-\infty}^{\infty} f(X) dX = 1$. This optimization problem (maximum entropy principle) with constraints is solved by introducing two Lagrange multipliers α and β , and maximizing with respect to $f(X)$ the functional $F(f)$ given by

$$F(f) = - \int_{-\infty}^{\infty} f(X) \log[f(X)] dX + \alpha \int_{-\infty}^{\infty} f(X) dX + \beta \int_{-\infty}^{\infty} |X - m|f(X) dX. \quad (\text{S2})$$

Expressing that the functional derivative of $F(f)$ with respect to $f(X)$ vanishes (maximum of $F(f)$), one obtains

$$- \log[f(X)] - 1 + \alpha + \beta|X - m| = 0, \quad (\text{S3})$$

leading to $f(X) = \exp(1 - \alpha - \beta|X - m|)$. Finally, α and β are computed by expressing the normalization and width constraints leading to the symmetric Laplace distribution

$$f(X, m, \sigma) = \frac{e^{-\frac{|X-m|}{\sigma}}}{2\sigma}. \quad (\text{S4})$$

92 The normalized variable $Z = \frac{X-m}{\sigma}$ has the standard Laplace distribution (i.e., with center 0
93 and width 1), $f(Z) = \exp(-|Z|)/2$.

94 Note that if the assumption of a symmetric distribution is relaxed and that one wants
95 for instance to penalize large or small values of X (in the context of human estimations,
96 this is related to the bias against large answers), one can add a term $\gamma \int_{-\infty}^{\infty} X f(X) dX$
97 in $F(f)$, adding an independent constraint on the mean of X . The same calculation as
98 above leads to an asymmetric Laplace distribution with different exponential tails below and
99 above the maximum of the distribution, reminiscent of what is observed experimentally in

100 Supplementary Figure S2c and S2d for estimates about questions with very large answers.

In physics, when the variable X is the energy of the considered system, the only constraint in addition to the normalization is the knowledge of its average energy $E = \int_{-\infty}^{\infty} X f(X) dX$. Repeating the same optimization procedure as above ($F(f)$ in Eq. (S2) is then proportional to the opposite of the free energy, to be maximized), one finds that the most likely distribution for the energy of the system is the famous Boltzmann distribution

$$f(X) = \frac{e^{-\beta X}}{\mathcal{Z}}, \quad (\text{S5})$$

101 where \mathcal{Z} is a normalization constant, and the Lagrange multiplier β can be ultimately shown
102 to be proportional to the inverse of the temperature \mathcal{T} , $\beta = (k_B \mathcal{T})^{-1}$, where k_B is the
103 Boltzmann constant.

104 **Model**

105 **Main experiment**

106 Our model simulates a sequence of $N = 20 + n$ successive estimates performed by 20 agents
107 and n *virtual influencers* for a given question. The model directly implements the log devia-
108 tions from the truth $X = \log\left(\frac{E}{T}\right)$. A typical run of the model consists of the following steps,
109 for an agent estimating a given quantity in a given condition (ρ, α, τ) :

- 110 1. An agent's personal estimate X_p is drawn from the Laplace distribution $f(X_p, m_p, \sigma_p)$,
111 where m_p and σ_p are respectively the median and average absolute deviation from the
112 median of the experimental distribution of estimates for each question. We impose that
113 estimates E_p are greater than 1, i.e., $X_p > -\log(T)$. This condition explains the fast
114 decay on the left side of the distribution in the main text, Figure 3a;
- 115 2. With probability $\rho = \frac{n}{n+20}$, a *virtual influencer* plays and provides the value $V = \alpha \sigma_{\text{pexp}}$
116 (the values of σ_{pexp} for each question are given in Supplementary Table S1). With

117 probability $(1 - \rho)$, the agent receives as social information the average M of the τ
118 previous estimates (possibly including estimates from *virtual influencers*). $M_0 = V$ is
119 used as initial condition for the first agent;

120 3. The agent's sensitivity to social influence S is drawn from a Gaussian distribution of
121 mean $m_g = 0.58$ and standard deviation $\sigma_g = 0.3$ with probability P_g , or takes the
122 value $S = 0$ or $S = 1$ with probability P_0 and $P_1 = 1 - P_0 - P_g$, respectively (main text,
123 Figure 3b). P_g and P_0 have a linear cusp relationship with the distance $D = M - X_p$
124 between the social information M and the personal estimate X_p (main text, Figure 3c),
125 while $P_1 = 0.012$ is kept independent of D .

126 For a given distance D , the average sensitivity to social influence is $\langle S \rangle = P_0 \times 0 + P_1 \times$
127 $1 + P_g \times m_g = a + b|D|$, where the intercept $a = 0.34$ and the slope $b = 0.09$ are the
128 coefficients of the linear cusp relationship extracted from the main text, Figure 3c. P_g
129 is hence given by $P_g = (a + b|D| - P_1)/m_g$. A plateau is added at $|D| > 3$ by setting
130 an upper limit $P_{g_{\max}} = 0.85$ on P_g .

131 4. The agent's final estimate X_s is the weighted average of the personal estimate and the
132 social information: $X_s = (1 - S) X_p + S M$. The condition $X_s > -\log(T)$ is also
133 imposed on the X_s .

134 5. One starts again from step 1 for the next agent.

135 Apart from α over which we loop in order to make it pseudo-continuous and to explore
136 values beyond those tested in the experiment, our model closely follows the experimental
137 structure: one subject is associated to $\rho = 0\%$, four subjects to $\rho = 20\%$ (5/25) and four
138 subjects to $\rho \approx 43\%$ (15/25). Moreover, each question is associated with the corresponding
139 value τ used experimentally (see Supplementary Table S1). One simulation of the model thus
140 mimics an experimental run, and results are calculated for each simulation, then averaged
141 over 10000 simulations.

142 **Isolated subjects**

143 In a parallel experiment, “isolated” subjects received as social information an estimate gen-
144 erated from a random value of α , uniformly distributed between -3 and 3 . The model for
145 isolated subjects is very similar to the one presented above, and also sticks closely to the ex-
146 perimental structure: instead of receiving the average of τ previous estimates, agents receive
147 a value of α randomly generated from a uniform distribution. There is no loop over α here,
148 and the variables ρ and τ do not exist in this separate experiment. All the rest is the same
149 as in the model of the main experiment. 10000 simulations were run with 200 agents in each.

150 **Computation of the error bars**

151 The error bars indicate the variability of our results depending on the $N_Q = 32$ questions
152 presented to the subjects. We call x_0 the actual measurement of a quantity appearing in the
153 figures by considering all N_Q questions asked. Then, we generate the results of $N_{\text{exp}} = 1000$
154 new effective experiments. For each effective experiment indexed by $j = 1, \dots, N_{\text{exp}}$, we
155 randomly draw $N'_Q = N_Q$ questions among the N_Q questions asked (so that some questions
156 can appear several times, and others may not appear) and recompute the quantity of interest
157 which now takes the value x_j . The upper error bar b_+ for x_0 is defined so that $C = 68,3\%$
158 (by analogy with the usual standard deviation for a normal distribution) of the x_j greater
159 than x_0 are between x_0 and $x_0 + b_+$. Similarly, the lower error bar b_- is defined so that
160 $C = 68,3\%$ of the x_j lower than x_0 are between $x_0 - b_-$ and x_0 . The introduction of these
161 upper and lower confidence intervals is adapted in cases where the distribution of the x_j is
162 unknown and potentially not symmetric.

163 **2 List of Questions**

164 We list below the 32 questions used in the experiment and the corresponding correct answers
165 T . In the original experiment, the questions were asked in French. Questions 1 to 29 were

166 classified into 4 clear-cut categories, and questions 30 to 32 do not clearly belong to either of
167 these categories.

168

169 **1. Visual perception (number or length of objects in an image):**

170 1. Marbles 1: How many marbles are in this jar? $T = 100$



171

172 2. Matches 1: How many matches can you see? $T = 240$



173

174 3. Matches 2: How many matches can you see? $T = 480$



175

176

4. Marbles 2: How many marbles are in this jar? $T = 450$



177

178

5. Matches 3: How many matches can you see? $T = 720$



179

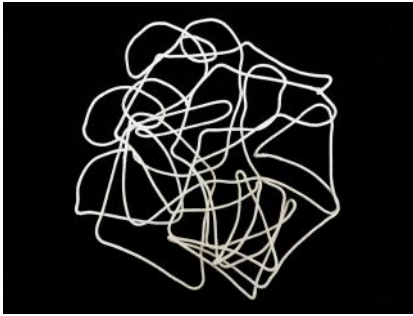
180

6. Rope 1: In your opinion, how long is this rope (in cm)? $T = 200$



181

182 7. Rope 2: In your opinion, how long is this rope (in cm)? $T = 700$



183

184 **2. Population of large cities in the world:**

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

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

187 10. What is the population of Amsterdam and its agglomeration? $T = 1,600,000$

188 11. What is the population of Tokyo and its agglomeration? $T = 38,000,000$

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

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

191 14. What is the population of Shanghai and its agglomeration? $T = 25,000,000$

192 **3. Daily life facts:**

193 15. How many kilometers does a professional cyclist typically bike a year? $T = 40,000$

194 16. What is the mean annual gross salary of a professional league 1 soccer player in France
195 (in euros)? $T = 600,000$

196 17. How many cell phones are sold in France every year? $T = 22,000,000$

197 18. How many cars are stolen in France every year? $T = 110,000$

198 19. How many ebooks were sold in France in 2014? $T = 10,700,000$

199 20. How many books does the American Library of Congress hold? $T = 23,000,000$

200 21. How many people die from cancer in the world every year? $T = 15,000,000$

201 **4. Extreme (astronomical or biological/geological) events:**

202 22. What is the radius of the Sun (in km)? $T = 696,000$

203 23. What is the distance between the Earth and the Moon (in km)? $T = 385,000$

204 24. What is the mean distance between planet Mercury and the Sun (in km)? $T =$
205 $57,800,000$

206 25. What is the total mass of oceans on Earth (in thousand billion tons)? $T = 1,400,000$

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

208 27. How many galaxies does the visible universe hold (in million galaxies)? $T = 100,000$

209 28. How many stars does the Milky way hold (in million stars)? $T = 235,000$

210 29. How many billions kilometers is worth a light-year? $T = 9,461$

211 **5. Other:**

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

213 31. What is the total length of the metal threads used in the Golden Gate Bridge's braided
214 cables (in km)? $T = 129,000$

215 32. How much did Burj Khalifa Tower, in Dubai, cost to build (in thousand dollars)?
216 $T = 1,500,000$

3 Supplementary Figures

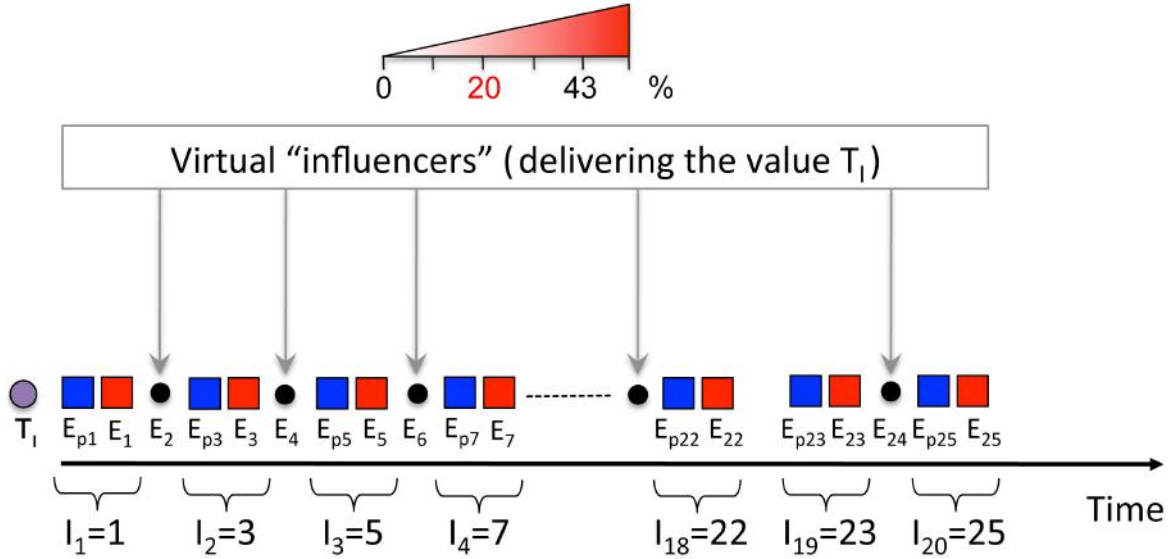


Figure S1: Sequence of estimates for a given question: one after the other, each individual i_k ($k = 1, \dots, 20$) first gives his/her personal estimate E_p (in blue). Then, after receiving social information (the geometric mean of the τ previous estimates), he/she gives his/her second estimate (in red). $n = 0, 5$, or 15 *virtual influencers* (black dots), corresponding to a fraction $\rho = \frac{n}{n+20} = 0, 20$, or 43% respectively, are added at random locations in the sequence, without the subjects being aware of it. The value T_1 provided by the *virtual influencers* is chosen as initial condition (violet circle). The figure shows an example with 5 virtual influencers, corresponding to a fraction $\rho = 20\%$ (highlighted in red on top of the panel). In this example, $\tau = 1$, such that subjects receive as social information the estimate of the subject or *virtual influencer* that comes right before them in the sequence. For instance, subject I_2 receives T_1 as social information from the first *virtual influencer* ($E_2 = T_1$), and subject I_{19} receives E_{22} as social information from subject I_{18} .

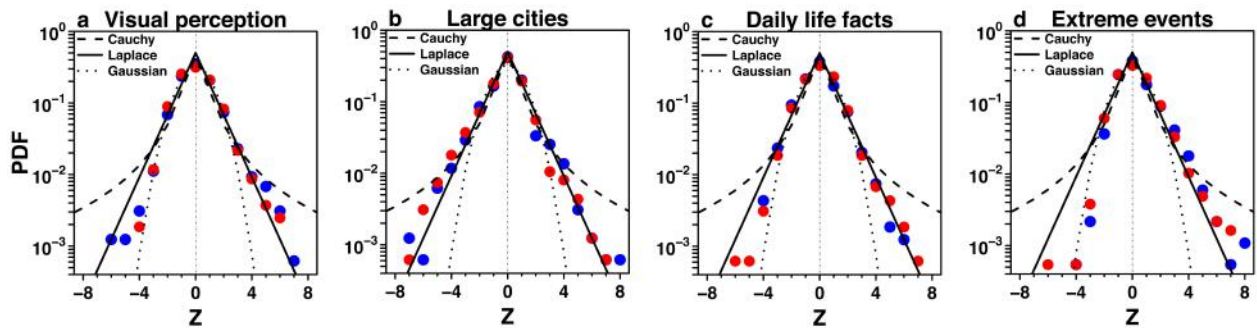


Figure S2: Distribution of fully normalized estimates $Z = \frac{X-m}{\sigma}$, before (blue) and after (red) social influence, for the four categories defined: (a) visual perception, (b) population of large cities in the world, (c) daily life events and (d) extreme (astronomical or biological/geological) events.

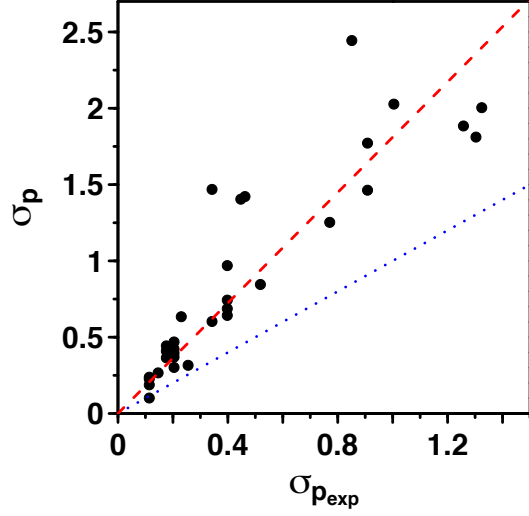


Figure S3: Dispersion $\sigma_p = \langle |X_p - m_p| \rangle$ of estimates $X_p = \log(\frac{E_p}{T})$, for the 32 questions asked in the experiment, against the expected values of the dispersion $\sigma_{p_{\text{exp}}}$ used to define the information T_I provided by the influencers. The red dashed line is a linear regression of slope about 1.8, while the blue dotted line has a slope of 1. The red line being above the blue line means that our expected values underestimated the actual dispersion of estimates.

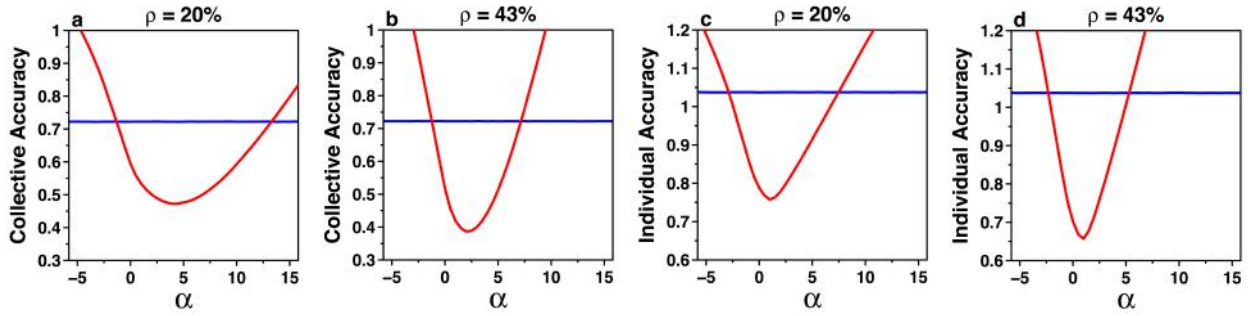


Figure S4: Model simulations of collective (a and b) and individual (c and d) accuracy, as a function of the quantifier of information quality α , before (blue) and after (red) social influence, for $\rho = 20\%$ (a and c) and $\rho = 43\%$ (b and d) of *virtual influencers* in the sequence of estimates. Collective and individual accuracy can improve when *virtual influencers* provide information that overestimate the truth by far, especially when $\rho = 20\%$, but decrease sharply when *virtual influencers* provide information that underestimate the truth: reinforcing the group bias has a strong negative impact on its accuracy.

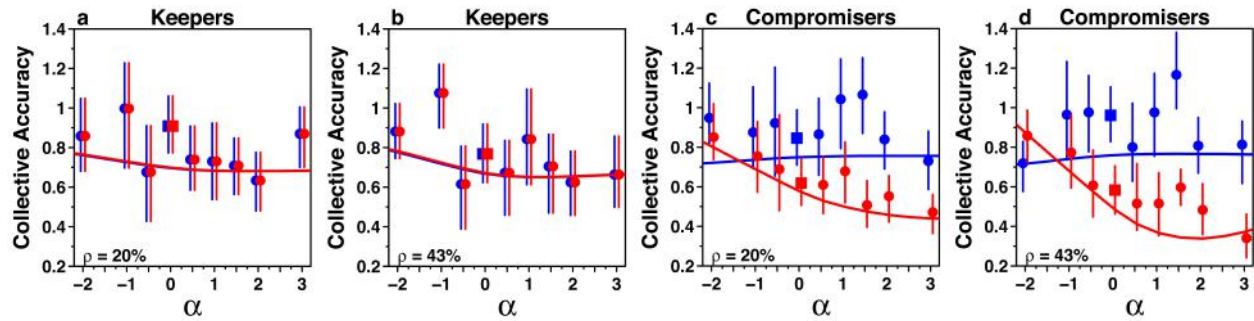


Figure S5: Collective accuracy as a function of the quantifier of information quality α , before (blue) and after (red) social influence, for $\rho = 20\%$ (a, c) and $\rho = 43\%$ (b, d) of influencers in the sequence of estimates. Keepers ($S = 0$) are shown in (a, b) and compromisers ($0 < S < 1$) in (c, d). Dots are experimental data from the experiment presented here, while squares at $\alpha = 0$ are experimental data from a previous study, in which the same percentage of influencers provided some perfectly accurate information (18). Full lines are model simulations. By disregarding social information, keepers are unable to improve in collective accuracy after social influence. Compromisers, however, by partly following social information, improve in collective accuracy after social influence, even when influencers provide some incorrect information.

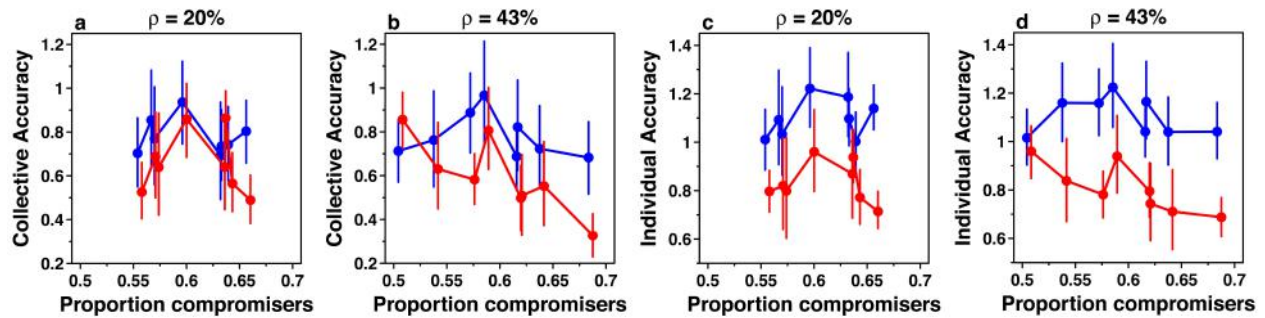


Figure S6: Collective (a and b) and individual (c and d) accuracy, as a function of the proportion of compromisers, before (blue) and after (red) social influence, for $\rho = 20\%$ (a and c) and $\rho = 43\%$ (b and d) of influencers in the sequence of estimates. Each dot corresponds to a value of α . When $\rho = 43\%$, the individual and collective accuracies significantly improve (i.e., gets closer to 0) after social influence (red) with an increasing proportion of compromisers, but not when $\rho = 20\%$.

4 Supplementary Table

Question	τ	σ_{pexp}	$\rho = 20\%$				$\rho = 43\%$			
			Subj. 2	Subj. 3	Subj. 4	Subj. 5	Subj. 6	Subj. 7	Subj. 8	Subj. 9
1	1	0.114	0.5	-0.5	1	-1	0.5	-0.5	1	-1
2	1	0.114	2	-2	3	1.5	2	-2	3	1.5
3	3	0.114	3	1.5	2	-2	3	1.5	2	-2
4	3	0.114	1	-1	0.5	-0.5	1	-1	0.5	-0.5
5	1	0.146	2	-2	3	1.5	2	-2	3	1.5
6	1	0.204	0.5	-0.5	2	-2	0.5	-0.5	2	-2
7	3	0.255	1	-1	3	1.5	1	-1	3	1.5
8	3	0.176	3	1.5	2	-2	3	1.5	2	-2
9	1	0.204	0.5	-0.5	2	-2	0.5	-0.5	2	-2
10	1	0.204	0.5	-0.5	1	-1	0.5	-0.5	1	-1
11	1	0.176	0.5	-0.5	1	-1	0.5	-0.5	1	-1
12	3	0.204	1	-1	3	1.5	1	-1	3	1.5
13	1	0.176	2	-2	3	1.5	2	-2	3	1.5
14	3	0.204	1	-1	0.5	-0.5	1	-1	0.5	-0.5
15	1	0.342	2	-2	0.5	-0.5	2	-2	0.5	-0.5
16	3	0.230	1	-1	0.5	-0.5	1	-1	0.5	-0.5
17	3	0.398	1	-1	2	-2	1	-1	2	-2
18	3	0.398	3	1.5	1	-1	3	1.5	1	-1
19	1	0.398	2	-2	1	-1	2	-2	1	-1
20	3	0.519	3	1.5	0.5	-0.5	3	1.5	0.5	-0.5
21	1	0.398	0.5	-0.5	3	1.5	0.5	-0.5	3	1.5
22	1	0.447	2	-2	1	-1	2	-2	1	-1
23	3	0.462	3	1.5	0.5	-0.5	3	1.5	0.5	-0.5
24	1	0.908	0.5	-0.5	2	-2	0.5	-0.5	2	-2
25	3	1.324	3	1.5	2	-2	3	1.5	2	-2
26	3	1.258	3	1.5	1	-1	3	1.5	1	-1
27	1	1.324	2	-2	0.5	-0.5	2	-2	0.5	-0.5
28	1	1.004	0.5	-0.5	3	1.5	0.5	-0.5	3	1.5
29	3	0.851	1	-1	2	-2	1	-1	2	-2
30	1	0.771	2	-2	1	-1	2	-2	1	-1
31	3	0.908	1	-1	3	1.5	1	-1	3	1.5
32	3	0.324	3	1.5	0.5	-0.5	3	1.5	0.5	-0.5

Table S1: Parameter values in a given session: each subject in a session is associated a specific value of ρ for all questions, but the values of τ and α vary across questions. The ‘‘Subjects’’ columns (Subj. i , $i = 2 \dots 9$) give the values of α for each question.