Equality bias impairs collective decision-making across cultures

Ali Mahmoodi[a,b], Dan Bang[c,d,e], Karsten Olsen[e,f], Yuanyuan Aimee Zhao[g], Zhenhao Shi[h], Kristina Broberg[f], Shervin Safavi[b,i], Shihui Han[g], Majid Nili Ahmadabadi[a,b], Chris D. Frith[e,f,j], Andreas Roepstorff[e,f], Geraint Rees[j,k], and Bahador Bahrami[e,k,1]

[a]Control and Intelligent Processing Centre of Excellence, School of Electrical and Computer Engineering, College of Engineering, University of Tehran, 14395-515 Tehran, Iran; [b]School of Cognitive Science, Institute for Research in Fundamental Sciences, 19395-5746 Tehran, Iran; [c]Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom; [d]Calleva Research Centre for Evolution and Human Sciences, Magdalen College, Oxford OX1 4AU, United Kingdom; [e]The Interacting Minds Centre, Aarhus University, 8000 Aarhus, Denmark; [f]Centre of Functionally Integrative Neuroscience, Aarhus University Hospital, 8000 Aarhus C, Denmark; [g]Department of Psychology, Peking University, Beijing 100871, China; [h]Annenberg Public Policy Center, University of Pennsylvania, Philadelphia, PA 19104-3806; [i]Department of Physiology of Cognitive Processes, Max Planck Institute for Biological Cybernetics, 72012 Tubingen, Germany; [j]Wellcome Trust Centre for Neuroimaging, Institute of Neurology, University College London, London WC1N 3BG, United Kingdom; and [k]Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, United Kingdom

Edited by Richard M. Shiffrin, Indiana University, Bloomington, IN, and approved February 6, 2015 (received for review November 16, 2014)

We tend to think that everyone deserves an equal say in a debate. This seemingly innocuous assumption can be damaging when we make decisions together as part of a group. To make optimal decisions, group members should weight their differing opinions according to how competent they are relative to one another; whenever they differ in competence, an equal weighting is suboptimal. Here, we asked how people deal with individual differences in competence in the context of a collective perceptual decision-making task. We developed a metric for estimating how participants weight their partner's opinion relative to their own and compared this weighting to an optimal benchmark. Replicated across three countries (Denmark, Iran, and China), we show that participants assigned nearly equal weights to each other's opinions regardless of true differences in their competence—even when informed by explicit feedback about their competence gap or under monetary incentives to maximize collective accuracy. This equality bias, whereby people behave as if they are as good or as bad as their partner, is particularly costly for a group when a competence gap separates its members.

social cognition | joint decision-making | bias | equality

When it comes to making decisions together, we tend to give everyone an equal chance to voice their opinion. This is not just good manners but reflects a long-standing consensus ("two heads are better than one"; *Ecclesiastes* 4:9–10) on the "wisdom of the crowd" (1–3). However, a question much contested (4, 5) is, once we have heard everyone's opinion, how should they be combined so as to make the most of them? One recent suggestion is to weight each opinion by the confidence with which it is expressed (6, 7). However, this strategy may fail dramatically (8, 9) when "the fool who thinks he is wise" is paired with "the wise who [thinks] himself to be a fool" (10). In the face of such a competence gap, knowing whose confidence to lean on is critical (11).

A wealth of research suggests that people are poor judges of their own competence—not only when judged in isolation but also when judged relative to others. For example, people tend to overestimate their own performance on hard tasks; paradoxically, when given an easy task, they tend to underestimate their own performance (the hard-easy effect) (12, 13). Relatedly, when comparing themselves to others, people with low competence tend to think they are as good as everyone else, whereas people with high competence tend to think they are as bad as everyone else (the Dunning–Kruger effect) (14–16). In addition, when presented with expert advice, people tend to insist on their own opinion, even though they would have benefitted from following the advisor's recommendation (egocentric advice discounting) (17–19). These findings and similar findings suggest that individual differences in competence may not feature saliently in social interaction. However, it remains unknown whether—and to what extent—people take into account such individual differences in collective decision-making.

To address this issue, we developed a computational framework inspired by recent work on how people learn the reliability of social advice (20). We used this framework to (*i*) quantify how members of a group weight each other's opinions and (*ii*) compare this weighting to that of an optimal model in the context of a simple perceptual task. On each trial, two participants (a dyad) viewed two brief intervals, with a target in either the first or the second one (Fig. 1*A*). They privately indicated which interval they thought contained the target, and how confident they felt about this decision. In the case of disagreement (i.e., they privately selected different intervals), one of the two participants (the arbitrator) was asked to make a joint decision on behalf of the dyad, having access to their own and their partner's responses. Last, participants received feedback about the accuracy of each decision before continuing to the next trial. Previous work predicts that participants would be able to use the trial-by-trial feedback to track the probability that their own decision is correct and that of their partner (21). We hypothesized that the arbitrator would make a joint decision by scaling these probabilities (p_{self}, and p_{other}) by the expressed levels of confidence ($p_{\text{self}} \times c_{\text{self}}$ and $p_{\text{other}} \times c_{\text{other}}$), thus making full use of the information available on a given trial, and then combining these scaled probabilities into a decision criterion ($DC = p_{\text{self}} \times c_{\text{self}} + p_{\text{other}} \times c_{\text{other}}$). In addition, to capture any bias in how the arbitrator weighted their partner's opinion, we included a free parameter (γ) that modulated the influence of the partner in the decision process ($DC = p_{\text{self}} \times c_{\text{self}} + \gamma \times p_{\text{other}} \times c_{\text{other}}$).

In four experiments, we tested whether and to what extent advice taking (γ) varied with differences in competence between dyad members. To anticipate our findings, we found that the

Significance

When making decisions together, we tend to give everyone an equal chance to voice their opinion. To make the best decisions, however, each opinion must be scaled according to its reliability. Using behavioral experiments and computational modelling, we tested (in Denmark, Iran, and China) the extent to which people follow this latter, normative strategy. We found that people show a strong equality bias: they weight each other's opinion equally regardless of differences in their reliability, even when this strategy was at odds with explicit feedback or monetary incentives.

Author contributions: A.M., D.B., K.O., C.D.F., A.R., G.R., and B.B. designed research; A.M., D.B., K.O., Y.A.Z., Z.S., K.B., S.S., S.H., and B.B. performed research; A.M., D.B., K.O., Y.A.Z., and B.B. analyzed data; and A.M., D.B., K.O., M.N.A., C.D.F., A.R., G.R., and B.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: bbahrami@ucl.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1421692112/-/DCSupplemental.

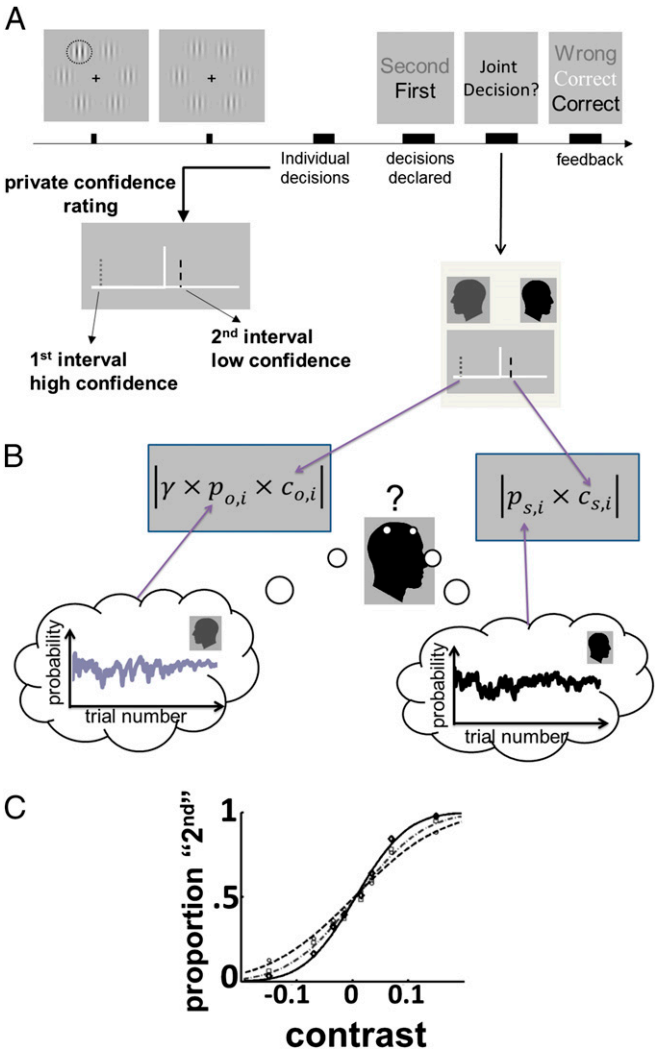

Fig. 1. (*A*) The schematic shows an example trial. Dyad members observed two consecutive intervals, with a target appearing in either the first or the second one (here indicated by the dotted outline). They then privately decided which interval they thought contained the target and indicated how confident they felt about this decision. Next, their individual decisions were shared, and in the case of disagreement (i.e., they privately selected different intervals), one of the two dyad members (the arbitrator) was randomly prompted to make a joint decision on behalf of the dyad. Feedback about the accuracy of each decision was provided at the end of the trial. (*B*) We assumed that each participant tracked the probability of being correct given their own decision (thought bubble on the right) and that of their partner (thought bubble on the left). See main text and *Methods* for details. (*C*) A psychometric function showing the proportion of trials in which the target was reported to be in the second interval given the contrast difference between the second and the first interval at the target location. Circles, performance of the less sensitive dyad member averaged across dyads; gray squares, performance of the more sensitive dyad member averaged across dyads; and black squares, average performance of the dyad.

worse members of each dyad underweighted their partner's opinion (i.e., assigned less weight to their partner's opinion than recommended by the optimal model), whereas the better members of each dyad overweighted their partner's opinion (i.e., assigned more weight to their partner's opinion than recommended by the optimal model). Remarkably, dyad members exhibited this "equality bias"—behaving as if they were as good as or as bad as their partner—even when they (*i*) received a running score of their own and their partner's performance, (*ii*) differed dramatically in terms of their individual performance, and (*iii*) had a monetary incentive to assign the appropriate weight to their partner's opinion.

Recently, psychological phenomena previously believed to be universal have been shown to vary across cultures—where culture is understood as a set of behavioral practices specific to a particular population (22)—calling the generalizability of studies based on Western, Educated, Industrialized, Rich, and Democratic (WEIRD) participants into question (23, 24). To test whether the pattern of advice taking observed here was culture specific, we conducted our experiments in Denmark, China, and Iran. Broadly speaking, we take these three countries to represent Northern European (Denmark), East Asian (China), and Middle Eastern (Iran) cultures. According to the latest World Values Survey,* 71.1% of Northern European respondents (data from Norway and Sweden) and 52.3% of Chinese respondents, but only 10.6% of Iranian respondents, favored the sentence "most people can be trusted" over "you can never be too careful when dealing with people." Reflecting this pattern of responses, research using public goods games has shown that Danish participants contribute more than Chinese and that Chinese participants in turn contribute more than Middle Eastern (data available from Turkey, Saudi Arabia, and Oman) (25, 26). Our sample should thus be sensitive to any cultural commonalities or differences in advice taking.

Results

Experiment 1. A total of 98 healthy adult males (15 dyads in Denmark, 15 dyads in Iran, and 19 dyads in China) were recruited for experiment 1. Each dyad completed 256 trials (Fig. 1*A*). Participants were separated by a screen and not allowed to communicate verbally with each other. The overall pattern of results was consistent between countries; we have included separate analyses for each country in *SI Results*. We quantified performance—that is, sensitivity to the contrast difference between the two intervals—as the slope of the psychometric function relating choices (i.e., individual choices and those made by participants on behalf of dyad) to the stimulus values (Fig. 1*C*; *SI Methods*).

To address the question of how individual differences in competence affect joint decisions, we first asked if the performance of the dyad depended on who arbitrated the trial. More precisely, we calculated dyadic sensitivity (*SI Methods*) separately for the trials arbitrated by the less and the more sensitive member of each dyad. We defined collective benefit as the ratio of the sensitivity of the dyad to that of the more sensitive dyad member (i.e., s_{dyad}/s_{max}). Collective benefit was significantly higher in trials arbitrated by the more sensitive dyad members [$t(46) = -4.18$, $P < 0.0004$; paired t test]. We then asked if participants' confidence reflected their individual differences in performance. Indeed, the more sensitive dyad members were, on average, more confident [Fig. 2*A*; $t(46) = 2.42$, $P < 0.02$, paired t test].

When collapsed across dyad members, participants showed a small egocentric bias, confirming their partner's decision in 45 ± 16% (mean ± SD) of disagreement trials. This result is consistent with previous works on egocentric advice discounting (13, 17). However, splitting the data as above showed that the less sensitive dyad members followed their partner's advice as often as they confirmed their own decision (i.e., 50 ± 20%; mean ± SD). In contrast, the more sensitive dyad members followed their partner's advice on 41 ± 11% (mean ± SD) of the disagreement trials (Fig. 2*B*). This difference in advice taking was significant [Fig. 2*B*; $t(46) = -2.35$, $P < 0.03$, paired t test]. Taken together, the results showed that the more sensitive dyad members had a larger—more positive—influence on the joint decisions.

A critical insight, however, came from assessing participants' confidence in their correct and wrong individual decisions. For each participant, we calculated the probability that a participant expressed high confidence in a correct decision ($P_{high}|correct$) and in an incorrect decision ($P_{high}|error$) separately. We defined high confidence trials as those in which the participant's confidence was higher than their own average confidence. A 2 (dyad member: less vs. more

*A cross-country project coordinated by the Institute for Social Research of the University of Michigan: www.worldvaluessurvey.org/.

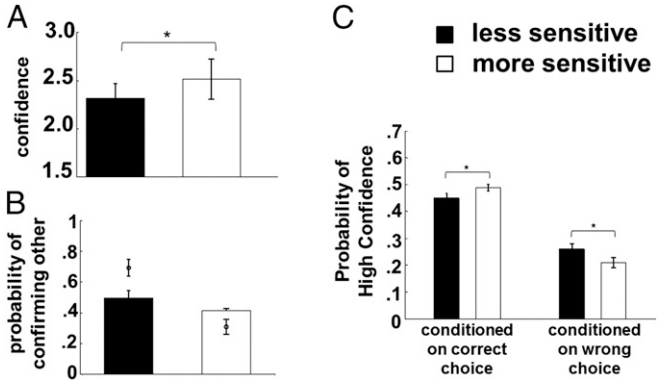

Fig. 2. (*A*) Average confidence. Black bars, less sensitive dyad member; white bars, more sensitive dyad member. (*B*) Probability of confirming partner's decision. Circles indicate the recommendation of the optimal model. (*C*) Probability of high confidence given a correct (left bars) and a wrong (right bars) decision. *$P < 0.05$. All error bars are 1 SEM.

sensitive) × 2 (accuracy: wrong vs. correct decision) mixed ANOVA (dependent variable: P_{high}) showed an expected significant main effect of accuracy [Fig. 2*C*; $F(1, 92) = 491.2$, $P < 0.0001$] but, importantly, no main effect of dyad member [Fig. 2*C*; $F(1, 92) = 0.001$, $P > 0.96$]. However, there was a significant interaction between dyad member and accuracy [Fig. 2*C*; $F(1, 92) = 18.31$, $P < 0.0001$]. In particular, the less sensitive dyad members were more likely (compared with their partner) to report high confidence in their incorrect decisions [Fig. 2*C*; $t(46) = -2.15$, $P < 0.04$, paired t test]. Conversely, the more sensitive members were more likely (compared with their partner) to report high confidence in their correct decision [Fig. 2*C*; $t(46) = 2.39$, $P < 0.03$, paired t test]. The interaction suggests that the less sensitive dyad members were more likely (than their partner) to lead the group astray by expressing high confidence in their errors; in contrast, the more sensitive dyad members were more likely, compared to their partner, to lead the group to the right answer by expressing high confidence when they were correct (see *SI Methods* for details).

To directly estimate how dyad members weighted each other's opinions, we adopted a recent computational model (21). When resolving a disagreement, the arbitrator would have to compare their own decision and confidence with those of their partner. Importantly, to make the most accurate judgments, the weight assigned to each member should be informed by the reliability of their individual responses (27, 28). To learn this information, the arbitrator must dynamically track the history of trial outcomes (29). The model solves this task using Bayesian reinforcement learning (21) (*SI Methods*). In particular, it learns the probability that a given dyad member has made a correct decision on the current trial given their accuracy on previous trials (Fig. 1*B*). On the first trial, the prediction of probability correct is set to 50%. The model then, trial by trial, updates this prediction according to the discrepancy between the prediction of probability correct and the observed accuracy (i.e., the prediction error). If this prediction error is negative (i.e., the participant was wrong), the prediction of probability correct is decreased, but if this prediction error is positive (i.e., the participant was correct), then the prediction of probability correct is increased. Critically, the magnitude of this update (i.e., the learning rate) varies according to the volatility of each dyad member's performance. Possible sources of volatility are lapses of attention, changes of arousal level, and/or perceptual learning. If performance is stable, then the trial-by-trial update is small (small learning rate), but if performance is volatile, then the trial-by-trial update is large (high learning rate).

In line with previous studies (30, 31), we assumed that the arbitrator, on a given trial, based their decision on their own predicted probability correct, $p_{s,i}$ and their partner's predicted probability correct, $p_{o,i}$, where s denotes self, o denotes other, and i denotes the trial number (Fig. 1*B*). Further, we hypothesized that the arbitrator scaled these predictions by the expressed levels of confidence, $p_{s,i} \times c_{s,i}$ and $p_{o,i} \times c_{o,i}$, where the sign of c indicates the participant's individual choice: negative if a participant chose the first interval and positive for the second interval. Last, we introduced a free parameter, γ, which captured the influence of the partner in the decision process. This gave us the following decision criterion (DC)

$$DC = p_{s,i} \times c_{s,i} + \gamma \times p_{o,i} \times c_{o,i}, \quad [1]$$

where the arbitrator would respond first interval if $DC < 0$ and second interval if $DC > 0$ (*SI Methods*). As such, $\gamma > 1$ indicates that the arbitrator is strongly influenced by their partner, whereas $\gamma < 1$ indicates that the arbitrator is weakly influenced by their partner. We estimated this social influence (γ) under two constraints: (*i*) the parameter value that maximized the fit between the model-derived decisions and the empirically observed decisions (γ_{fit}) and (*ii*) the parameter value that maximized the accuracy of the model-derived decisions (γ_{opt}). Critically, by comparing γ_{fit} with γ_{opt} for each participant, we could judge whether the participant overweighted ($\gamma_{fit} > \gamma_{opt}$) or underweighted ($\gamma_{fit} < \gamma_{opt}$) their partner's opinion relative to their own.

We first tested if the less and the more sensitive dyad members differed in how they weighted their partner's opinion relative to their own (γ_{fit} smaller or larger than γ_{opt}). Interestingly, the less sensitive dyad members (s_{min}) underweighted (i.e., $\gamma_{fit} < \gamma_{opt}$) the opinion of their better-performing partner [Fig. 3*A*, black bars; $t(46) = -5.78$, $P < 10^{-6}$, paired t test]. In contrast, the more sensitive dyad members (s_{max}) overweighted (i.e., $\gamma_{fit} > \gamma_{opt}$) the opinion of their poorer-performing partner [Fig. 3*A*, white bars; $t(46) = 2.27$, $P < 0.03$, paired t test]. Critically, the optimal model (γ_{opt}) allowed us to calculate the proportion of trials in which the arbitrator should have taken their partner's advice. Echoing the previous results, the less sensitive dyad members took their partner's advice significantly less often than they should have [compare the black bar (empirical data) and the circle above it (optimal model) in Fig. 2*B*; $t(46) = -4.96$, $P < 10^{-5}$, paired t test]. In contrast, the more sensitive dyad members took their partner's advice significantly more often than they should have [compare white bar and corresponding circle in Fig. 2*B*; $t(46) = -3.66$, $P < 0.007$, paired t test]. We return to the implication of these findings for egocentric advice discounting in *Discussion*.

We then tested if the less and the more sensitive dyad members differed in terms of how effectively they weighted their partner's opinion relative to their own (the absolute difference between the optimal and the empirical weight, $|\gamma_{fit} - \gamma_{opt}|$). Interestingly, the less sensitive dyad members were significantly worse at weighting their partner's opinion [$|\gamma_{fit} - \gamma_{opt}|$ larger for s_{min} vs. s_{max}; $t(46) = 4.58$, $P < 10^{-4}$, paired t test]. This result is in line with an earlier finding (15) that misjudgments of one's own competence is more severe in individuals with low competence than those with high competence (the Dunning–Kruger effect). However, our results go beyond those earlier findings by probing people's implicit beliefs about relative competence as revealed through their decisions rather than a direct, one-shot report. Moreover, we found that, across participants, the absolute difference between the optimal and the empirical weight was negatively correlated with individual sensitivity (Fig. 3*B*; $r = -0.29$, $P < 0.004$, Pearson's correlation). In other words, the more sensitive dyad members were better at judging their own competence relative to their partner.

The above results can be described as an equality bias for social influence—that is, participants appeared to arbitrate the disagreement trials as if they were as good as or as bad as their partner. In the context of our model, such an equality bias would entail that the empirical weights, γ_{fit}, were clustered around 1, whereas the optimal weights, γ_{opt}, were scattered away from 1. Indeed, we found that the empirical weights were significantly closer to 1 [i.e., $|1 - \gamma_{fit}| < |1 - \gamma_{opt}|$; $t(93) = -4.19$, $P < 10^{-5}$, paired t test]. Critically, an equality bias should be useful when true. More precisely, if dyad members are indeed equally sensitive, then the empirical weights and the optimal

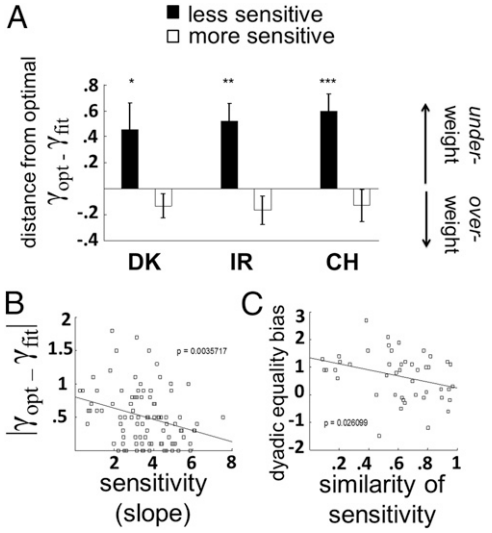

Fig. 3. (*A*) Deviation from optimal performance. Data plotted separately for each country. Positive and negative values indicate over- and underweighting, respectively. Black bars, less sensitive dyad member; white bars, more sensitive dyad member. *$P < 0.05$; **$P < 0.003$; ***$P < 0.001$. Error bars are 1 SEM. (*B*) Absolute deviation from optimal weighting, plotted against sensitivity (slope of psychometric function). Each data point represents one participant. (*C*) Combined equality bias is plotted against similarity of sensitivity. Each data point represents a dyad. For *B* and *C*, data are collapsed across cultures, and the lines show the least-square linear regression.

weight should be close to 1. However, for dyad members with different levels of sensitivity, the empirical and the optimal weight should diverge dramatically as equality bias would be counterproductive. This hypothesis entails that the similarity of dyad members' sensitivity, s_{min}/s_{max}, should be negatively correlated with combined equality bias (i.e., the summed equality bias of both dyad members), which we defined as $(\gamma_{opt,smin} - \gamma_{fit,smin}) - (\gamma_{opt,smax} - \gamma_{fit,smax})$. More positive values indicate a higher equality bias (see *SI Methods* for details). Indeed, this predication was borne out by the data (Fig. 2*C*; $r = -0.32$, $P < 0.03$, Pearson's correlation).

It is important to note that only the positive values of $\gamma_{opt} - \gamma_{fit}$ indicate egocentric advice discounting. Indeed, when collapsed across all participants, $\gamma_{opt} - \gamma_{fit}$ was significantly positive [one-sample *t* test comparing $\gamma_{opt} - \gamma_{fit}$ (mean ± SD = 0.19 ± 0.63) to zero; $t(93) = 2.95$, $P < 0.004$]. Had we followed the convention established by previous studies on advice taking and investigated the overall average (rather than splitting s_{min} and s_{max}) of $\gamma_{opt} - \gamma_{fit}$, we might have interpreted our results purely in terms of egocentric advice discounting, whereas our results described above indicate that both types of advice discounting, egocentric and self-discarding, may coexist in social interaction.

Experiment 2. One critique of the results may be that limitations on memory capacity may have made it too difficult for participants to track the history of trial outcomes ($p_{s,i}$ and $p_{o,i}$), leading to their inability to decide whose opinion is more accurate. We tested this hypothesis in a second study in which participants performed the same task as in experiment 1, except that they were informed about their own and their partner's cumulative (as well as trial-by-trial) accuracy at the end of each trial. A total of 22 healthy adult males (11 dyads in Iran) were recruited for experiment 2 (see *SI Methods* for details). Despite the change in feedback, the results were similar to those obtained in experiment 1. In particular, the less sensitive dyad members underweighted (i.e., $\gamma_{fit} < \gamma_{opt}$) the opinion of their better-performing partner [Fig. S1, black bar; $t(10) = -4.69$, $P < .001$, paired *t* test]. In contrast, the more sensitive dyad members overweighted (i.e., $\gamma_{fit} > \gamma_{opt}$) the opinion of their poorer-performing partner [Fig. S1, white bar; $t(10) = 3.73$, $P < 0.004$, paired *t* test]. With cumulative feedback, holding previous trial outcomes in memory was not necessary. However, dyad members still exhibited an equality bias, ruling out limitations on memory capacity as an explanation of the above results.

Experiment 3. Another possible interpretation of our results may be that dyad members exhibited an equality bias because their individual differences in sensitivity were, at least in some cases, relatively small or accidental. We tested this hypothesis in experiment 3 in which participants performed the same task as in experiment 1, except that we now covertly made the task much harder for one of the two participants. A total of 46 healthy adult males (13 dyads in Denmark and 10 dyads in Iran) were recruited for experiment 3 (see *SI Methods* for details). Briefly, we first ran a practice session (128 trials), without any social interaction (joint decisions and joint feedback), in which we estimated the sensitivity of each participant. We then ran the experimental session (128 trials), with social interaction as in the previous experiments, in which we covertly introduced noise to the visual stimulus presented to the less sensitive member of each dyad. Critically, this manipulation ensured that dyad members' performance differed significantly [Fig. S2; s_{min} vs. s_{max} $t(22) = 8.85$, $P < 10^{-7}$, paired *t* test]. Despite the conspicuous difference in performance, dyad members still exhibited an equality bias. The less sensitive dyad members underweighted (i.e., $\gamma_{fit} < \gamma_{opt}$) the opinion of their better-performing partners [Fig. S3, black bars; $t(22) = -3.38$, $P < 0.003$, paired *t* test]. In contrast, the more sensitive dyad members overweighted (i.e., $\gamma_{fit} > \gamma_{opt}$) the opinion of their poorer-performing partners [Fig. S3, white bars; $t(22) = 2.11$, $P < 0.05$, paired *t* test].

Experiment 4. Finally, we tested whether monetary incentives would reduce—or even eradicate—the equality bias. We conducted a version of experiment 1 in which participants first bet on their choice individually and then collectively, with the payoff (correct: reward; error: cost) determined by the sum of bets (see *SI Methods* for details). Replicating our previous results, the less sensitive dyad members underweighted (i.e., $\gamma_{fit} < \gamma_{opt}$) their partner's opinion [Fig. S4, black bar; $t(16) = -3.91$, $P < 0.002$, paired *t* test]. Conversely, the more sensitive dyad members overweighted (i.e., $\gamma_{fit} > \gamma_{opt}$) their partners [Fig. S4, white bar; $t(16) = 5.04$, $P < 0.001$, paired *t* test]. The addition of monetary incentives did not affect the equality bias.

Discussion

An important challenge for group decision-making (e.g., jury voting, managerial boards, medical diagnosis teams) is to take into account individual differences in competence among group members. Although theory tells us that the opinion of each group member should be weighted by its reliability (11, 32), empirical research cautions that this is easier said than done. To start with, we tend to grossly misestimate our own competence—not only when judged in isolation but also relative to others (14–16). This raises the question of whether—and to what extent—people take into account individual differences in competence when they engage in collective decision-making.

To address this question, we developed a computational framework inspired by previous work on how people learn the reliability of social advice (21). We quantified how members of a group weighted each other's opinion in the context of a visual perceptual task. Research on advice taking (18, 19, 33) predicts that group members (irrespective of their relative competence) would insist on their own opinion more often than following their partner. Looking only at the raw behavioral results (Fig. 2*B* and Fig. S5, bars), one may have concluded that only the better-performing members of each group displayed egocentric discounting and insisted on following their own opinion more often than they followed that of their partner. However, the model-based analysis (Fig. 2*B*, circles) showed that the better-performing group members should (optimally) have followed their partner's opinion even less often than they actually did (compare white bar and circle). Conversely, their poorer-performing counterparts should have followed their partner's opinion even more often than they actually did (compare black bar and circle). Thus, thinking of advice taking as monolithically egocentric may miss out the nuances due to individual differences in

competence arising in social interaction. Importantly, this divergence between the conclusions from the raw behavioral data and the model-based analysis underscores the power of the latter approach for understanding social behavior.

Our participants exhibited an equality bias—behaving as if they were as good or as bad as their partner. We excluded a number of alternative explanations. We replicated the equality bias in experiment 2 in which participants were presented with a running score of each other's performance. Therefore, it is unlikely that the bias reflects a previously reported self-enhancement or self-superiority bias (34) as these were (at least for the less sensitive group members) explicitly contradicted by the running score. In addition, experiment 2 ruled out poor memory for partner's recent performance as an alternative explanation. Moreover, in experiment 3, we replicated the equality bias when participants' performance was (covertly) manipulated to differ conspicuously from that of their partner. It is therefore unlikely that our findings are an artifact of small, accidental differences in performance (i.e., due to noise and not competence per se) as this manipulation ensured that all differences in competence were large. Additionally, the results of experiments 2 and 3 also rule out ignorance of competence (15) as an explanation of the bias. We note that our results cannot be explained as a statistical artifact of regression to the main (see *SI Methods* for detailed discussion of this issue). Finally, we observed the equality bias across three cultures (Denmark, Iran, and China) that differ widely in terms of their norms for interpersonal trust (see Introduction), suggesting that this bias may reflect a general strategy for collective decision-making rather than a culturally specific norm.

So, why do people display an equality bias? People may try to diffuse the responsibility for difficult group decisions (35). In this view, participants are more likely to alternate between their own opinion and that of their partner when the decision is particularly difficult (high uncertainty), thus sharing the responsibility for potential errors. To test the shared responsibility hypothesis directly, one could vary the magnitude of the outcome of the group decisions and quantify the equality bias when (*i*) the reward is small and the cost of an error is high vs. (*ii*) the reward is large and the cost is low. The shared responsibility hypothesis would predict a stronger equality bias under condition *i*. Alternatively, the equality bias may arise from group members' aversion to social exclusion (36). Social exclusion invokes strong aversive emotions that—some have even argued—may resemble actual physical pain (37). By confirming themselves more often than they should have, the inferior member of each dyad may have tried to stay relevant and socially included. Conversely, the better performing member may have been trying to avoid ignoring their partner. Finally, the equality bias may arise because people "normalize" their partner's confidence to their own. More precisely, although the better-performing members of each group were more confident (Fig. 2*A* and Fig. S6), they also overweighted the opinion of their respective partners and vice versa. This suggests that participants may have rescaled their partners' confidence to their own and made the joint decisions on the basis of these rescaled values.

It is also constructive to consider what cognitive purpose equality bias might serve. Cognitive biases often serve as shortcut solutions to complex problems (38). When group members are similar in terms of their competence, the equality bias simplifies the decision process (and lowers cognitive load) by reducing it to a direct comparison of confidence. Similarly, equality bias may facilitate social coordination, with participants trying to get the task done with minimal effort but inadvertently at the expense of their joint accuracy. Indeed, it has been argued (38) that interacting agents quickly converge on social norms (i.e., coordination devices) to reduce disharmony and chaos in joint tasks. Lastly, a wealth of research on "homophily" in human social networks (39) suggests that our tendency for associating and bonding with similar others may exploit the benefits of an equality bias. However, when a wide competence gap separates group members, the normative strategy requires that each opinion is weighted by its reliability (28). In such situations, an equality bias can be damaging for the group. Indeed, previous research has shown that group performance in the task described here depends critically on how similar group members are in terms of their competence (6, 8, 9). Future research could refine our understanding of this bias by assessing the actual real-life probability of the breakdown of the similarity assumption.

Previous research suggests that mixed-sex dyads are particularly likely to elicit task-irrelevant sex-stereotypical behavior (40, 41). We therefore chose to test male participants only. Our choice leaves open the possibility that such an equality bias may be different in female-only or female-male dyads. However, previous studies (42, 43) have shown that women follow an equality norm more often than men when allocating rewards, suggesting that the equality bias may be found for female dyads too. Future studies can use our framework to explore the equality bias in male-female and female-female dyads.

An important question arising from our approach is the extent to which the model parameters are biologically interpretable at neural level. A number of recent studies have identified the neural correlates of confidence in perceptual and economic decision-making (44, 45). Importantly, these studies focused on single participants, leaving open the possibility that confidence as communicated in social interaction may have a different set of neural correlates, or, at least, involve other stages of processing (46). Indeed, it has been suggested that limitations on the processes underlying communicated confidence may explain why groups sometimes fail to outperform their members (8). Future research should test whether the neural substrates that give rise to confidence for communication differ from when confidence guides individual behavior (wagering, slowing down, etc.), and, if so, at what stages of processing. Future research could also identify the neural correlates of the arbitration process that ensues once confidence has been expressed.

In the early years of the 20th century, Marcel Proust, a sick man in bed but armed with a keen observer's vision, wrote "Inexactitude, incompetence do not modify their assurance, quite the contrary" (47). Indeed, our results show that, when making decisions together we do not seem to take into account each other's inexactitudes. However, are people able to learn how they should weight each other's opinions or are they, as implied by Monsieur Proust's melancholic tone, forever trapped in their incompetence?[†]

Methods

In four experiments, pairs of participants made individual and joint decisions in a perceptual task (6). We implemented a visual search task with a two-interval forced choice design. Discrimination sensitivity for individuals and dyads was assessed using the method of constant stimuli. Social influence was assessed by using and adapting a previously published computational model of social learning (30). See *SI Methods* for details.

ACKNOWLEDGMENTS. We thank Tim Behrens for sharing the code for implementing the computational model and Christos Sideras for helping collect the data for experiment 4. This work was supported by European Research Council Starting Grant "NeuroCoDec #309865" (to B.B.), the British Academy (B.B.), the Calleva Research Centre for Evolution and Human Sciences (D.B.), the European Union Mind Bridge Project (D.B., K.O., A.R., C.D.F., and B.B.), and the Wellcome Trust (G.R.).

[†]We did address this question by splitting our data (experiment 1) into two sessions to test whether participants moved closer toward the optimal weight over time. However, we found no statistically reliable difference between the two sessions. This could be due to participants' stationary behavior or that our data and analysis did not have sufficient power to address the issue of learning.

1. Condorcet M (1785) *Essai sur l'application de l'analyse á la probabilité des décisions rendues á la pluralité des voix* (de l'Impr. Royale, Paris, France).
2. Galton F (1907) Vox populi. *Nature* 75(1949):450–451.
3. Surowiecki J (2005) *The Wisdom of Crowds* (Doubleday, Anchor, New York).
4. Galton F (1907) One vote, one value. *Nature* 75(1948):414–414.
5. Perry-Coste FH (1907) The ballot-box. *Nature* 75(1952):509–509.
6. Bahrami B, et al. (2010) Optimally interacting minds. *Science* 329(5995):1081–1085.
7. Koriat A (2012) When are two heads better than one and why? *Science* 336(6079):360–362.
8. Bahrami B, et al. (2012) What failure in collective decision-making tells us about metacognition. *Philos Trans R Soc Lond B Biol Sci* 367(1594):1350–1365.

9. Bang D, et al. (2014) Does interaction matter? Testing whether a confidence heuristic can replace interaction in collective decision-making. *Conscious Cogn* 26:13–23.
10. Shakespeare W (1993) *As You Like It* (Wordsworth Editions, London).
11. Nitzan S, Paroush J (1985) *Collective Decision Making: An Economic Outlook* (Cambridge Univ Press, Cambridge, UK).
12. Gigerenzer G, Hoffrage U, Kleinbölting H (1991) Probabilistic mental models: A Brunswikian theory of confidence. *Psychol Rev* 98(4):506–528.
13. Soll JB (1996) Determinants of overconfidence and miscalibration: The roles of random error and ecological structure. *Organ Behav Hum Dec* 65(2):117–137.
14. Falchikov N, Boud D (1989) Student self-assessment in higher-education: A meta-analysis. *Rev Educ Res* 59(4):395–430.
15. Kruger J, Dunning D (1999) Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *J Pers Soc Psychol* 77(6):1121–1134.
16. Metcalfe J (1998) Cognitive optimism: Self-deception or memory-based processing heuristics? *Pers Soc Psychol Rev* 2(2):100–110.
17. Yaniv I, Kleinberger E (2000) Advice taking in decision making. *Organ Behav Hum Decis Process* 83(2):260–281.
18. Yaniv I (2004) Receiving other people's advice: Influence and benefit. *Organ Behav Hum Dec* 93(1):1–13.
19. Yaniv I, Kleinberger E (2000) Advice taking in decision making: Egocentric discounting and reputation formation. *Organ Behav Hum Dec* 83(2):260–281.
20. Behrens TE, Hunt LT, Rushworth MF (2009) The computation of social behavior. *Science* 324(5931):1160–1164.
21. Behrens TE, Hunt LT, Woolrich MW, Rushworth MF (2008) Associative learning of social value. *Nature* 456(7219):245–249.
22. Roepstorff A, Niewöhner J, Beck S (2010) Enculturing brains through patterned practices. *Neural Netw* 23(8-9):1051–1059.
23. Henrich J, Heine SJ, Norenzayan A (2010) Most people are not WEIRD. *Nature* 466(7302):29.
24. Henrich J, Heine SJ, Norenzayan A (2010) The weirdest people in the world? *Behav Brain Sci* 33(2-3):61–83, discussion 83–135.
25. Gächter S, Herrmann B, Thöni C (2010) Culture and cooperation. *Philos Trans R Soc Lond B Biol Sci* 365(1553):2651–2661.
26. Herrmann B, Thöni C, Gächter S (2008) Antisocial punishment across societies. *Science* 319(5868):1362–1367.
27. Mannes AE, Soll JB, Larrick RP (2014) The wisdom of select crowds. *J Pers Soc Psychol* 107(2):276–299.
28. Nitzan S, Paroush J (1982) Optimal decision rules in uncertain dichotomous choice situations. *Int Econ Rev* 23(2):289–297.
29. Daw ND, Doya K (2006) The computational neurobiology of learning and reward. *Curr Opin Neurobiol* 16(2):199–204.
30. Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10(9):1214–1221.
31. Yu AJ, Dayan P (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46(4):681–692.
32. Bovens L, Hartmann S (2004) *Bayesian Epistemology* (Oxford Univ Press, Oxford, UK).
33. Bonaccio S, Dalal RS (2006) Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organ Behav Hum Dec* 101(2):127–151.
34. Hoorens V (1993) Self-enhancement and superiority biases in social comparison. *Eur Rev Soc Psychol* 4(1):113–139.
35. Harvey N, Fischer I (1997) Taking advice: Accepting help, improving judgment, and sharing responsibility. *Organ Behav Hum Dec* 70(2):117–133.
36. Williams KD (2007) Ostracism. *Annu Rev Psychol* 58:425–452.
37. Eisenberger NI, Lieberman MD, Williams KD (2003) Does rejection hurt? An FMRI study of social exclusion. *Science* 302(5643):290–292.
38. Schelling TC (1980) *The Strategy of Conflict* (Harvard Univ Press, Cambridge, MA).
39. McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: Homophily in social networks. *Annu Rev Sociol* 27:415–444.
40. Buchan NR, Croson RTA, Solnick S (2008) Trust and gender: An examination of behavior and beliefs in the Investment Game. *J Econ Behav Organ* 68(3-4):466–476.
41. Diaconescu AO, et al. (2014) Inferring on the intentions of others by hierarchical Bayesian learning. *PLOS Comput Biol* 10(9):e1003810.
42. Kahn A, Krulewitz JE, O'Leary VE, Lamm H (1980) Equity and equality: Male and female means to a just end. *Basic Appl Soc Psych* 1(2):173–197.
43. Mikula G (1974) Nationality, performance, and sex as determinants of reward allocation. *J Pers Soc Psychol* 29(4):435–440.
44. De Martino B, Fleming SM, Garrett N, Dolan RJ (2013) Confidence in value-based choice. *Nat Neurosci* 16(1):105–110.
45. Fleming SM, Huijgen J, Dolan RJ (2012) Prefrontal contributions to metacognition in perceptual decision making. *J Neurosci* 32(18):6117–6125.
46. Shea N, et al. (2014) Supra-personal cognitive control and metacognition. *Trends Cogn Sci* 18(4):186–193.
47. Proust M (2006) *Remembrance of Things Past* (Wordsworth Editions Ltd., Ware, UK).