Social learning strategies regulate the wisdom and madness of interactive online crowds

Wataru Toyokawa[1,2,3,†], Andrew Whalen[4], and Kevin N. Laland[1]

[1] *School of Biology, University of St Andrews, Harold Mitchel Building, St Andrews, Fife, KY16 9TH, Scotland*

[2] *Japan Society for the Promotion of Science, Kojimachi, Chiyoda-ku, Tokyo, 102-0083, Japan*

[3] *Department of Evolutionary Studies of Biosystems, School of Advanced Sciences, SOKENDAI (The Graduate University for Advanced Studies), Shonan Village, Hayama, Kanagawa 240-0193, Japan*

[4] *The Roslin Institute, University of Edinburgh, Midlothian, SCT, EH25 9RG, Scotland*

[†] *e-mail address: wt25@st-andrews.ac.uk*

Abstract

Decentralised social interactions can generate swarm intelligence, but may concurrently increase the risk of maladaptive herding. Here we present an individual-based model analysis suggesting that the conflict between the 'wisdom' and 'madness' of interactive crowds is regulated by selectively choosing which social learning strategy to use. We used an interactive online experiment with 699 participants to measure the patterns of human social-information use, varying both task uncertainty and group size. Hierarchical Bayesian analyses identified the individual learning strategies, revealing that conformity bias increased with the task's uncertainty, whereas reliance on social learning increased with group size. Mapping individual strategies onto collective behaviour, we show that maladaptive herding occurred more frequently when larger groups were engaged in more uncertain tasks. Our computational modelling approach provides novel evidence that the likelihood of swarm intelligence versus herding can be predicted using knowledge of social learning strategies. (currently 144 words)

Keywords: swarm intelligence, herding, social learning, computational modelling, web-based experiment, hierarchical Bayesian approach

1 Introduction

Understanding the mechanisms that account for accurate collective decision-making amongst groups of animals has been a central focus of animal behaviour research (Bonabeau et al., 1999; Camazine et al., 2001; Sumpter, 2010). There are a large number of biological examples showing that collectives of poorly informed individuals can achieve a high performance in solving cognitive problems under uncertainty (Krause et al., 2010). Examples of such 'swarm intelligence' – the emergent wisdom of interactive crowds – have been found in a broad range of biological systems (Table 1). Although these findings suggest fundamental cognitive benefits of grouping (Krause and Ruxton, 2002), there is also a long-standing recognition, especially for humans, that interacting individuals may sometimes be overwhelmed by the '*extraordinary popular delusions and madness of crowds*' (Mackay, 1841). Herd behaviour (i.e. an alignment of thoughts or behaviours of individuals in a group) occurs because individuals imitate each others (Kameda and Hastie, 2015; Le Bon, 1896; Raafat et al., 2009), and it is thought to be a cause of financial bubbles (Chari and Kehoe, 2004; Mackay, 1841), 'groupthink' (Janis, 1972) and volatility in cultural markets (Muchnik et al., 2013; Salganik et al., 2006). More generally, herding is known to undermine the wisdom of crowds effect (Lorenz et al., 2011), whilst maladaptive aspects of information transfer are well-recognised in the biological literature (e.g. Giraldeau et al., 2002). It seems that information transmission among individuals, and making decisions collectively, is a double-edged sword: combining decisions may provide the benefits of swarm intelligence, but at the same time, increase the risk of maladaptive herding. Collectively, an understanding of whether and, if so, how it is possible to prevent or reduce the risk of maladaptive herd behaviour, while concurrently keeping or enhancing swarm intelligence, is largely lacking.

Table 1

Examples of swarm intelligence in diverse biological systems

Taxonomic families	Examples and references
Slime moulds	Finding conditions favorable to spore survival and dispersal (Reid and Latty, 2016)
Social insects	Collective foraging (Seeley et al., 1991; Shaffer et al., 2013) and nest-site selection (Franks et al., 2003; Sasaki and Pratt, 2012; Sasaki et al., 2013; Seeley and Visscher, 2004)
Fish	Collective sensing (Berdahl et al., 2013; Sumpter et al., 2008), predator avoidance (Ward et al., 2011) and foraging decisions (Webster et al., 2017)
Birds	Collective foraging (Liker and Bokony, 2009; Morand-Ferron and Quinn, 2011) and homing decisions (Sasaki and Biro, 2017)
Non-human primates	Group coordination in where and when to move (King and Sueur, 2011)
Humans	Decision-making in an estimation task (Krause et al., 2011; Rosenberg and Pescetelli, 2017) and in a multi-armed bandit task (Toyokawa et al., 2014)

A balance between using individual and social information may play a key role in determining the trade-off between collective wisdom and maladaptive herding (List et al., 2009). If individuals are too reliant on copying others' behaviour, any ideas, even a maladaptive one, can propagate in the social group (i.e. the 'informational cascade'; Bikhchandani et al., 1992; Giraldeau et al., 2002; Richerson and Boyd, 2005). On the other hand, however, if individuals completely ignore social information so as to be independent, they will fail to exploit the benefits of aggregating information through social interactions. The extent to which individuals should use social information should fall between these two extremes. Theoretical models predict that the balance between independence and interdependence in collective decision-making may be changeable, contingent upon the individual-level flexibility and inter-individual variability associated with the social learning strategies deployed in diverse environmental states (e.g. Arbilly et al., 2011; Boyd and Richerson, 1985; Feldman et al., 1996; Laland, 2004).

Animals (including humans) are reported to increase their use of social information as returns from asocial learning become more unreliable (e.g. Kameda and Nakanishi, 2002; Kendal et al., 2004; Morgan et al., 2012; Toyokawa et al., 2017; Webster and Laland, 2008, 2011). In addition, individuals are predicted to be more likely to rely on social learning larger the number of individuals that share information (Boyd and Richerson, 1989; Bond, 2005; Kline and Boyd, 2010; Morgan et al., 2012; Muthukrishna et al., 2014; Street et al., 2017). This selectivity in the predicted use of social information may have a substantial impact on collective decision-making because only a slight difference in the parameter values of social information use is known to be able to alter qualitatively the collective behavioural dynamics (e.g. Bonabeau et al., 1999; Camazine et al., 2001; Nicolis and Deneubourg, 1999; Pratt and Sumpter, 2006). Therefore, researchers should expect populations to exhibit a higher risk of being trapped with maladaptive behaviour with increasing group size and decreasing reliability of asocial learning (and concomitant increased reliance on social learning).

From the viewpoint of the classic wisdom of crowds theory, increasing group size may increase collective accuracy (List, 2004; King and Cowlishaw, 2007; Wolf et al., 2013; Becker et al., 2017; Laan et al., 2017). The relative advantage of the collective over solitary individuals may also be highlighted by increased task difficulty, because there would be more room in the performance to be improved compared to easier tasks in which high accuracy can already be achieved by asocial learning only (Cronin, 2016). To understand the potential conflict between swarm intelligence and the risk of maladaptive herding requires fine-grained quantitative studies of social learning strategies and their relations to collective dynamics, linked to sophisticated computational analysis.

The aims of this study were twofold. First, we set out to examine whether altering both the reliability of asocial learning and group size would induce heavier use of social information in humans, and thereby alter the balance between swarm intelligence and the risk of maladaptive herding. To do this, we focused on human groups exposed to a simple gambling task, where both asocial and social sources of information were available. Second, we sought to conduct a detailed analysis of the complex relationship between individual-level decision, learning strategies and population-level behavioural outcomes. Our use of an abstract decision-making task allowed us to implement a computational modelling approach, which has been increasingly deployed in quantitative studies of animal social learning strategies (Ahn et al., 2014; Aplin et al., 2017; Barrett et al., 2017; McElreath et al., 2005, 2008; Toyokawa et al., 2017). In particular, computational modelling allowed us to conduct a parametric description of different information-gathering processes and to estimate these parameter values at an individual-level resolution.

Below, we firstly described our experimental task and summarise the computational model. Then, we deploy agent-based simulation to illustrate how the model parameters relating to social learning can in principle affect the collective-level behavioural dynamics. The simulation provides us with precise, quantitative predictions on the complex relationship between individual behaviour and group dynamics. Finally, we present the findings of a multi-player web-based experiment with human participants that utilises the gambling task framework. Applying a hierarchical Bayesian statistical method, we estimated the model's parameters for each of 699 different individuals, allowing us to (*i*) examine whether and, if so, how social information use is affected by different group size and task uncertainty, and (*ii*) whether and how social-information use affects the balance between swarm intelligence and maladaptive herding.

1.1 Task overview

To study the relationship between social information use and collective behavioural dynamics, we focused on a well-established learning-and-decision problem called a 'multi-armed bandit' task, represented here as repeated choices between three slot machines (Figure S1, Video 1, for detail see Materials and methods). Individuals play the task for 70 rounds. The slots paid off money noisily, varying around two different means during the first 40 rounds such that there was one 'good' slot and two other options giving poorer average returns. From the round 41st, however, one of the 'poor' slots abruptly increased its mean payoff to become 'excellent' (i.e. superior to 'good'). The purpose of this environmental change was to observe the effects of maladaptive herding by potentially trapping groups in the out-of-date suboptimal (good) slot, as individuals did not know whether or how an environmental change would occur. Through making choices and earning a reward from each choice, individuals could gradually learn which slot generated the highest rewards.

In addition to this asocial learning, we provided social information for each member of the group specifying the frequency with which group members chose each slot. All group members played the same task with the same conditions simultaneously, and all individuals had been instructed that this was the case, and hence understood that the social information would be informative.

Task uncertainty was experimentally manipulated by changing the difference between the mean payoffs for the slot machines. In the task with the least uncertainty, the distribution of payoffs barely overlapped, whilst in the task with the greatest uncertainty the distribution of payoffs overlapped considerably (Figure S3).

1.2 Overview of the computational learning-and-decision-making model

We modelled individual behavioural processes by assuming that individual i makes a choice for option m at round t, in accordance with the choice-probability $P_{i,t}(m)$ that is a weighted average of social and asocial influences:

$$P_{i,t}(m) = \sigma_{i,t} \times \mathit{Social\ influence}_{i,m,t} + (1 - \sigma_{i,t}) \times \mathit{Asocial\ influence}_{i,m,t}, \tag{1}$$

where $\sigma_{i,t}$ is the *social learning weight* $(0 \leq \sigma_{i,t} \leq 1)$.

For the social influence, we assumed a frequency-dependent copying strategy by which an individual copies others' behaviour in accordance with the distribution of social frequency information (McElreath et al., 2005, 2008; Aplin et al., 2017; Barrett et al., 2017):

$$\mathit{Social\ influence}_{i,m,t} = \frac{\left(\mathit{frequency}_{m,t-1}\right)^{\theta_i}}{\sum_{k \in \mathit{options}} \left(\mathit{frequency}_{k,t-1}\right)^{\theta_i}}, \tag{2}$$

where $frequency_{m,t-1}$ is a number of choices made by other individuals for the option m in the preceding round $t-1$ ($t \geq 2$). The exponent θ_i is individual i's *conformity exponent* ($-\infty \leq \theta_i \leq +\infty$). When this exponent is larger than zero ($\theta_i > 0$), higher social influence is afforded to an option chosen by more individuals (i.e. positive frequency bias), with conformity bias arising when $\theta_i > 1$, such that disproportionally more social influence is given to the most common option (Boyd and Richerson, 1985). When $\theta_i < 0$, on the other hand, higher social influence is afforded to the option that fewest individuals chose in the preceding round $t-1$ (i.e. negative frequency bias). Note, there is no social influence when $\theta_i = 0$ because in this case the 'social influence' favours an uniformly random choice, i.e., $Social\ influence_{i,m,t} = f_m^0/(f_1^0 + f_2^0 + f_3^0) = 1/3$, independent of the social frequency distribution.

For the asocial influence, we used a standard 'softmax' choice rule well-established in the reinforcement-learning literature (Sutton and Barto, 1998) and widely applied in human social learning studies (e.g. McElreath et al., 2005, 2008; Toyokawa et al., 2017).

In summary, the model has two key social learning parameters, the *social learning weight* $\sigma_{i,t}$ and the *conformity exponent* θ_i, with $\sigma_{i,t}$ a time-dependent variable (i.e. individuals could modify their reliance on social learning as the task proceeded). Varying these parameters systematically, we conducted an individual-based simulation so as to establish quantitative predictions concerning the relationship between social information use and collective behaviour. We then fitted this model to our experimental data using a hierarchical Bayesian approach. This method allows us to specify with precision how each individual subject learns (i.e. which learning strategy or strategies they deploy), and thereby to describe the range and distribution of learning strategies deployed across the sample, and to investigate their population-level consequences.

2 Results

2.1 The relationship between social information use and the collective behaviour

Figure 1 shows the relationship between the average decision accuracy and individual-level social information use obtained from our individual-based model simulations. Figure 1a and 1c show that individuals in larger groups perform better both before and after the environmental change when the mean conformity exponent $\bar{\theta}$ is small (i.e. $\bar{\theta} = (\sum_i \theta_i)/individuals = 1$). In the absence of conformity, even when the average social learning weight is very high (i.e. $\bar{\sigma} = (\sum_i \sum_t \sigma_{i,t})/(individuals \times rounds) = 0.9$), larger groups are still able to recover the decision accuracy after the location of the optimal option has been switched.

On the other hand, when the mean conformity exponent is large (i.e. $\bar{\theta} = 3$; strong conformity bias), the group dynamics become less flexible, and become vulnerable to getting stuck on a suboptimal option after environmental change. Here, the recovery of performance after environmental change takes more time in larger compared to smaller groups (Figure 1b). When both the conformity exponent $\bar{\theta}$ and the social learning weight $\bar{\sigma}$ are large (Figure 1d), performance is no longer monotonically improving with increasing group size, and it is under these circumstances that the strong herding effect becomes prominent. Figure 2c and 2d indicate that when both $\bar{\theta}$ and $\bar{\sigma}$ are large the collective choices converged either on the good option or on one of the poor options almost randomly, regardless of the option's quality, and that once individuals start converging on an option the population gets stuck. As a result, the distribution of the groups' average performance over the replications becomes a bimodal 'U-shape'. Interestingly, however, the maladaptive herding effect remains relatively weak in smaller groups (see Figure 2c; the black histograms). This is because the majority of individuals in smaller groups (i.e. two individuals out of three) are more likely to break the cultural inertia by simultaneously exploring for another option than the majority in larger groups (e.g. six out of ten). As expected, herding does not occur in the absence of conformity (Figure 2a, 2b).

In summary, the model simulation suggests an interaction between social learning weight $\bar{\sigma}$ and conformity exponent $\bar{\theta}$ on decision accuracy and the risk of maladaptive herding: When the conformity exponent is not too large, swarm intelligence is prominent across a broad range of the mean social learning weights (i.e. increasing group size can increase decision accuracy while concurrently retaining decision flexibility). When the conformity bias becomes large, however, the risk of maladaptive herding arises, and, when both social learning parameters are large, swarm intelligence is rare and maladaptive herding dominates.

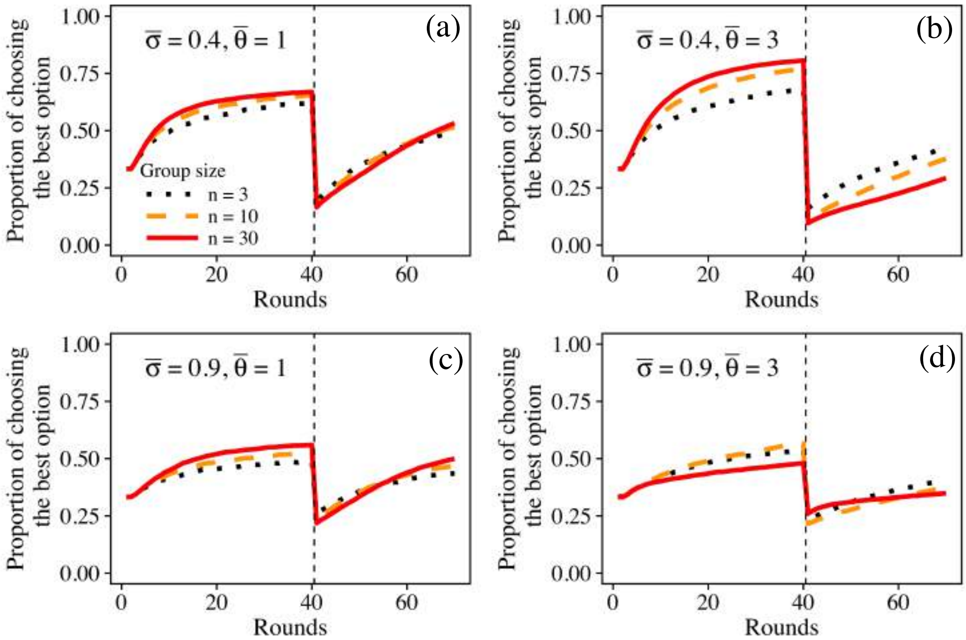

Figure 1: Findings of the individual-based model showing the effects of social information use on the average decision accuracy over replications. The x-axis gives the round and y-axis gives the proportion of individuals expected to choose the optimal slot (i.e. decision accuracy) averaged over all replications. The vertical dashed line indicates the timing of environmental (i.e. payoff) change (at $t = 41$). Different group sizes are shown by different styles (black (dotted): $n = 3$, orange (dashed): $n = 10$, red (solid): $n = 30$). We set the average slopes for the *social learning weight* to be equal to zero for the sake of simplicity; namely, $\mu_\delta = 0$. Other free parameter values (i.e. μ_α, $\mu_{\beta_0^*}$, μ_ϵ,ν_α, $\nu_{\beta_0^*}$, ν_ϵ, ν_σ, ν_δ and ν_θ) are best approximates to the experimental fitted values (see Table 2 and Table S1).

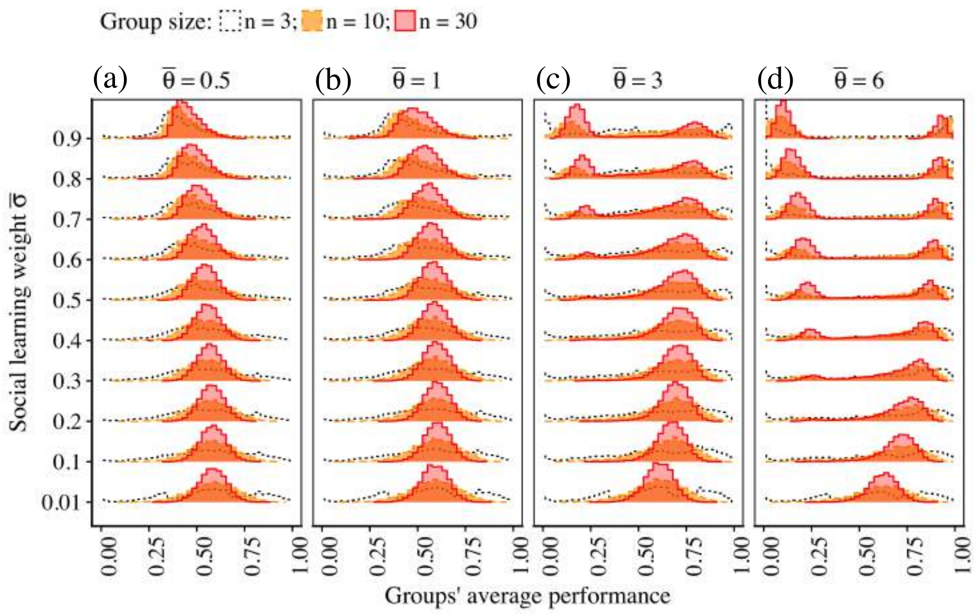

Figure 2: Results from the individual-based model simulations showing the distribution of each group's mean accuracy before environmental change. The x-axis gives the mean decision accuracy over the first 40 rounds (i.e. the environment 1) for each replication. Different group sizes are shown by different styles (black (dotted): $n = 3$, orange (dashed): $n = 10$, red (solid): $n = 30$). Again, $\mu_\delta = 0$, and other free parameter values (i.e. μ_α, $\mu_{\beta_0^*}$, μ_ϵ,ν_α, $\nu_{\beta_0^*}$, ν_ϵ, ν_σ, ν_δ and ν_θ), we approximated using experimental data (see Table 2 and Table S1).

2.2 Estimation of human social information use

Table 2 reveals how the *social learning weight* $\sigma_{i,t}$ and *conformity exponent* θ_i were influenced by task uncertainty in our behavioral experiment. It gives posterior estimation values for each of the global means of the learning model parameters, obtained by the hierarchical Bayesian model fitting method applied to the experimental data (see the Materials and methods). The fitted global variance parameters (i.e. ν) are shown in the Supporting Table S1.

Table 2

The mean and the 95% Bayesian credible intervals of the posterior global means for the parameter values. The number of participants (N) for each experimental condition are also shown.

	Group condition			Solitary condition		
	Uncertainty			Uncertainty		
Parameters	Low	Moderate	High	Low	Moderate	High
μ_{α^*} (learning rate)	0.99	0.90	0.61	0.85	-0.17	0.46
	[0.34, 1.73]	[0.43, 1.44]	[0.21, 1.03]	[-0.07, 1.95]	[-1.27, 0.89]	[-0.39, 1.36]
$\mu_{\beta_0^*}$ (inv. temp.)	1.84	1.68	1.38	1.10	1.44	0.85
	[1.15, 2.70]	[1.25, 2.18]	[1.16, 1.62]	[0.69, 1.54]	[0.80, 2.07]	[0.46, 1.22]
μ_{ϵ} (inv. temp.)	3.70	3.01	2.97	2.39	2.81	2.27
	[1.98, 5.71]	[1.88, 4.27]	[2.37, 3.60]	[1.46, 3.53]	[1.64, 4.07]	[1.40, 3.31]
$\mu_{\sigma_0^*}$ (soc. wight)	-1.55	-2.37	-2.16	–	–	–
	[-2.71, -0.71]	[-4.12, -1.01]	[-2.81, -1.63]	–	–	–
μ_{δ} (soc. wight)	-1.39	-1.55	-1.87	–	–	–
	[-2.66, -0.03]	[-4.29, 0.91]	[-3.04, -0.81]	–	–	–
μ_{θ} (conformity coeff.)	1.65	3.00	2.67	–	–	–
	[0.83, 2.82]	[1.57, 4.85]	[1.80, 3.73]	–	–	–
N	77	98	398	36	34	56

We were able to categorize the participants as deploying three different learning strategies based on their fitted conformity exponent values; namely, the 'positive frequency-dependent copying' strategy ($\theta_i \gg 0$), the 'negative-frequency dependent copying' strategy ($\theta_i \ll 0$) and the 'random choice' strategy ($\theta_i \approx 0$). Note that we could not reliably detect the 'weak positive' frequency-dependent strategy ($0 < \theta_i \leq 1$) due to the limitation of statistical power (Figure S10 and S17). Some individuals whose 'true' conformity exponent fell between zero and one would have been categorised as exhibiting a random choice strategy (Figure S10). Individuals identified as exhibiting a positive frequency-dependent copiers were mainly those whose conformity exponent was larger than one ($\theta_i > 1$).

Figure 3a-c show the estimated frequencies of different learning strategies. Generally speaking, participants were more likely to utilize a positive frequency-dependent copying strategy than the other two strategies (the 95% Bayesian CI of the intercept of the GLMM predicting the probability to use the positive frequency-dependent copying strategy is above zero, [1.05, 2.50]; Table S2). We found that positive frequency-dependent copying decreased with increasing task uncertainty (the 95% Bayesian CI of task uncertainty effect is below zero, [-1.88, -0.25]; Table S2). We found no clear effects of either the group size, age or gender on adoption of the positive frequency-dependent copying strategy, except for the negative interaction effect between age and task uncertainty (the 95% Bayesian CI of the age × uncertainty interaction = [-1.46, -0.15]; Table S2).

We also investigated the effects of group size and task uncertainty on the fitted individual parameter values. We found that the individual mean *social learning weight* parameter (i.e. $\bar{\sigma}_i = (\sum_t \sigma_{i,t})/70$) increased with group size (the 95% Bayesian CI = [0.15, 0.93]; Figure 3d-f;

Table S3), and decreased with uncertainty (the 95% Bayesian CI = [-0.98, -0.22]), and age of subject (the 95% Bayesian CI = [-0.36, -0.02]). However, the negative effects of task uncertainty and age disappeared when we focused only on $\bar{\sigma}_i$ of the positive frequency-dependent copying individuals, and only the positive effect of the group size was confirmed (Table S4; Figure S16). It is worth noting that the meaning of the social learning weight is different between these three different strategies: The social learning weight regulates positive reactions to the majorities' behaviour for positive frequency-dependent copiers, whereas it regulates avoidance of the majority for negative-frequency dependent copiers, and determines the probability of random decision-making for the random choice strategists.

The individual *conformity exponent* parameter θ_i increased with task uncertainty (the 95% Bayesian CI = [0.38, 1.41]), but we found no significant effects of group size, age, gender or interactions (Figure 3g-i; Table S5). These results were qualitatively unchanged when we focused only on the positive frequency-dependent copying individuals (Table S6; Figure S16).

We observed extensive individual variation in social information use. The greater the task's uncertainty, the larger were individual variances in both the mean social learning weight and the conformity exponent (the 95% Bayesian CI of the GLMM's variation parameter for $\bar{\sigma}_i$ was [1.11, 1.62] (Table S3) and for θ_i was [1.07, 1.54] (Table S5)). This was confirmed when focusing only on the positive frequency-dependent copying individuals: The Bayesian 95% CIs were [1.14, 1.80] (Table S4) and [0.71, 1.10] (Table S6), respectively.

The manner in which individual variation in social-information use of positive frequency-dependent copying individuals changes over time is visualised in Figure 4a-c. The social learning weights generally decreased with experimental round. However, some individuals in the

Moderate- and the High-uncertain conditions accelerated rather than decreased their reliance on social learning over time. Interestingly, those accelerating individuals tended to have a larger conformity exponent (Figure S18). In addition, the time-dependent $\theta_{i,t}$ in our alternative model generally increased with experimental round in the Moderate- and the High-uncertainty conditions (see the appendix; Figure S26), although the fitting of $\theta_{i,t}$ in the alternative model was relatively unreliable (Figure S20). These findings suggest that conformists tended to use asocial learning at the outset but increasingly started to conform as the task proceeded.

Extensive variation in the temporal dynamics of the social learning weight $\sigma_{i,t}$ was also found for the negative-frequency dependent copying individuals but not found for the random choice individuals (Figure S14). Individuals deploying a random choice strategy exhibited a $\sigma_{i,t}$ that approached to zero, indicating that their decision-making increasingly relied exclusively on asocial reinforcement learning as the task proceeded.

No significant fixed effects were found in other asocial learning parameters such as the learning rate α_i and the mean inverse temperature $\bar{\beta}_i = (\sum_t \beta_{i,t})/70$ (Table S7, Table S8 and Figure S15).

In summary, our experiments on adult humans revealed asymmetric influences of increasing task uncertainty and increasing group size on the social learning parameters. The conformity exponent increased with task uncertainty on average but the proportion of positive frequency-dependent copying individuals showed a corresponding decrease, due to the extensive individual variation emerging in the High-uncertain condition. Conversely, group size had a positive effect on the mean social learning weight, but did not affect conformity (Figure 3, 4a-c).

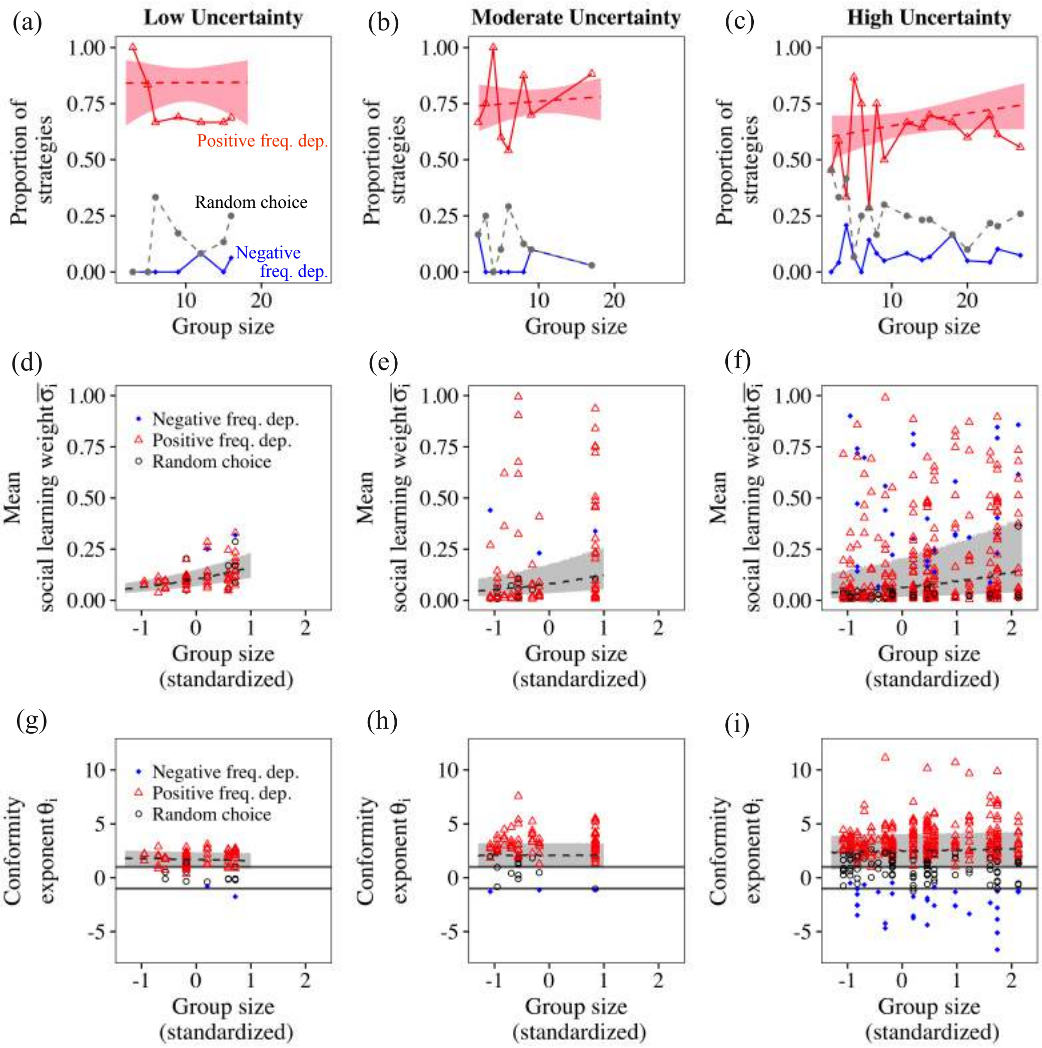

Figure 3: Model fitting for the three different task's uncertain conditions (the Low-, Moderate- and High-uncertainty) and the different group size. Three different learning strategies are shown in different styles (red-triangle: positive frequency-dependent learning, blue-circle: negative frequency-dependent learning; grey-circle: nearly random choice strategy). (a-c) Frequencies of three different learning strategies. Note that a sum of the frequencies of these three strategies in the same group size does not necessarily equal to 1, because there are a small number of individuals eliminated from this analysis due to insufficient data. (d-f) Estimated social learning weight, and (g-i) estimated conformity exponent, for each individual shown for each learning strategy. The 50% Bayesian CIs of the fitted GLMMs are shown by dashed lines and shaded areas. The horizontal lines in (g-i) show a region $-1 < \theta_i < 1$.

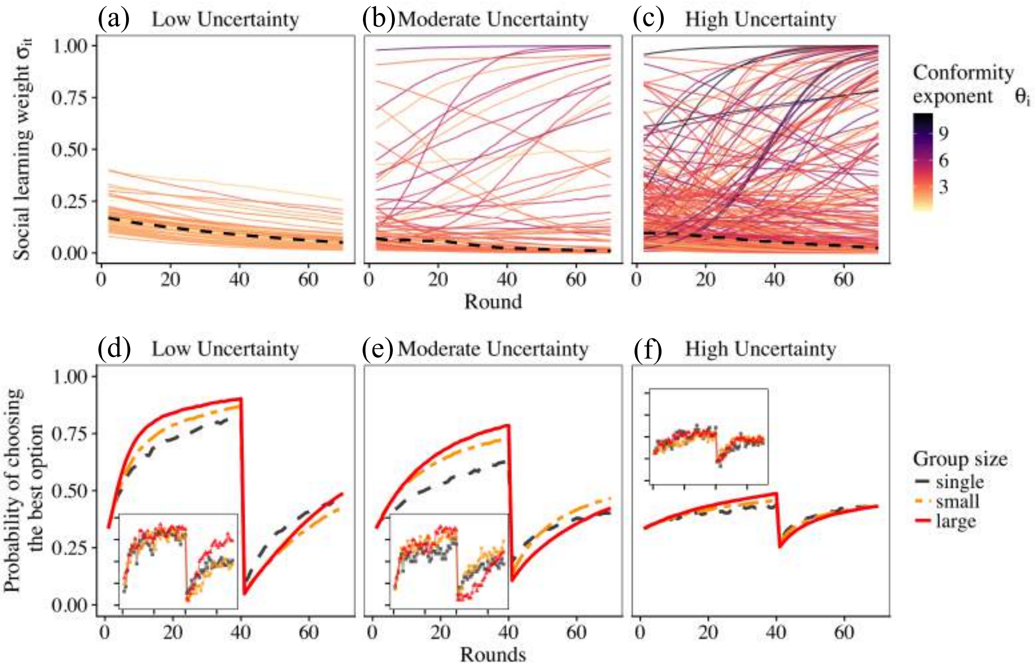

Figure 4: (a-c) Change in fitted values (i.e. median of the Bayesian posterior distribution) of the social learning weight $\sigma_{i,t}$ with time for each individual, for each level of task uncertainty. Thick dashed lines are the median values of $\sigma_{i,t}$ across the subjects for each uncertainty condition. Individual conformity exponent values θ_i are shown in different colours (higher θ_i is darker). (d-f) Change in average decision accuracy of the individual-based post-hoc model simulations using the experimentally fit parameter values (main panels). The inner panels show the average decision accuracies of the experimental participants. Each line indicates different group-size categories (red-solid: large groups, orange-halfdashed: small groups, grey-dashed: lone individuals). All individual performances were averaged within the same size category. The large or small groups were categorised using the median sizes for each experimental condition, i.e. small groups were: $n \leq 9$, $n \leq 6$ and $n \leq 11$ for the Low-, Moderate- and High-uncertain conditions, respectively.

2.3 A balance between the collective decision accuracy and the herding effect

Figure 4d-f show the change over time in performance with different group sizes and different uncertainty conditions, generated by the post-hoc simulations of the parameter-fitted model. The mean decision accuracies of the experimental groups are shown in the inner windows. Because the post-hoc simulations were run for 5,000 replications for each group size, which should generate more robust pattern than the raw experimental data basing only on a limited number of experimental replications, and given the correspondence between simulations and data, below we concentrate our interpretation on the simulated results.

Prior to the environmental change (Round 1 to 40), larger groups performed better on average than did both smaller groups and lone individuals across all the uncertainty levels, suggesting swarm intelligence was operating. However, after the environmental change (i.e. from Round 41) performance differed between the conditions. In the Low-uncertain condition, where we found that the participants were most likely to have a relatively weak positive frequency-dependence (i.e. $\bar{\theta} = 1.65$), large groups again made more accurate decisions than small groups (Figure 4d, from Round 41). However, in the Moderate- and the High-uncertain condition, where we found that participants were most likely to have strong positive frequency dependence ($\bar{\theta} = 3.00$ and 2.67, c.f. 1.65 in the Low-uncertainty condition), the large groups seemed to get stuck on the suboptimal option after the change (Figure 4e and 4f, from Round 41), although the decision accuracy did not substantially differ with group size in the High-uncertain condition.

Lone individuals in the Low-uncertain condition recovered performance more quickly than did both the small and large groups even though the lone individuals performed worse in the first-half of the task (Figure 4d), suggesting that asocial learners are more capable of detecting the environmental change than individuals in groups. This might be due to the higher exploration rate of lone individuals (both $\mu_{\beta_0^*}$ and μ_ϵ of solitary individuals were smaller than those of grouping individuals; Table 2).

Overall, the pattern of results was broadly consistent with our predictions (Figure 1). We confirmed that in the Low-uncertainty condition, where individuals have weaker positive frequency bias, larger groups were more accurate than smaller groups while retaining flexibility in their decision-making (i.e. swarm intelligence dominates). However, in the Moderate- and the High-uncertain conditions, larger groups performed better prior to environmental change but were vulnerable to getting stuck with an out-dated maladaptive option due to the larger estimated conformity exponent, thereby generating the conflict between swarm intelligence and maladaptive herding.

3 Discussion

We investigated whether and how human social learning strategies regulate the conflict between swarm intelligence and herding behaviour using a collective learning-and-decision-making task combined with simulation and model fitting. We examined whether manipulating the reliability of asocial learning and group size would affect the use of social information, and thereby alter the collective decision dynamics, as suggested by our computational model simulation. Although a theoretical study has suggested that reliance on social learning and conformity bias would play a role in collective dynamics (Kandler and Laland, 2013), thus far no empirical studies have quantitatively investigated the population-level consequences of these two different social learning processes. Our high-resolution, model-based behavioural analysis using a hierarchical Bayesian statistics enabled us to identify individual-level patterns and variation of different learning parameters and to explore their population-level outcomes. The results provide strong support for our hypothesis that the conflict between the swarm intelligence effect and maladaptive herding can be predicted with knowledge of human social learning strategies.

Consistent with previous empirical findings (e.g., Morgan et al., 2012; Muthukrishna et al., 2014), adult human participants were increasingly likely to make a conformity-biased choice as the uncertainty of the task went up (i.e. as it became more difficult to determine the best option. Figure 3g-i). The fitted global mean values of the conformity exponent parameters were 3.0 and 2.7 in the Moderate- and the High-uncertain conditions, respectively (Table 2), and these values were sufficiently high to cause larger populations to get stuck on a suboptimal option following environmental change (Figure 1b; Figure 4e, 4f). Conversely, in the Low-uncertain condition individuals exhibited relatively weak conformity (i.e. $\bar{\theta} \approx 1.65$), allowing larger groups to escape the suboptimal option, and retain their swarm intelligence (Figure 1a; Figure 4d). Although the social learning weight was also found to be contingent upon the environmental factors, the estimated mean value was $\bar{\sigma}_i = 0.3$ (Figure 3d-f; Figure S14). This implies a weaker social than asocial influence on decision-making as reported in several other experimental studies (e.g. Efferson et al., 2008; McElreath et al., 2005; Mesoudi, 2011; Toyokawa et al., 2017). Thanks to this relatively weak reliance of social learning, the kind of herding that would have blindly led a group to any option regardless of its quality (like the ‘symmetry breaking’ known in social insect collective foraging systems. Figure 2c,d; Camazine et al., 2001; Sumpter, 2010), did not occur. Research that explores the factors that can induce higher social learning weights in humans, in order to understand under which circumstances herd behaviour would dominate, would be valuable.

Individual differences in exploration might also play a crucial role in shaping collective decision dynamics. Although a majority of participants adopted a positive frequency-dependent copying strategy, some individuals exhibited negative frequency dependent or random decision-making strategy (Figure 3a-c). It is worth noting that the random choice strategy was associated with more exploration than the other strategies, because it led to an almost random choice at a rate σ_i, irrespective of the options' quality. In addition, negative-frequency dependent copying individuals could also be highly exploratory. These individuals tended to avoid choosing an option upon which the other people had converged and would explore the other two 'unpopular' options. Interestingly, in the High-uncertain condition the mean social learning weights of the negative-frequency dependent copying individuals ($\bar{\sigma}_i \approx 0.5$) were larger than that of the other two strategies ($\bar{\sigma}_i \approx 0.1$, Figure S14), indicating that these individuals engaged in such majority-avoiding exploration relatively frequently. Such high exploratory tendencies would prevent individuals from converging on a better option, leading to a diminishing of swarm intelligence in high-uncertainty circumstances (Figure 4f).

Individual differences have received increasing attention in both collective behaviour and animal social learning studies (e.g. Jolles et al., 2018; Michelena et al., 2010; Planas-sitja et al., 2015), and across the human behavioural sciences (e.g. Gray et al., 2017; Mesoudi et al., 2016). Our finding that the effects of individual variation depend on uncertainty implies that human subjects' use of social learning strategies is deployed plastically, and is not a fixed propensity (i.e. personality trait), that differs rigidly between individuals (Dingemanse et al., 2010; Toyokawa et al., 2017). Our approach of combining with individual-based simulation and experimentation could potentially prove a powerful tool with which to explore decision-making in other animals.

Another methodological advantage of using computational models to study social learning strategies is its explicitness of assumptions about the temporal dynamics of behaviour. It has been argued that just observing the final frequencies of learned behaviour does not provide enough information to determine what asocial and/or social learning processes might have been used because multiple learning-and-decision mechanisms are equally likely to produce the same population-level patterns (Barrett, 2018; Hoppitt and Laland, 2013). For example, very exploitative asocial reinforcement learners (i.e. exploitation parameter $\beta_{i,t}$ is large and the social learning weight $\sigma_{i,t}$ is nearly zero) and conformity-biased social learners (conformity exponent θ_i is large and $\sigma_{i,t}$ is positive) would eventually converge on the same option, resulting in the same final behavioural steady state. However, how they explored the environment, as well as how they reacted to the other individuals in the same group, are significantly different and they could produce qualitatively different collective temporal dynamics. A time-depth perspective is crucially important in order to model the relationship between individual behavioural mechanisms and group behavioural dynamics (Biro et al., 2016).

The Internet-based experimentation allowed us to conduct a real-time interactive behavioural task with larger subject pools than a conventional laboratory-based experiment. This enabled us not only to quantify the individual-level learning-and-decision processes (e.g. Ahn et al., 2014; Daw et al., 2006) but also to map these individual-level processes on to the larger-scale collective behaviour (Raafat et al., 2009; Salganik et al., 2006; Sumpter, 2010). Although there are always questions about the validity of participants' behaviour when deploying the web-based method, we believe that the computational modelling approach coupled with higher statistical power due to the large sample size, compensates for any drawbacks. The fact that our learning model could approximate the participants' decision trajectories effectively suggest that most of the participants engaged seriously with solving the task. An increasing body of evidence supports the argument that web-based behavioural experiments are as reliable as results from the laboratory (e.g. Dandurand et al., 2008; Hergueux and Jacquemet, 2015).

The diverse effects of social influence on the collective wisdom of a group has been drawing substantial attention (e.g. Becker et al., 2017; Jayles et al., 2017; Lorenz et al., 2011; Lorge et al., 1958; Muchnik et al., 2013). The bulk of this literature, including many jury models and election models (Hastie and Kameda, 2005; List, 2004), has focused primarily on the static estimation problem, where the 'truth' is fixed from the outset. However, in reality, there are many situations under which the state of the true value is changing over time so that monitoring and tracking the pattern of change is a crucial determinant of decision performance (Payzan-Lenestour and Bossaerts, 2011). In such temporally dynamic environments, decision-making and learning are coordinated to affect future behavioural outcomes recursively (Sutton and Barto, 1998). Our experimental task provides a simple vehicle for exploring collective intelligence in a dynamic situation, which encompasses this learning-and-decision-making feedback loop. Potentially, integrating the wisdom of crowds with social learning and collective dynamics research will facilitate the more tractable use of swarm intelligence in a temporary changing world.

4 Material and methods

4.1 Computational learning-and-decision model

We modelled a learning and decision process based on standard reinforcement-learning theory (Sutton and Barto, 1998). Following previous empirical studies of social learning strategies in humans (e.g. McElreath et al., 2005, 2008; Toyokawa et al., 2017), our model consists of two steps. First, an individual i updates the estimated average reward associated with an option m at round t, namely Q-value ($Q_{i,t}(m)$), according to the Rescorla-Wagner rule (Trimmer et al., 2012) as follows:

$$Q_{i,t+1}(m) = Q_{i,t}(m) + \alpha_i \mathbb{1}(m, m_{i,t})\Big(r_{i,t}(m) - Q_{i,t}(m)\Big), \tag{3}$$

where α_i ($0 \leq \alpha_i \leq 1$) is a *learning rate* parameter of individual i determining the weight given to new experience and $r_{i,t}(m)$ is the amount of monetary reward obtained from choosing the option m in round t. $\mathbb{1}(m, m_{i,t})$ is the binary action-indicator function of individual i, given by

$$\mathbb{1}(m, m_{i,t}) = \begin{cases} 1, & \text{if } m_{i,t} = m \text{ or } t = 1, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Therefore, $Q_{i,t}(m)$ is updated only when the option m was chosen; when the option m was not chosen, $Q_{i,t}(m)$ is not updated (i.e. $Q_{i,t+1}(m) = Q_{i,t}(m)$). Note that, only in the first round $t = 1$, all Q-values are updated by using the chosen option's reward $r_{i,1}(m)$, so that the individual can set a naive 'intuition' about the magnitude of reward values she/he would expect to earn from a choice in the task; namely, $Q_{i,t=2}(1) = Q_{i,t=2}(2) = Q_{i,t=2}(3) = \alpha_i r_{i,t=1}(m)$. In practical terms, this prevents the model from being overly sensitive to the first experience. Before the first choice, individuals had no prior preference for either option (i.e. $Q_{i,1}(1) = Q_{i,1}(2) = Q_{i,1}(3) = 0$).

Second, a choice is made for an option m by individual i at the choice probability $P_{i,t}(m)$ that is determined by a weighted average of social and asocial influences:

$$P_{i,t}(m) = \sigma_{i,t} S_{i,t}(m) + (1 - \sigma_{i,t}) A_{i,t}(m), \tag{5}$$

where $\sigma_{i,t}$ is the *social learning weight* ($0 \le \sigma_{i,t} \le 1$), and $S_{i,t}(m)$ and $A_{i,t}(m)$ are social and asocial influences on the choice probability, respectively ($0 \le S_{i,t}(m) \le 1$ and $0 \le A_{i,t}(m) \le 1$). Note that the sum of choice probabilities, the sum of social influences and the sum of asocial influences are all equal to 1; namely, $\sum_{k \in options} P_{i,t}(k) = 1$, $\sum_k S_{i,t}(k) = 1$ and $\sum_k A_{i,t}(k) = 1$.

As for the asocial influence $A_{i,t}$, we assumed the so-called softmax (or logit choice) function, which is widely used in the reinforcement-learning literature:

$$A_{i,t}(m) = \frac{\exp\left(\beta_{i,t} Q_{i,t}(m)\right)}{\sum_{k \in options} \exp\left(\beta_{i,t} Q_{i,t}(k)\right)}, \tag{6}$$

where $\beta_{i,t}$, called *inverse temperature*, manipulates individual i's sensitivity to the Q-values (in other words, controlling the proneness to explore). As $\beta_{i,t}$ goes to zero, asocial influence approximates to a random choice (i.e. highly explorative). Conversely, if $\beta_{i,t} \to +\infty$, the asocial influence leads to a deterministic choice in favour of the option with the highest Q-value (i.e. highly exploitative). For intermediate values of $\beta_{i,t}$, individual i exhibits a balance between exploration and exploitation (Daw et al., 2006; Toyokawa et al., 2017). We allowed for the possibility that the balance between exploration-exploitation could change as the task proceeds. To depict such time dependence in exploration, we used the equation: $\beta_{i,t} = \beta^*_{i,0} + \epsilon_i t/70$. If the slope ϵ_i is positive (negative), asocial influence $A_{i,t}$ becomes more and more exploitative (explorative) as round t increases. For a model fitting purpose, the time-dependent term $\epsilon_i t$ is scaled by the total round number 70.

We modelled the social influence (i.e. the frequency-dependent copying) on the probability that individual i chooses option m at round t as follows (McElreath et al., 2005, 2008; Aplin et al., 2017; Barrett et al., 2017):

$$S_{i,t}(m) = \frac{\left(F_{t-1}(m) + 0.1\right)^{\theta_i}}{\sum_{k \in options}\left(F_{t-1}(k) + 0.1\right)^{\theta_i}}, \tag{7}$$

where $F_{t-1}(m)$ is a number of choices made by other individuals (excluding her/his own choice) for the option m in the preceding round $t-1$ ($t \geq 2$). θ_i is individual i's *conformity exponent*, $-\infty \leq \theta_i \leq +\infty$. When this exponent is larger than zero, higher social influence is given to an option which was chosen by more individuals (i.e. positive frequency bias). When $\theta_i < 0$, on the other hand, higher social influence is given to an option that fewer individuals chose in the preceding round $t-1$ (i.e. negative frequency bias). To implement the negative frequency dependence, we added a small number 0.1 to F so that an option chosen by no one (i.e. $F_{t-1} = 0$) could provide the highest social influence when $\theta_i < 0$. Note, there is no social influence when $\theta_i = 0$ because in this case the 'social influence' favours an uniformly random choice, $S_{i,t}(m) = 1/(1+1+1) = 1/3$, independent of the social frequency distribution. Note also that, in the first round $t = 1$, we assumed that the choice is only determined by the asocial softmax function because there is no social information available.

We considered that the social learning weight $\sigma_{i,t}$ could change over time as assumed in the inverse temperature $\beta_{i,t}$. To let $\sigma_{i,t}$ satisfy the constraint $0 \leq \sigma_{i,t} \leq 1$, we used the following sigmoidal function:

$$\sigma_{i,t} = \frac{1}{1 + \exp(-(\sigma^*_{i,0} + \delta_i t/70))}. \tag{8}$$

If the slope δ_i is positive (negative), the social influence increases (decreases) over time. We set the social learning weight equal to zero when group size is one (i.e. when an individual participated in the task alone and/or when $\sum_{k \in options} F_{t-1}(k) = 0$).

We modelled both the inverse temperature $\beta_{i,t}$ and the social learning weight $\sigma_{i,t}$ as a time function since otherwise it would be challenging to distinguish different patterns of learning in this social learning task (Barrett, 2018). The parameter recovery test confirmed that we were able to differentiate such processes under these assumptions (Figure S8-S12). While we also considered the possibility of the conformity exponent being time-dependent (i.e. $\theta_{i,t} = \theta^*_{i,0} + \gamma_i t/70$), the parameter recovery test suggested that the individual slope parameter γ_i was not reliably recovered (Figure S20 and S21), and hence we concentrated our analysis on the time-independent θ_i model. We confirmed that instead using the alternative model where both social learning parameters were time-dependent (i.e. $\sigma_{i,t}$ and $\theta_{i,t}$) did not qualitatively change our results (Figure S25 and S26).

In summary, the model has six free parameters that were estimated for each individual human participant; namely, α_i, $\beta^*_{i,0}$, ϵ_i, $\sigma^*_{i,0}$, δ_i, and θ_i. To fit the model, we used a hierarchical Bayesian method (HBM), estimating the global means (μ_α, $\mu_{\beta^*_0}$, μ_ϵ, $\mu_{\sigma^*_0}$, μ_δ, and μ_θ) and the global variations (ν_α, $\nu_{\beta^*_0}$, ν_ϵ, $\nu_{\sigma^*_0}$, ν_δ, and ν_θ) for each of the three experimental conditions (i.e. the Low-, Moderate- and High-uncertain condition), which govern overall distributions of individual parameter values. It has become recognised that the HBM can provide more robust and reliable parameter estimation than conventional maximum likelihood point estimation in complex cognitive models (e.g. Ahn et al., 2014), a conclusion with which our parameter recovery test agreed (Figure S10-S12).

4.2 Agent-based model simulation

We ran a series of individual-based model simulations assuming that a group of individuals play our three-armed bandit task (under the Moderate-uncertainty condition) and that individuals behave in accordance with the computational learning-and-decision model. We varied the group size ($n \in \{3, 10, 30\}$), the mean social learning weight ($\bar{\sigma} \in \{0.01, 0.1, 0.2, 0.3, ..., 0.9\}$) and the mean conformity exponent ($\bar{\theta} \in \{0.5, 1, 3, 6\}$), running 10,000 replications for each of the possible parameter $\times$ group size combinations. As for the other parameter values (e.g. the asocial reinforcement learning parameters; $\alpha, \beta_0^*, \epsilon$), here we used the experimentally fitted global means (Table 2 and Table S1). Relaxation of this assumption (i.e. using a different set of asocial learning parameters) does not qualitatively change our story (e.g. Figure S4-S7). Note that each individual's parameter values were randomly drawn from the distributions centred by the global mean parameter values fixed to each simulation run. Therefore, the actual composition of individual parameter values were different between individuals even within the same social group.

4.3 Participants in the online experiment

A total of 755 subjects (354 females, 377 males, 2 others and 22 unspecified; mean age (1 $s.d.$) = 34.33 (10.9)) participated in our incetivised economic behavioural experiment (Figure S2). The experimental sessions were conducted in December 2015 and in January 2016. We excluded subjects who disconnected to the online task before completing at least the first 30 rounds from our learning model fitting analysis, resulted in 699 subjects (573 subjects entered the group (i.e. $n \geq 2$) condition and 126 entered the solitary (i.e. $n = 1$) condition). The task was advertised using Amazon's Mechanical Turk (AMT; `https://www.mturk.com`; see Video S1; Video S2), so that the participants could enter anonymously through their own internet browser window. Upon connecting to the experimental game web page, the participants might be required to wait on other participants at the virtual 'waiting room' for up to 5 minutes or until the requisite number of participants arrived, whichever was sooner, before the task starts. The participants were payed 25 cents for a show-up fee plus a waiting-bonus at a rate of 12 cents per minute (i.e. *pro rata* to 7.2 USD per hour) and a game bonus ($mean \pm 1s.d. = 1.7 \pm 0.79$ USD) depending on their performance in the task. The total time, including net time spent in the waiting room, tended to be less than 10 minutes.

4.4 The online three-armed bandit task

The participants performed a three-armed bandit task for 70 rounds. Each round started with the choice stage at which three slot machines appeared on the screen (Figure S1; Video 1). Participants chose a slot by clicking the mouse pointer (or tapping it if they used a tablet computer). Participants had a maximum of 8 seconds to make their choices. If no choice was made during the choice stage, a 'TIME OUT' message appeared in the centre of the screen without a monetary reward (average number of missed rounds per participant was 0.18 out of 70 rounds). Participants were able to know the rest of the choice time by seeing a 'count-down bar' shown at the top of the experimental screen.

Each option yielded monetary rewards randomly drawn from a normal probability distribution unique to each slot, rounded up to the next integer, or truncated to zero if it would have been a negative value (Figure S3). The standard deviations of the probabilistic payoff distributions were identical for all slots and did not change during the task (the $s.d. = 0.55$; although it actually was slightly smaller than 0.55 due to the zero-truncation). The mean values of the probabilistic payoff were different between the options. 'Poor', 'good' and 'excellent' slots generated the lowest, intermediate and the highest rewards on average, respectively. In the first 40 rounds, there were two poor and one good options. After the round 40th, one of the poor option abruptly changed to an excellent option (i.e. environmental change), and from the 41st round there were poor, good and excellent options.

Once all the participants in the group made a choice (or had been time-outed), they proceeded to the feedback stage in which they could see their own payoff from the current choice for two seconds ('0' was shown if they had been time-outed), while they could not see others' reward values. After this feedback stage, subjects proceeded to the next round's choice stage. From the second round, a distribution of choices made by all participants in the group at the preceding round (i.e. the social frequency information) was shown below each slot.

Before the task started, participants had read an illustrated instruction which told them that they would play 70 rounds of the task, that the payoff would be randomly generated per choice but associated with a probability distribution unique to each slot machine, i.e. the profitability of the slot might be different from each other, that the environment might change during the task so that the mean payoff from the slots might secretly change during the task, and that their total payout were decided based on the sum of all earnings they achieved in the task. We also explicitly informed subjects that all participants in the same group played the identical task so that they could infer that the social information was informative. However, we did not specify either the true mean payoff values associated with each option, or when and how the mean payoff would actually change. After reading these instructions, participants proceeded to a 'tutorial task' without any monetary reward and without the social frequency information, so as to become familiar with the task.

After they completed the behavioural task or were excluded from the task due to a bad internet connection or due to opening another browser window during the task (see the 'Reducing the risk of cheating' section in the appendix), subjects proceeded to a brief questionnaire page asking about demographic information, which were skippable. Finally, the result screen was shown, informing the total monetary reward she/he earned as well as a confirmation code unique for each participant. Participants could get monetary reward through AMT by inputting the confirmation code into the form at the AMT's task page.

4.5 Manipulating the group size and uncertainty

To manipulate the size of each group, we varied the capacity of the waiting room from 10 to 30. Because the task was being advertised on the Worker website at AMT for approximately 2 hours, some participants occasionally arrived after the earlier groups had already started. In that case the participant entered the newly opened waiting room which was open for the next 5 minutes. The number of participants arriving declined with time because newly posted alternative tasks were advertised on the top of the task list, which decreased our task's visibility. This meant that a later-starting session tended to begin before reaching maximum room capacity, resulting in the smaller group size. Therefore, the actual size differed between groups.

To investigate the effect of the task uncertainty, we manipulated the closeness of each option's mean payoff value, setting three different conditions in a between-group design. The three conditions were: Low-uncertainty condition (differences between mean payoffs were 1.264; $N = 113$), Moderate-uncertainty condition (differences between mean payoffs were 0.742; $N = 132$) and High-uncertainty condition (differences between mean payoffs were 0.3; $N = 454$). The mean payoff associated with the 'excellent' slot in all three conditions was fixed to 3.1 cents (Figure S3). These conditions were randomly assigned for each experimental session. However, we recruited more participants in the High-uncertainty condition compared to the other two because we expected that larger group sizes would be needed to generate the collective wisdom in noisier environments.

4.6 Statistical analysis

We used a hierarchical Bayesian method (HBM) to estimate the free parameters of our statistical models, including the computational learning-and-decision-making model. The HBM allows us to estimate individual differences, while ensures these individual variations are bounded by the group-level global parameters. The HBM was performed under Stan 2.16.2 (http://mc-stan.org) in R 3.4.1 (https://www.r-project.org) software. The models contained at least 4 parallel chains and we confirmed convergence of the MCMC using both the Gelman-Rubin statistics and the effective sample sizes. Full details of the model fitting procedure and prior assumptions are shown in the appendix.

4.6.1 Parameter recovery test

To check the validity of our model-fitting method, we conducted a 'parameter recovery test' so as to examine how well our model fitting procedure had been able to reveal true individual parameter values. To do this, we generated synthetic data by running a simulation with the empirically fitted global parameter values, and then re-fitted the model with this synthetic data using the same procedure. The parameter recovery test showed that the all true global parameter values were fallen into the 95% Bayesian credible interval (Figure S8), and at least 93% of the true individual parameter values were correctly recovered (i.e. 96% of α_i, 93% of $\beta^*_{i,0}$, 95% of ϵ_i, 97% of $\sigma^*_{i,0}$, 96% of δ_i and 97% of θ_i values were fallen into the 95% Bayesian CI. Figure S9-S12).

4.6.2 Categorisation of individual learning strategies

Based on the 50% CI of the individual *conformity exponent* parameter values θ_i, we divided the participants into the following three different social learning strategies. If her/his 50% CI of θ_i fell above zero ($\theta_{lower} > 0$), below zero ($\theta_{upper} < 0$) or including zero ($\theta_{lower} \leq 0 \leq \theta_{upper}$), she/he was categorised as a 'positive frequency-dependent copier', a 'negative frequency-dependent copier', or a 'random choice individual', respectively. We used the 50% Bayesian CI to conduct this categorisation instead of using the more conservative 95% CI because the latter would cause much higher rates of 'false negatives', by which an individual who applied either a positive frequency-dependent copying or a negative-frequency dependent copying strategy was falsely labelled as an asocial random choice individual (Figure S10d). Four hundred agents out of 572 ($\approx$ 70%) were falsely categorised as a random choice learner in the recovery test when we used the 95% criterion (Figure S10d). On the other hand, the 50% CI criterion seemed to be much better in terms of the false negative rate which was only 18.5% (i.e. 106 agents), although it might be slightly worse in terms of 'false positives': Thirty-seven agents (6.5%) were falsely labelled as either a positive frequency-dependent copier or a negative-frequency dependent copier by the 50% CI, whereas the false positive rate of the 95% CI was only 0.2% (Figure S10e). To balance the risk of false positives and false negatives, we decided to use the 50% CI which seemed to have more strategy detecting power.

4.6.3 Generalised linear mixed models

To examine whether increasing group size and increasing task uncertainty affected individual use of the positive frequency-dependent copying strategy, we used a hierarchical Bayesian logistic regression model with a random effect of groups. The dependent valuable was whether the participant used the positive frequency-dependent copying (1) or not (0). The model includes fixed effects of group size (standardised), task uncertainty (0: Low, 0.5: Moderate, 1: High), age (standardised), gender (0: male, 1: female, NA: others or unspecified), and possible two-way interactions between these fixed effects.

We also investigated the effects of both group size and the task's uncertainty on the fitted values of the learning parameters. We used a hierarchical Bayesian gaussian regression model predicting the individual fitted parameter values. The model includes effects of group size (standardised), task uncertainty (0: Low, 0.5: Moderate, 1: High), age (standardised), gender (0: male, 1: female, NA: others or unspecified), and two-way interactions between these fixed effects. We assumed that the variance of the individual parameter values might be contingent upon task uncertainty because we had found in the computational model-fitting result that the fitted global variance parameters (i.e. $v_{\sigma_0^*}$, v_δ and v_θ) were larger in more uncertain conditions (Table S1).

4.6.4 Post-hoc model simulation for Figure 4d-f

So as to evaluate how accurately our model can generate observed decision pattern in our task setting, we ran a series of individual-based model simulation using the fitted individual parameter values (i.e. means of the individual posterior distributions) for each group size for each uncertainty condition. At the first step of the simulation, we assigned a set of fitted parameters of a randomly-chosen experimental subject from the same group size and the same uncertain condition to an simulated agent, until the number of agents reaches the simulated group size. We allowed duplicate choice of experimental subject in this parameter assignment. At the second step, we let this synthetic group of agents play the bandit task. We repeated these steps 5,000 times for each group size, task uncertainty.

4.7 Code and data availability

The browser based online task was built by Node.js (`https://nodejs.org/en/`) and socket.io (`https://socket.io`), and the code are available on a GitHub repository (`https://github.com/WataruToyokawa/MultiPlayerThreeArmedBanditGame`). Analyses were conducted in R (`https://www.r-project.org`) and simulations of individual based models were conducted in Mathematica (`https://www.wolfram.com`), both of which including data are available on an online repository (URL will appear after acceptance from a journal).

5 Ethics statement

This study was approved by University of St Andrews (BL10808).

6 Competing interest

We have no competing interest.

7 Authors' contributions

WT, AW and KNL planned the study and built the computational model. WT ran simulations. WT and AW made the experimental material, ran the web-base experiment, and collected the experimental data. WT, AW and KNL analysed the data and wrote the manuscript.

8 Acknowledgements

9 Funding

This research was supported by The John Templeton Foundation (KNL; 40128) and Suntory Foundation Research Support (WT; 2015-311). WT was supported by JSPS Overseas Research Fellowships (H27-11).

References

Ahn, W. Y., Vasilev, G., Lee, S. H., Busemeyer, J. R., Kruschke, J. K., Bechara, A., and Vassileva, J. (2014). Decision-making in stimulant and opiate addicts in protracted abstinence: Evidence from computational modeling with pure users. *Frontiers in Psychology*, 5(AUG):1–15. (doi: 10.3389/fpsyg.2014.00849.

Aplin, L. M., Sheldon, B. C., and McElreath, R. (2017). Conformity does not perpetuate suboptimal traditions in a wild population of songbirds. *Proceedings of the National Academy of Sciences*, page 201621067. (doi: 10.1073/PNAS.1621067114.

Arbilly, M., Motro, U., Feldman, M. W., and Lotem, A. (2011). Evolution of social learning when high expected payoffs are associated with high risk of failure. *Journal of the Royal Society, Interface / the Royal Society*, 8(64):1604–15. (doi: 10.1098/rsif.2011.0138.

Barrett, B. J. (2018). Equifinality in empirical studies of cultural transmission. *Behavioural Processes*. (doi: 10.1016/j.beproc.2018.01.011.

Barrett, B. J., Mcelreath, R. L., Perry, S. E., and Barrett, B. J. (2017). Pay-off-biased social learning underlies the diffusion of novel extractive foraging traditions in a wild primate. *Proceedings of Royal Society B*, 284.

Becker, J., Brackbill, D., and Centola, D. (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences*, page 201615978. (doi: 10.1073/pnas.1615978114.

Berdahl, A., Torney, C., Ioannou, C., Faria, J., and Couzin, I. (2013). Emergent Sensing of

Complex Environments by Mobile Animal Groups. *Science*, 339:574–576. (doi: 10.1126/science.1229262.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100(5):992. (doi: 10.1086/261849.

Biro, D., Sasaki, T., and Portugal, S. J. (2016). Bringing a Time-Depth Perspective to Collective Animal Behaviour. *Trends in Ecology and Evolution*, 31(7):550–562. (doi: 10.1016/j.tree.2016.03.018.

Bonabeau, E., Dorigo, M., and Theraulaz, G. (1999). *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press, New York, NY.

Bond, R. (2005). Group Size and Conformity. *Group Processes & Intergroup Relations*, 8(4):331–354. (doi: 10.1177/1368430205056464.

Boyd, R. and Richerson, P. J. (1985). *Culture and the Evolutionary Process*. University of Chicago Press, Chicago, IL.

Boyd, R. and Richerson, P. J. (1989). Social learning as an adaptation. *Lectures on Mathematics in the Life Sciences*, 20:1–26.

Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Theraulaz, G., and Bonabeau, E. (2001). *Self-Organization in Biological Systems*. Princeton University Press, Princeton, NJ.

Chari, V. V. and Kehoe, P. J. (2004). Financial crises as herds: Overturning the critiques. *Journal of Economic Theory*, 119(1 SPEC. ISS.):128–150. (doi: 10.1016/S0022-0531(03)00225-4.

Cronin, A. L. (2016). Group size advantages to decision making are environmentally contingent in house-hunting Myrmecina ants. *Animal Behaviour*, 118:171–179. (doi: 10.1016/j.anbehav.2016.06.010.

Dandurand, F., ShultzEmail, T. R., and Onishi, K. H. (2008). Comparing online and lab methods in a problem-solving experiment. *Behavior Research Methods*, 40(2):428–434. (doi: 10.3758/BRM.40.2.428.

Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B., and Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, 441(7095):876–9. (doi: 10.1038/nature04766.

Dingemanse, N. J., Kazem, A. J., Réale, D., and Wright, J. (2010). Behavioural reaction norms: animal personality meets individual plasticity. *Trends in Ecology and Evolution*, 25(2):81–89. (doi: 10.1016/j.tree.2009.07.013.

Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., and Lubell, M. (2008). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior*, 29:56–64. (doi: 10.1016/j.evolhumbehav.2007.08.003.

Feldman, M. W., Aoki, K., and Kumm, J. (1996). Individual Versus Social Learning: Evolutionary Analysis in a Fluctuating Environment. *Anthropological Science*, 104(3):209–231. (doi: 10.1537/ase.104.209.

Franks, N. R., Mallon, E. B., Bray, H. E., Hamilton, M. J., and Mischler, T. C. (2003). Strategies for choosing between alternatives with different attributes: exemplified by house-hunting ants. *Animal Behaviour*, 65(1):215–223. (doi: 10.1006/anbe.2002.2032.

Giraldeau, L.-A., Valone, T. J., and Templeton, J. J. (2002). Potential disadvantages of using socially acquired information. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 357(1427):1559–66. (doi: 10.1098/rstb.2002.1065.

Gray, J., Hilton, J., and Bijak, J. (2017). Choosing the choice: Reflections on modelling decisions and behaviour in demographic agent-based models. *Population Studies*, 71:85–97. (doi: 10.1080/00324728.2017.1350280.

Hastie, R. and Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological review*, 112(2):494–508. (doi: 10.1037/0033-295X.112.2.494.

Hergueux, J. and Jacquemet, N. (2015). Social preferences in the online laboratory: a randomized experiment. *Experimental Economics*, 18(2):251–283. (doi: 10.1007/s10683-014-9400-5.

Hoppitt, W. and Laland, K. N. (2013). *Social Learning: An Introduction to Mechanisms, Methods, and Models*. Princeton University Press.

Janis, I. L. (1972). *Victims of groupthink: A psychological study of foreign policy*. Houghton Mifflin Company, Boston, NY.

Jayles, B., Kim, H.-r., Escobedo, R., Cezera, S., Blanchet, A., Kameda, T., Sire, C., and Theraulaz, G. (2017). How social information can improve estimation accuracy in human groups. *Proceedings of the National Academy of Sciences*, 114(47):201703695. (doi: 10.1073/pnas.1703695114.

Jolles, J. W., Laskowski, K. L., Boogert, N. J., and Manica, A. (2018). Repeatable group differences in the collective behaviour of stickleback shoals across ecological contexts. *Proceedings of Royal Society B*, pages 13–16.

Kameda, T. and Hastie, R. (2015). Herd Behavior. In *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*, pages 1–14. Wiley Online Library.

Kameda, T. and Nakanishi, D. (2002). Cost–benefit analysis of social/cultural learning in a nonstationary uncertain environment. *Evolution and Human Behavior*, 23(5):373–393. (doi: 10.1016/S1090-5138(02)00101-0.

Kandler, A. and Laland, K. N. (2013). Tradeoffs between the strength of conformity and number of conformists in variable environments. *Journal of theoretical biology*, 332:191–202. (doi: 10.1016/j.jtbi.2013.04.023.

Kendal, R. L., Coolen, I., and Laland, K. N. (2004). The role of conformity in foraging when personal and social information conflict. *Behavioral Ecology*, 15(2):269–277. (doi: 10.1093/beheco/arh008.

King, A. J. and Cowlishaw, G. (2007). When to use social information: the advantage of large group size in individual decision making. *Biology letters*, 3(2):137–139. (doi: 10.1098/rsbl.2007.0017.

King, A. J. and Sueur, C. (2011). Where Next? Group Coordination and Collective Decision Making by Primates. *International Journal of Primatology*, 32(6):1245–1267. (doi: 10.1007/s10764-011-9526-7.

Kline, M. A. and Boyd, R. (2010). Population size predicts technological complexity in Oceania. *Proceedings of the Royal Society B: Biological Sciences*, 277(1693):2559–2564. (doi: 10.1098/rspb.2010.0452.

Krause, J. and Ruxton, G. D. (2002). *Living in groups*. Oxford University Press, Oxford, New York.

Krause, J., Ruxton, G. D., and Krause, S. (2010). Swarm intelligence in animals and humans. *Trends in ecology & evolution*, 25(1):28–34. (doi: 10.1016/j.tree.2009.06.016.

Krause, S., James, R., Faria, J. J., Ruxton, G. D., and Krause, J. (2011). Swarm intelligence in humans: diversity can trump ability. *Animal Behaviour*, 81(5):941–948. (doi: 10.1016/j.anbehav.2010.12.018.

Laan, A., Madirolas, G., and Polavieja, G. G. D. (2017). Rescuing collective wisdom when the average group opinion is wrong. *Frontiers in Robotics and AI*, 4(November):1–28. (doi: 10.3389/frobt.2017.00056.

Laland, K. N. (2004). Social learning strategies. *Animal Learning & Behavior*, 32(1):4–14. (doi: 10.3758/BF03196002.

Le Bon, G. (1896). *The crowd: study of the popular mind*. Unwin, London, UK, 4th edition.

Liker, A. and Bokony, V. (2009). Larger groups are more successful in innovative problem solving in house sparrows. *Proceedings of the National Academy of Sciences*, 106(19):7893–7898. (doi: 10.1073/pnas.0900042106.

List, C. (2004). Democracy in animal groups: a political science perspective. *Trends in Ecology and Evolution*, 19(4):166–168. (doi: 10.1016/j.tree.2004.02.001.

List, C., Elsholtz, C., and Seeley, T. D. (2009). Independence and interdependence in collective decision making: an agent-based model of nest-site choice by honeybee swarms. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1518):755–762. (doi: 10.1098/rstb.2008.0277.

Lorenz, J., Rauhut, H., Schweitzer, F., and Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences of the United States of America*, 108(22):9020–5. (doi: 10.1073/pnas.1008636108.

Lorge, I., Fox, D., Davitz, J., and Brenner, M. (1958). A survey of studies contrating the quality of group performance and individual performance, 1920-1957. *Psychological Bulletin*, 55(6):337–372. (doi: 10.1037/h0042344.

Mackay, C. (1841). *Extraordinary Popular Delusions and the Madness of Crowds*. Richard Bentley, London, UK.

McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., and Waring, T. (2008). Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 363(1509):3515–28. (doi: 10.1098/rstb.2008.0131.

McElreath, R., Lubell, M., Richerson, P. J., Waring, T. M., Baum, W., Edsten, E., Efferson, C., and Paciotti, B. (2005). Applying evolutionary models to the laboratory study of social learning. *Evolution and Human Behavior*, 26(6):483–508. (doi: 10.1016/j.evolhumbehav.2005.04.003.

Mesoudi, A. (2011). An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, 32(5):334–342. (doi: 10.1016/j.evolhumbehav.2010.12.001.

Mesoudi, A., Chang, L., Dall, S. R., and Thornton, A. (2016). The Evolution of Individual and Cultural Variation in Social Learning. *Trends in Ecology & Evolution*, xx:1–11. (doi: 10.1016/j.tree.2015.12.012.

Michelena, P., Jeanson, R., Deneubourg, J.-L., and Sibbald, A. M. (2010). Personality and collective decision-making in foraging herbivores. *Proceedings. Biological sciences / The Royal Society*, 277(1684):1093–9. (doi: 10.1098/rspb.2009.1926.

Morand-Ferron, J. and Quinn, J. L. (2011). Larger groups of passerines are more efficient problem solvers in the wild. *Proceedings of the National Academy of Sciences*, 108(38):15898–15903. (doi: 10.1073/pnas.1111560108.

Morgan, T. J. H., Rendell, L. E., Ehn, M., Hoppitt, W., and Laland, K. N. (2012). The evolutionary basis of human social learning. *Proceedings. Biological sciences / The Royal Society*, 279(1729):653–62. (doi: 10.1098/rspb.2011.1172.

Muchnik, L., Aral, S., and Taylor, S. J. (2013). Social influence bias: a randomized experiment. *Science (New York, N.Y.)*, 341(6146):647–51. (doi: 10.1126/science.1240466.

Muthukrishna, M., Shulman, B. W., Vasilescu, V., and Henrich, J. (2014). Sociality influences cultural complexity. *Proceedings of Royal Society B*, 281(November 2013):20132511.

Nicolis, S. and Deneubourg, J. (1999). Emerging patterns and food recruitment in ants: an analytical study. *Journal of theoretical biology*, 198(4):575–92. (doi: 10.1006/jtbi.1999.0934.

Payzan-Lenestour, E. and Bossaerts, P. (2011). Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Computational Biology*, 7(1). (doi: 10.1371/journal.pcbi.1001048.

Planas-sitja, I., Deneubourg, J.-l., Gibon, C., and Sempo, G. (2015). Group personality during collective decision-making : a multi-level approach. *Proceedings of the Royal Society B: Biological Sciences*, 282:20142515.

Pratt, S. C. and Sumpter, D. J. T. (2006). A tunable algorithm for collective decision-making. *Proceedings of the National Academy of Sciences of the United States of America*, 103(43):15906–10. (doi: 10.1073/pnas.0604801103.

Raafat, R. M., Chater, N., and Frith, C. (2009). Herding in humans. *Trends in cognitive sciences*, 13(10):420–8. (doi: 10.1016/j.tics.2009.08.002.

Reid, C. R. and Latty, T. (2016). Collective behaviour and swarm intelligence in slime moulds. *FEMS Microbiology Reviews*, 40(6):798–806. (doi: 10.1093/femsre/fuw033.

Richerson, P. J. and Boyd, R. (2005). *Not By Genes Alone*. University of Chicago Press, Chicago, IL.

Rosenberg, L. and Pescetelli, N. (2017). Amplifying Prediction Accuracy using Swarm A . I . *Intelligent Systems Conference 2017*, (September):1–5.

Salganik, M. J., Dodds, P. S., and Watts, D. J. (2006). Experimental Study of Inequality and Cultural Market. *Science*, 311(February).

Sasaki, T. and Biro, D. (2017). Cumulative culture can emerge from collective intelligence in animal groups. *Nature Communications*, 8:1–6. (doi: 10.1038/ncomms15049.

Sasaki, T., Granovskiy, B., Mann, R. P., Sumpter, D. J. T., and Pratt, S. C. (2013). Ant colonies outperform individuals when a sensory discrimination task is difficult but not when it is easy. *Proceedings of the National Academy of Sciences*, 110(34):13769–13773. (doi: 10.1073/pnas.1304917110.

Sasaki, T. and Pratt, S. C. (2012). Groups have a larger cognitive capacity than individuals. *Current biology : CB*, 22(19):R827–9. (doi: 10.1016/j.cub.2012.07.058.

Seeley, T., Camazine, S., and Sneyd, J. (1991). Collective decision-making in honey bees: how colonies choose among nectar sources. *Behavioral Ecology and Sociobiology*, 28:277–290.

Seeley, T. D. and Visscher, P. K. (2004). Quorum sensing during nest-site selection by honeybee swarms. *Behavioral Ecology and Sociobiology*, 56(6):594–601. (doi: 10.1007/s00265-004-0814-5.

Shaffer, Z., Sasaki, T., and Pratt, S. C. (2013). Linear recruitment leads to allocation and flexibility in collective foraging by ants. *Animal Behaviour*, 86(5):967–75. (doi: 10.1016/j.anbehav.2013.08.014.

Street, S. E., Navarrete, A. F., Reader, S. M., and Laland, K. N. (2017). Coevolution of cultural intelligence , extended life history , sociality , and brain size in primates. *Proceedings of the National Academy of Sciences*, 114(30):1–7. (doi: 10.1073/pnas.1620734114.

Sumpter, D. J. T. (2010). *Collective Animal Behavior*. Princeton University Press, Princeton, NJ.

Sumpter, D. J. T., Krause, J., James, R., Couzin, I. D., and Ward, A. J. W. (2008). Consensus decision making by fish. *Current biology : CB*, 18(22):1773–7. (doi: 10.1016/j.cub.2008.09.064.

Sutton, R. S. and Barto, a. G. (1998). *Reinforcement learning: an introduction.* MIT Press, Cambridge, MA. (doi: 10.1109/TNN.1998.712192.

Toyokawa, W., Kim, H.-R., and Kameda, T. (2014). Human Collective Intelligence under Dual Exploration-Exploitation Dilemmas. *PloS one*, 9(4):e95789. (doi: 10.1371/journal.pone.0095789.

Toyokawa, W., Saito, Y., and Kameda, T. (2017). Individual differences in learning behaviours in humans: asocial exploration tendency does not predict reliance on social learning. *Evolution and Human Behavior*, 38(3):325–333. (doi: 10.1016/j.evolhumbehav.2016.11.001.

Trimmer, P. C., McNamara, J. M., Houston, A. I., and a.R. Marshall, J. (2012). Does natural selection favour the Rescorla–Wagner rule? *Journal of Theoretical Biology*, 302:39–52. (doi: 10.1016/j.jtbi.2012.02.014.

Ward, A. J. W., Herbert-Read, J. E., Sumpter, D. J. T., and Krause, J. (2011). Fast and accurate decisions through collective vigilance in fish shoals. *Proceedings of the National Academy of Sciences*, 108(6):2312–2315. (doi: 10.1073/pnas.1007102108.

Webster, M. and Laland, K. (2008). Social learning strategies and predation risk: minnows copy only when using private information would be costly. *Proceedings of the Royal Society B: Biological Sciences*, 275(1653):2869–2876. (doi: 10.1098/rspb.2008.0817.

Webster, M. M. and Laland, K. N. (2011). Reproductive state affects reliance on public information in sticklebacks. *Proceedings. Biological sciences / The Royal Society*, 278(1705):619–627. (doi: 10.1098/rspb.2010.1562.

Webster, M. M., Whalen, A., and Laland, K. N. (2017). Fish pool their experience to solve problems collectively. *Nature Ecology and Evolution*, 1(5):1–5. (doi: 10.1038/s41559-017-0135.

Wolf, M., Kurvers, R. H. J. M., Ward, A. J. W., Krause, S., and Krause, J. (2013). Accurate decisions in an uncertain world: collective cognition increases true positives while decreasing false positives. *Proceedings. Biological sciences / The Royal Society*, 280(1756):20122777. (doi: 10.1098/rspb.2012.2777.

1 Appendix 1 Supplementary experimental procedure

1.1 Amazon Mechanical Turk

The online task was advertised as a 'HIT (Human Intelligence Task)' entitled 'Lottery Choice Experiment (about 15 mins + fun + bonus!)' in Amazon Mechanical Turk (AMT). The HIT was only available for individuals whose 'HIT Approval Rate' was greater than or equal to 90% and who live in the US.

On the advertisement page, we stressed that there could be extra payoff to subjects depending on their performance in the task; that stability of their Internet connection was necessary; and that they could participate in the task only once (Video S1).

At the bottom of the advertisement page, an URL link to our experimental server would appear. The link was not shown until the participants decided to join the task. There were also text forms into which the participants could input their own confirmation code, which they would get by finishing the experimental task, and where they could add comments on the task. By clicking the submit button below they could complete the task, allowing the monetary reward to be paid through Amazon system.

1.2 A written consent form and an instruction for the task

On clicking the URL link shown at the bottom of the HIT advertisement page, participants proceeded to a consent form that emphasised data anonymity and asked them not to interact with anyone during the task (Video S1). Scrolling this consent form page, participants proceeded to sign the form. After answering all the YES/NO questions and inputting the CAPTCHA, participants could proceed to instructions.

On the first page of the instructions participants were informed that the study was split into two parts: an interactive economic decision-making game and a short survey. On the second page of the instructions the details of the decision task were explained with illustrations.

Full details and code of the task are available in GitHub (`https://github.com/WataruToyokawa/MultiPlayerThreeArmedBanditGame`).

1.3 Reducing the risk of cheating

To minimise the risk of multiple accesses from the same person, we introduced the restriction that a single 'worker ID' associated with participants' AMT accounts, could participate only once in the experiment. We rejected access from the same IP address: If a participant's IP address had already been stored in our database, the participant directly proceeded from the instruction page to the questionnaire page. In that case, 25 cents show-up fee was still paid because it was possible that different persons might use the same IP address.

To minimise the risk of opening other browser windows during the task (for example, browsing other websites), we used 'Page Visibility API' (`https://developer.mozilla.org/en-US/docs/Web/API/Page_Visibility_API`) to track whether the experimental browser window was always active and not hidden by other browser windows or tabs for more than 1 second. If it was detected that the experimental window was in a hidden state, the participant was automatically redirected to the questionnaire page. In that case, 25 cents show-up fee plus a waiting-bonus (if applicable) and a game-bonus earned so far were paid. In the instruction, participants were warned not to open any other browser windows/tabs during the task and were informed that they

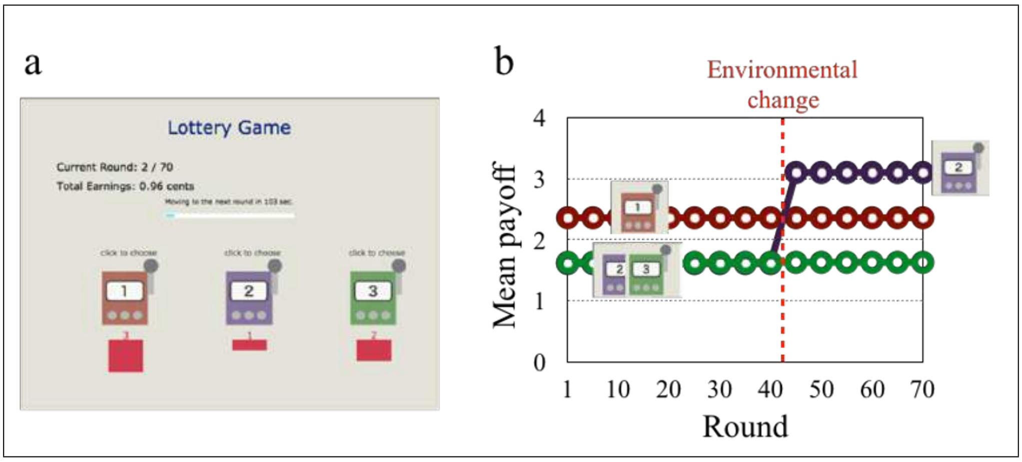

Figure S1. The three-armed bandit task. (a) Illustration of the user interface of the task. Participants could choose an option by clicking one of the slot machines. The frequency distribution of choices made by participants in the same group in the preceding round (i.e. the social frequency information) is shown by red numbers below each option. The lengths of the red bars are proportional to the social frequency distributions. Participants could also see their current total earnings (0.96 cents in this example) as well as the current round number (Round 2 in this example). (b) Example of mean payoffs for each option in the task. The payoff received for a particular choice is drawn from a Gaussian distribution with noise around each mean (with a fixed standard deviation: $1S.D. = 0.55$). Note that the most profitable slot (the 'optimal option') was switched (i.e. environmental change) after 40 rounds: One 'poor' slot (slot-2 in this example) changed to an 'excellent' one at the beginning of the 41st round, while the other machines' expected returns were not changed. The association between each option's number and its payoff was randomly assigned across experimental sessions.

would be eliminated from the session if they did so.

2 Appendix 2 Supplementary computational modelling procedure

2.1 Hierarchical Bayesian parameter estimation

We used hierarchical Bayesian method (HBM) to estimate the free parameters of our learning model. HBM allows us to estimate individual differences, while this individual variation is bounded by the group-level (i.e. hyper) parameters. To perform HBM, we used Stan 2.16.2 in R 3.4.1.

In our model, there are 6 individual parameters; namely, α_i, $\beta^*_{i,0}$, ϵ_i, $\sigma^*_{i,0}$, δ_i, and θ_i. Because the learning rate α_i is bounded between 0 and 1, we estimated α^*_i rather than α_i itself ($-\infty \leq \alpha^*_i \leq +\infty$), which is given by the following sigmoidal function:

$$\alpha_i = \frac{1}{1 + \exp(-\alpha^*_i)}. \tag{1}$$

We assumed the Student's t distributions for individual random effects of each parameter so as to allow a few 'outliers', because the Student's t distribution has a longer tail compared to a normal distribution. To do so, we used the following reparameterization for each $parameter(.) \in \{\alpha^*_i, \beta^*_{i,0}, \epsilon_i, \sigma^*_{i,0}, \delta_i, \theta_i\}$:

$$parameter(.)_i = \mu_{(.),c} + \nu_{(.),c} * parameter(.)\_raw_i, \tag{2}$$

where $\mu_{(.),c}$ is a global mean of the $parameter(.)$ in the condition c ($c \in$ {Low- , Moderate-, High-uncertainty condition}), and $\nu_{(.),c}$ is a global scale parameter of the individual variations in condition c, which is multiplied by a standardised individual random variable $parameter(.)\_raw_i$ drawn from

$$parameter(.)\_raw_i \sim Student\_t\big(df = 4, location = 0, scale = 1\big). \tag{3}$$

As for the global parameters, we used a normal and a half-normal prior distributions for $\mu_{(.),c}$ and $\nu_{(.),c}$, respectively:

$$\mu_{(.),c} \sim Normal\big(mean = 0, sd = 5\big), \tag{4a}$$

$$\nu_{(.),c} \sim Normal^{+}\big(mean = 0, sd = 3\big). \tag{4b}$$

In summary, there are 36 global free parameters (= 6 μs and 6 νs for 3 different conditions each). A total of 2000 iterations were performed after 1000 warm-up with thin = 5 for each of 8 chains (= 2000 samples / 5 steps $\times$ 8 chains = a total of 3200 samples). We used the Gelman-Rubin statistics (as known as $\hat{R}$) as well as the effective sample sizes (ESS) so as to check the convergence of the MCMC samples. All global parameter values had $\hat{R} \approx 1.00 \leq 1.10$ indicating that chains are converged to the target distributions. The ESS of model parameters were typically greater than 500 (out of 3200 total samples).The minimum ESS of global-parameters was 233 (on $\nu_{\epsilon,Low}$). Visual inspection of the parameters with smaller ESSs confirmed their convergence to target distributions. We confirmed that changing both *df* (i.e. broadness of the tail) of the Student's t prior distributions and *sd* of the Normal prior distributions did not change our findings.

2.2 Parameter recovery test

To assess the adequacy of the hierarchical Bayesian model-fitting method, we tested how well the HBM could recover 'true' parameter values that were used to simulate synthetic data. We simulated participants' behaviour assuming that they behave according to the model with each parameter setting. We generated 'true' parameter values for each simulated agent based on the experimentally fit global parameters (Table 1 in the main text). We then simulated synthetic behavioural data and recovered their parameter values using the HBM described above.

2.3 Time-dependent conformity exponent $\theta_{i,t}$ model

We also considered the possibility of the conformity exponent being time-dependent (i.e. $\theta_{i,t} = \theta^*_{i,0} + \gamma_i t/70$). If the slope γ_i is positive (negative), the frequency-dependent bias increases (decreases) over time. In this model, there are 7 individual parameters; namely, α_i, $\beta^*_{i,0}$, ϵ_i, $\sigma^*_{i,0}$, δ_i, $\theta^*_{i,0}$ and γ_i. We fitted this model to the experimental data using the HBM descried above.

84 ## 2.4 Other figures related to the methods

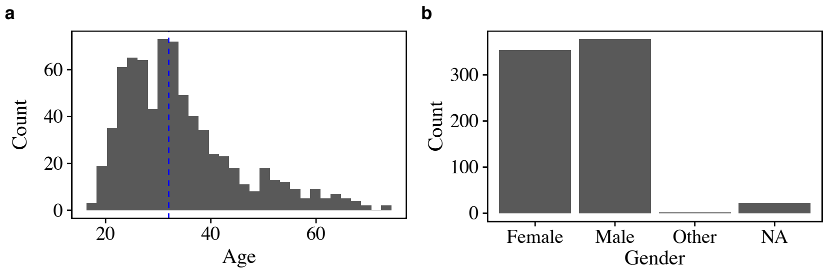

Figure S2. Histograms of the participants' age and gender. The mean age is indicated by a blue dashed line. Note that these data were inputted by participants themselves on the questionnaire forms.

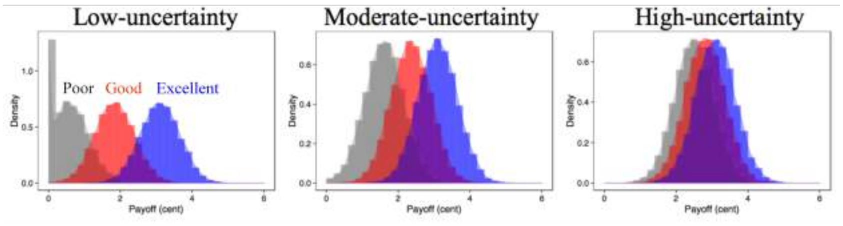

Figure S3. The distributions of payoffs generated by each of the slot machines for each condition. The poor, good and excellent slot are indicated by grey, red and blue, respectively. The payoff was truncated to zero if it would have been a negative value.

3 Appendix 3 Supplementary results

3.1 Individual-based simulation using other parameter sets

Individual-based model simulations using a different set of asocial learning parameters suggest that our main findings from the simulation (Figure 1, 2 in the main text) are broadly robust in a range of parameter combinations (Figure S4, S5, S6, S7).

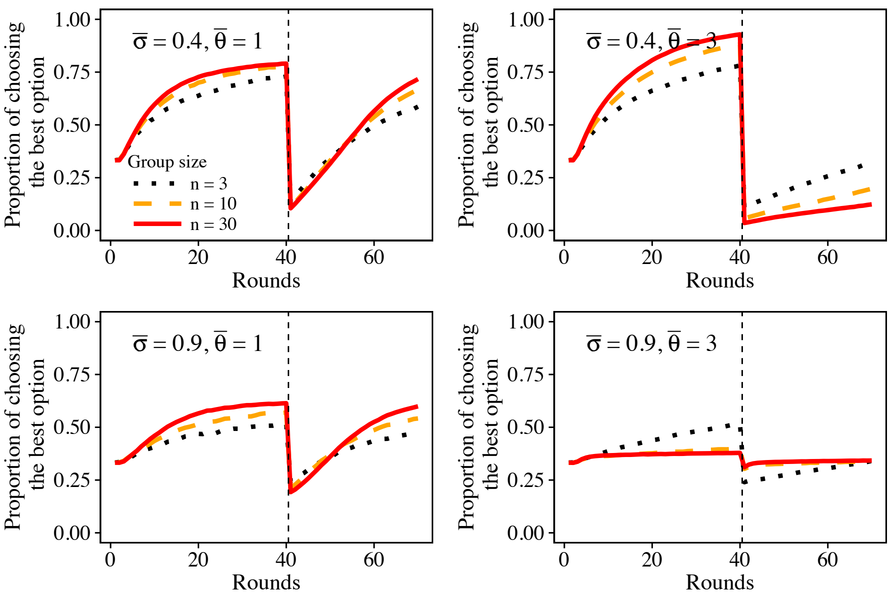

Figure S4. The same figure as Figure 1 in the main text, except for the asocial learning parameter settings (i.e. $\mu_\alpha = 0.7$, $\mu_{\beta_0^*} = 2$, $\mu_\epsilon = 4$, $\nu_\alpha = 1$, $\nu_{\beta_0^*} = 1$, $\nu_\epsilon = 1$, $\nu_\sigma = 1$, and $\nu_\theta = 1$).

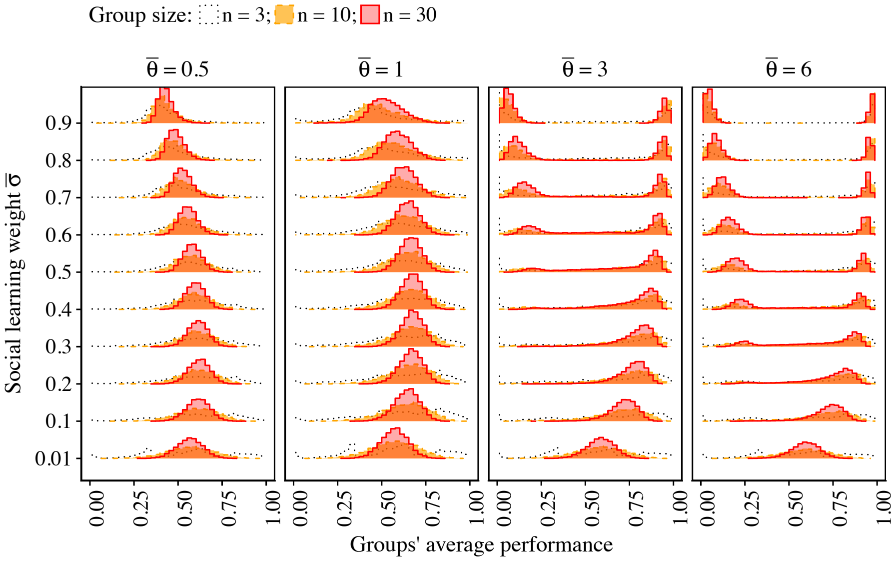

Figure S5. The same figure as Figure 2 in the main text, except for the asocial learning parameter settings. Parameter values were the same in Figure S4.

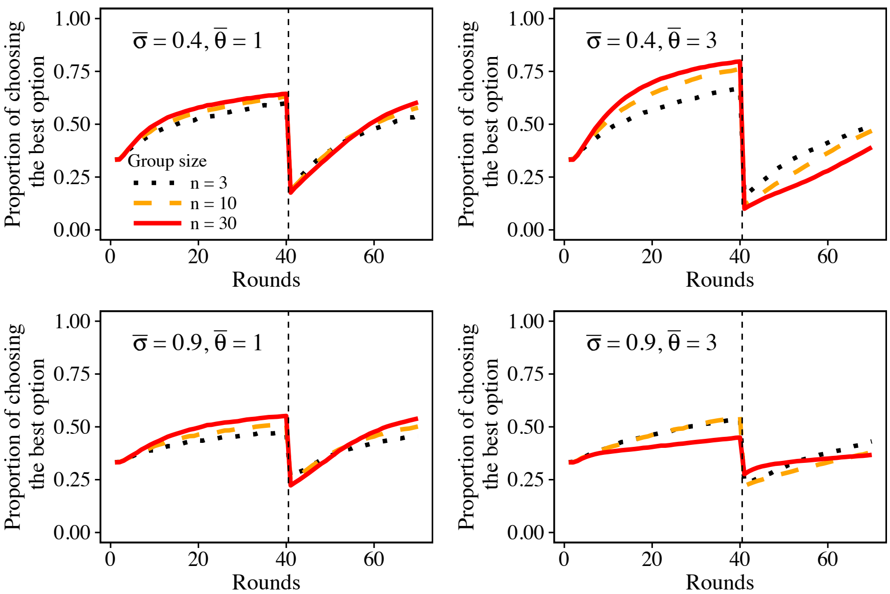

Figure S6. The same figure as Figure 1 in the main text, except for the asocial learning parameter settings (i.e. $\mu_\alpha = 0.8$, $\mu_{\beta_0^*} = 0.5$, $\mu_\epsilon = 3$, $\nu_\alpha = 1$, $\nu_{\beta_0^*} = 2$, $\nu_\epsilon = 2$, $\nu_\sigma = 2$, and $\nu_\theta = 2$).

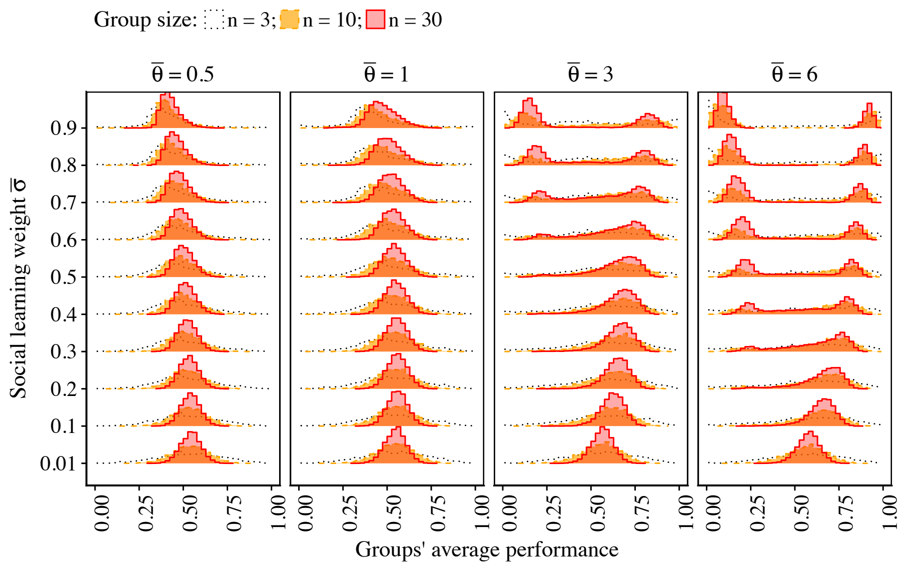

Figure S7. The same figure as Figure 2 in the main text, except for the asocial learning parameter settings. Parameter values were the same in Figure S6.

90 ## 3.2 Parameter recovery test

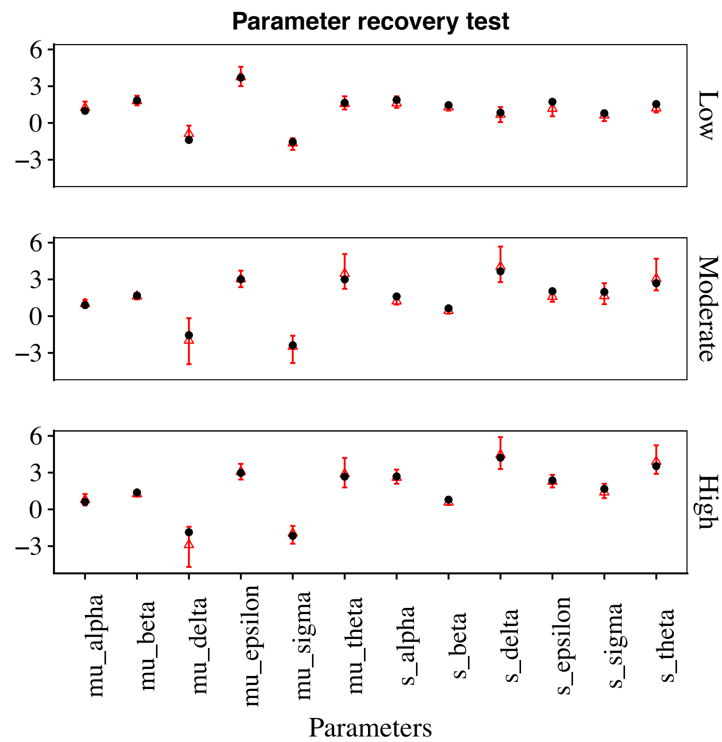

Figure S8. The parameter recovery performance on the global parameters. The black points are the true values and the red triangles are the mean posterior values (i.e. recovered values). The 95% Bayesian credible intervals are shown by the error bars.

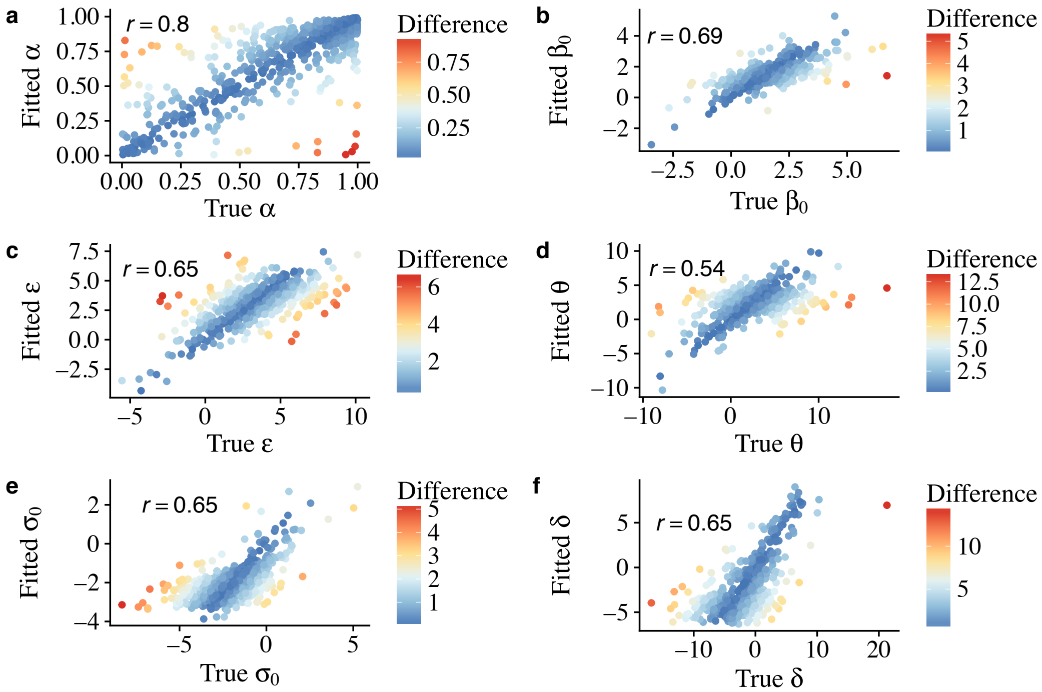

Figure S9. The parameter recovery performance on the individual parameters. The x-axis is the true value and the y-axis is the fitted (i.e. the mean posterior) value. The differences between the true value and the fitted value are shown in different colours. The correlation coefficients between the true value and the fitted value are shown.

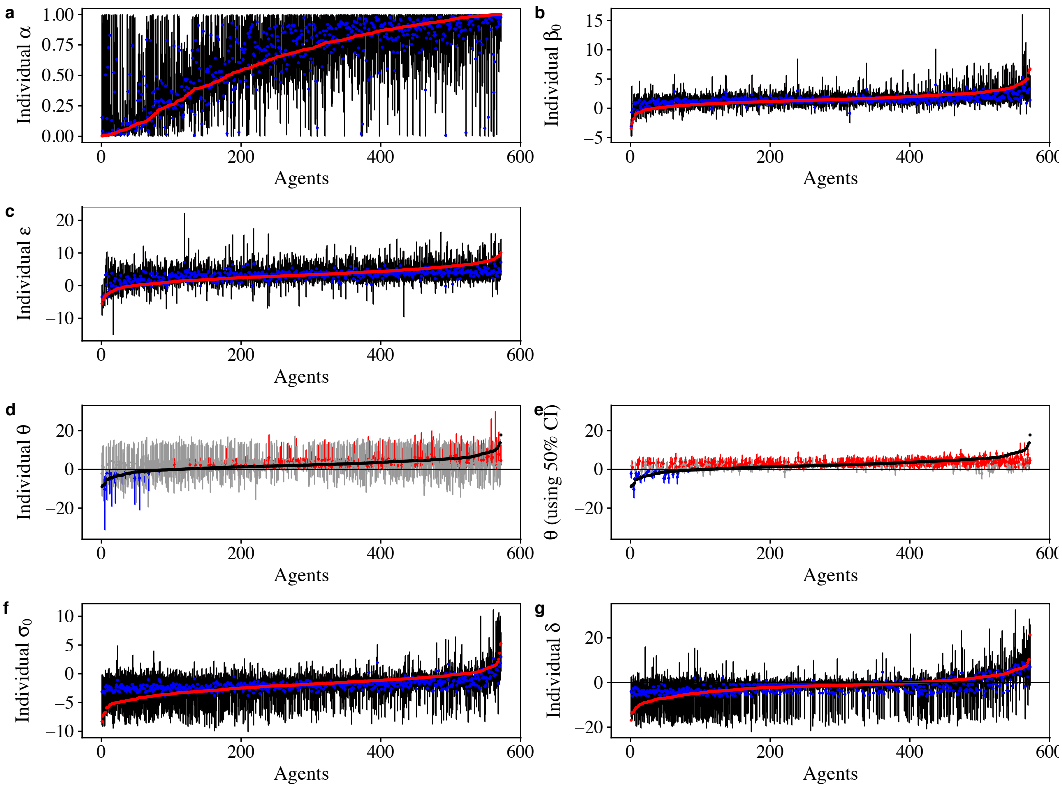

Figure S10. The parameter recovery performance on the individual parameters. (a,b,c,f,g) The red points are the true individual parameter values, the blue points are the mean posterior fitted values and the black lines are the 95% Bayesian CI. (d) The categorisation of three different strategies based on the 95% Bayesian CI and (e) on the 50% Bayesian CI of individual θ_i values. The red coloured individuals are categorised as the positive frequency-dependent copiers (positive frequency-biased choice), the blue coloured individuals are categorised as the negative frequency-dependent copiers (negative frequency-biased choice) and the grey individual are categorised as the asocial random copiers (nearly random decision-making). The black points are the true θ_i values. The horizontal lines indicate the categorisation threshold where $\theta_i = 0$.

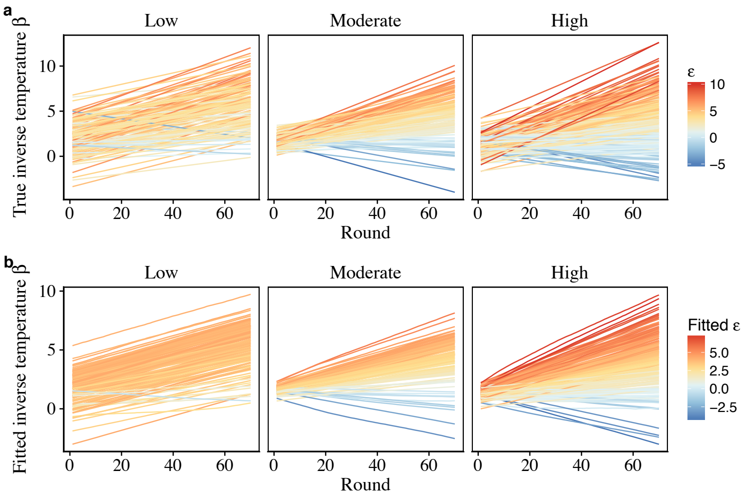

Figure S11. The temporal evolution of (a) the true individual *inverse temperature* $\beta_{i,t}$ parameters and (b) the recovered $\beta_{i,t}$ value. The magnitude of individual slope parameter ϵ_i are shown in different colours.

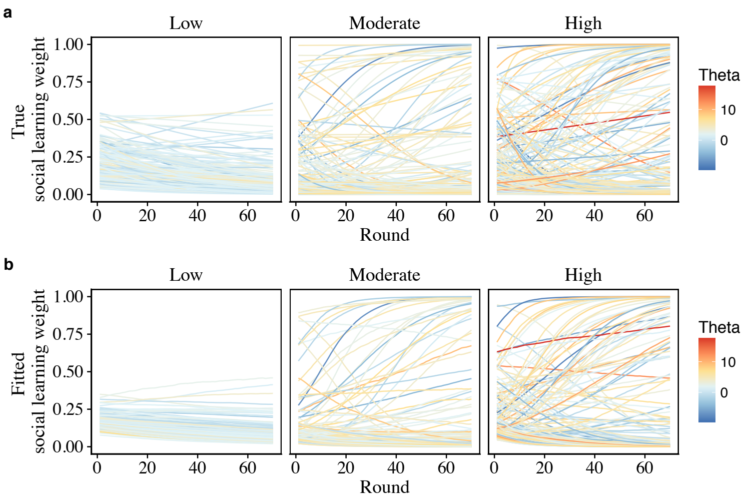

Figure S12. The temporal evolution of (a) the true individual *social learning weight* $\sigma_{i,t}$ parameters and (b) the recovered $\sigma_{i,t}$ value. The magnitude of individual conformity exponent θ_i are shown in different colours.

91 ## 3.3 Model fitting to our experimental data

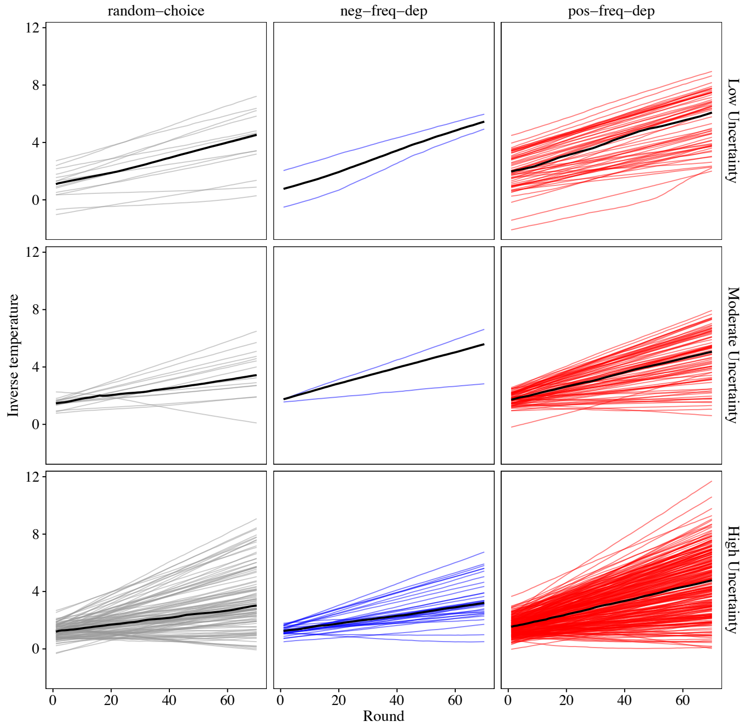

Figure S13. Individual inverse temperature $\beta_{i,t}$ fit for each experimental participant.

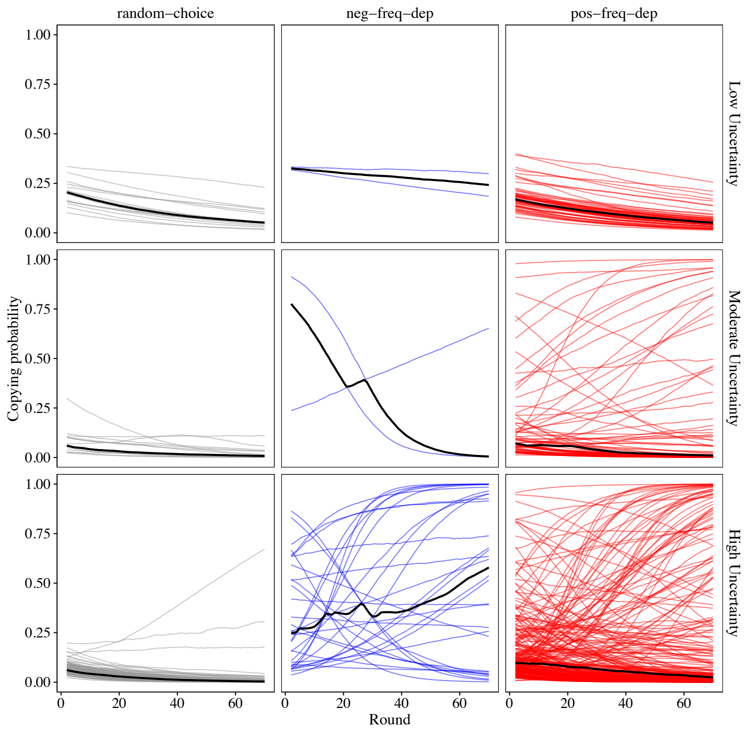

Figure S14. Individual social learning weights $\sigma_{i,t}$ fit for each experimental participant.

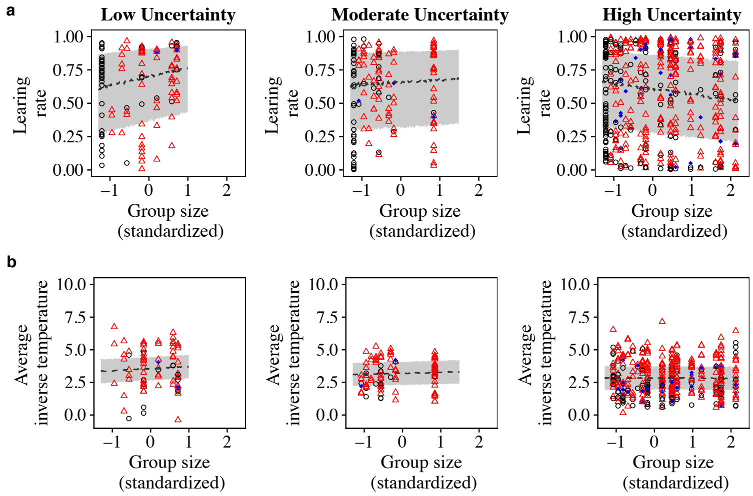

Figure S15. (a) Estimated learning rate α_i and (b) estimated mean inverse temperature $\bar{\beta}_i = (\sum_t \beta_{i,t})/70$ for each individual shown for each different learning strategy (red-triangle: positive frequency-biased choice, blue-diamond: negative frequency-biased choice; grey open circle: nearly random decision-making). Predictions from the fitted generalised mixed models are shown by dashed lines (the shaded areas indicate 50% Bayesian credible intervals).

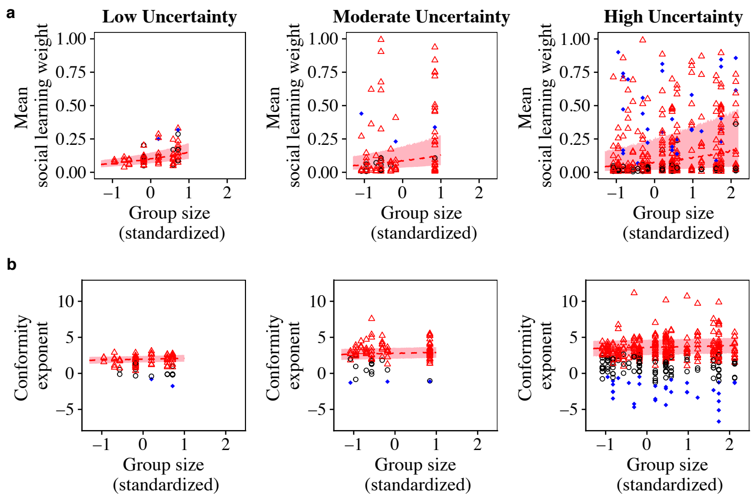

Figure S16. (a) Estimated social learning weight and (b) estimated conformity exponent for each individual shown for each different learning strategy (red-triangle: positive frequency-biased choice, blue-diamond: negative frequency-biased choice; grey open circle: nearly random decision-making). The dashed lines show regressions of the fitted generalised mixed models for only the positive frequency-biased choice individuals (the shaded areas indicate 50% Bayesian credible intervals).

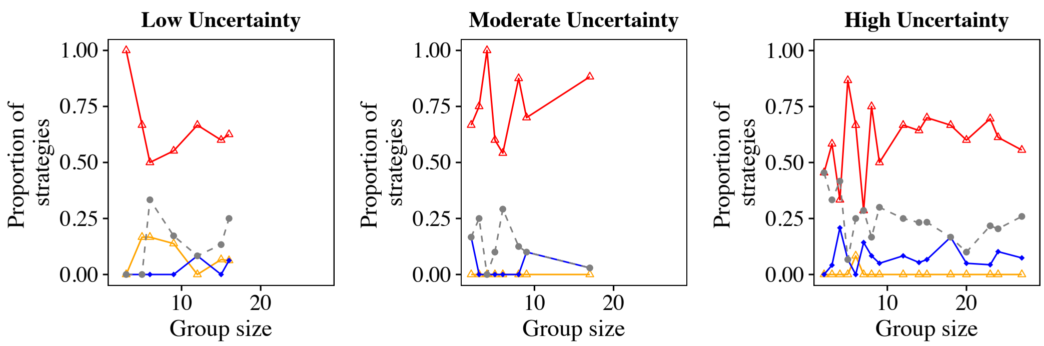

Figure S17. Model fitting for the three different task's uncertain conditions (the Low-, Moderate- and High-uncertainty) and the different group size. Frequencies of four different learning strategies are shown in different styles (red-triangle: strong positive frequency-dependent learning $\theta_i > 1$, orange-triangle: weak positive frequency-dependent learning $0 < \theta_i \leq 1$, blue-circle: negative frequency-dependent learning; grey-circle: nearly random choice strategy).

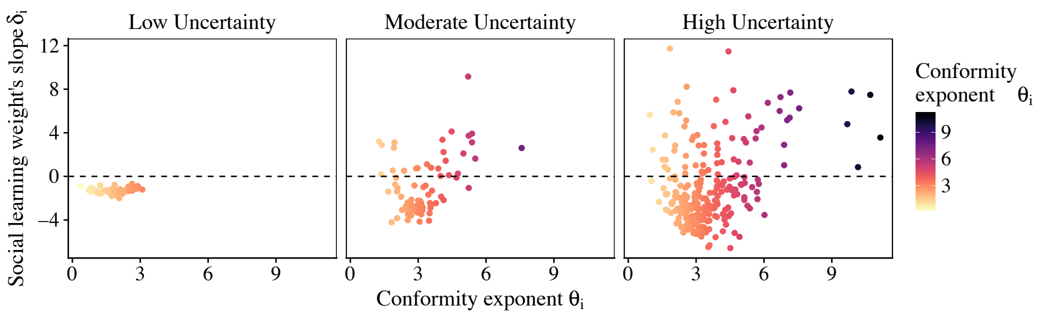

Figure S18. The relationship between θ_i and the slope of social learning weight δ_i. The horizontal dashed lines indicate a threshold at which $\sigma_{i,t}$ does not change with time (i.e. $\delta_i = 0$).

Table S1

The mean and the 95% Bayesian credible intervals of the posterior global variance parameters governing the magnitude of individual variations in each free parameter. The number of participants (N) for each experimental condition are also shown.

	Group condition			Solitary condition		
	Uncertainty			Uncertainty		
Parameters	Low	Moderate	High	Low	Moderate	High
v_{α^*} (learning rate)	1.88	1.61	2.69	1.79	2.37	2.34
	[1.20, 2.82]	[1.14, 2.23]	[2.23, 3.21]	[0.98, 3.03]	[1.46, 3.58]	[1.39, 3.77]
$v_{\beta_0^*}$ (inv. temp.)	1.45	0.64	0.79	0.77	0.96	0.71
	[0.91, 2.16]	[0.34, 1.00]	[0.61, 1.00]	[0.42, 1.25]	[0.46, 1.61]	[0.40, 1.08]
v_{ϵ} (inv. temp.)	1.73	2.04	2.36	1.75	1.76	2.23
	[0.41, 3.34]	[1.27, 2.96]	[1.89, 2.85]	[0.85, 2.94]	[0.85, 2.87]	[1.43, 3.08]
$v_{\sigma_0^*}$ (soc. wight)	0.79	1.98	1.67	–	–	–
	[0.17, 1.50]	[1.17, 3.15]	[1.26, 2.18]	–	–	–
v_{δ} (soc. wight)	0.84	3.66	4.22	–	–	–
	[0.05, 1.94]	[1.92, 5.78]	[3.30, 5.40]	–	–	–
v_{θ} (conformity coeff.)	1.54	2.69	3.53	–	–	–
	[0.87, 2.61]	[1.60, 4.32]	[2.67, 4.65]	–	–	–
N	77	98	398	36	34	56

3.4 Results of a time-dependent conformity model ($\theta_{i,t}$)

3.4.1 Parameter recovery test

The parameter recovery test showed that the all true global parameter values were fallen into the 95% Bayesian credible interval (Figure S19). Correlations between individual true parameters and recovered parameters were all positive, while the correlation coefficients of both $\theta^*_{i,0}$ and γ_i were lower than other parameters (Figure S20). At least 89% of the true individual parameter values were correctly recovered (i.e. 97% of α_i, 96% of $\beta^*_{i,0}$, 97% of ϵ_i, 96% of $\sigma^*_{i,0}$, 94% of δ_i, 96% of $\theta^*_{i,0}$ and 89% of γ_i were fallen into the 95% Bayesian CI; Figure S21).

Figure S22, S23 and S24 show that overall patterns of temporal dynamics of these parameters were well recovered.

3.4.2 Fitting to our experimental data

In order to compared the findings from the time-independent θ_i model (see Results section in the main text), we again categorized the participants as deploying three different learning strategies based on their mean fitted conformity exponent values; namely, the 'positive frequency-dependent copying' strategy ($\bar{\theta}_i \gg 0$), the 'negative-frequency dependent copying' strategy ($\bar{\theta}_i \ll 0$) and the 'random choice' strategy ($\bar{\theta}_i \approx 0$). Note, the conformity exponent here is averaged over time: $\bar{\theta}_i = (\sum_t \theta_{i,t})/70$. Figure S25 suggests that the patterns were consistent with Figure 3 in the main text, and hence our conclusion was not changed.

Individual frequency dependence changed over time (Figure S26). The conformity exponents generally increased with experimental round, while some individuals in the High-uncertain con-

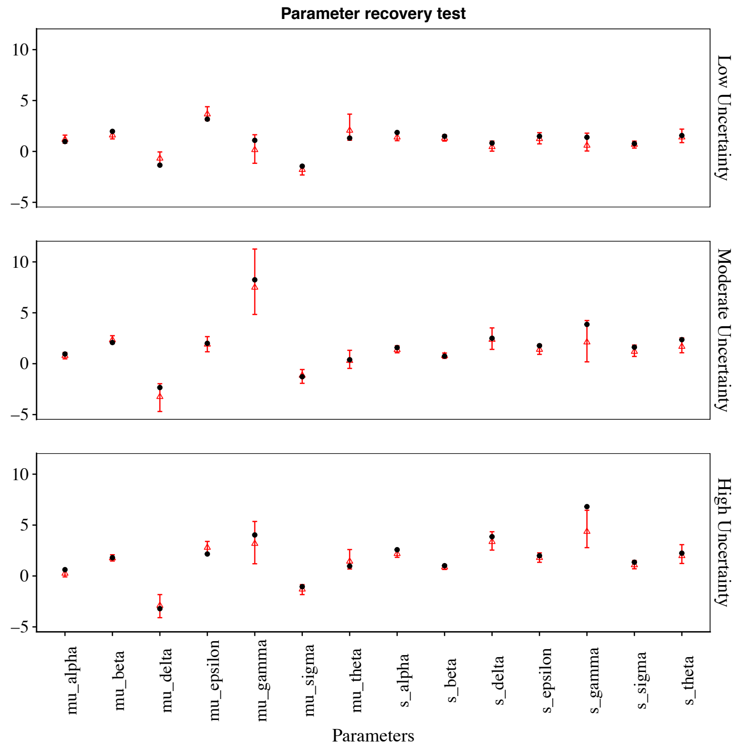

Figure S19. The parameter recovery performance on the global parameters. The black points are the true values and the ref triangles are the mean posterior values (i.e. recovered values). The 95% Bayesian credible intervals are shown by the error bars.

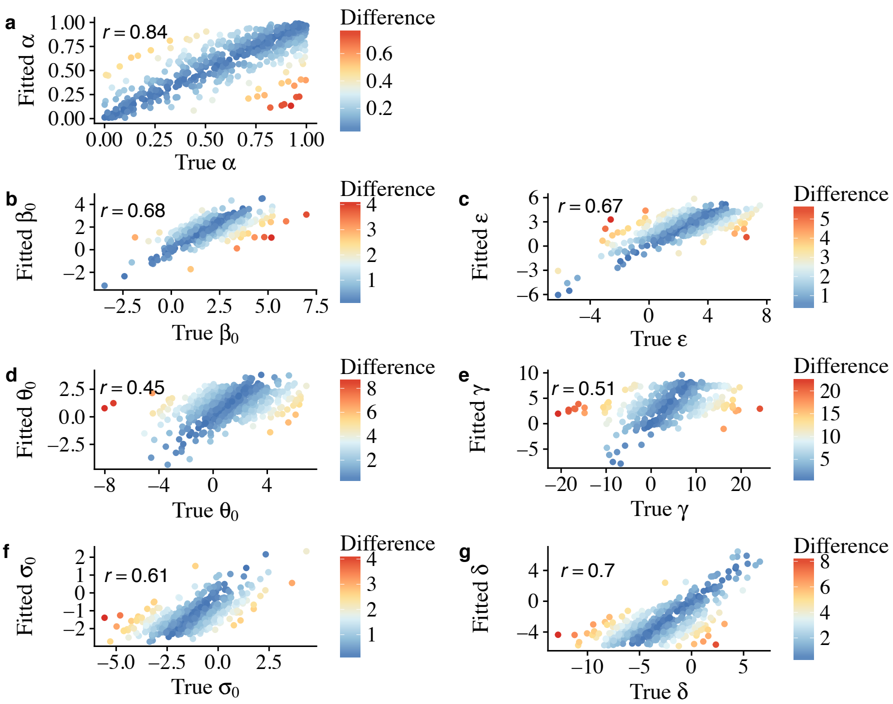

Figure S20. The parameter recovery performance on the individual parameters. The x-axis is the true value and the y-axis is the fitted (i.e. the mean posterior) value. The differences between the true value and the fitted value are shown in different colour. The correlation coefficients between the true value and the fitted value are shown.

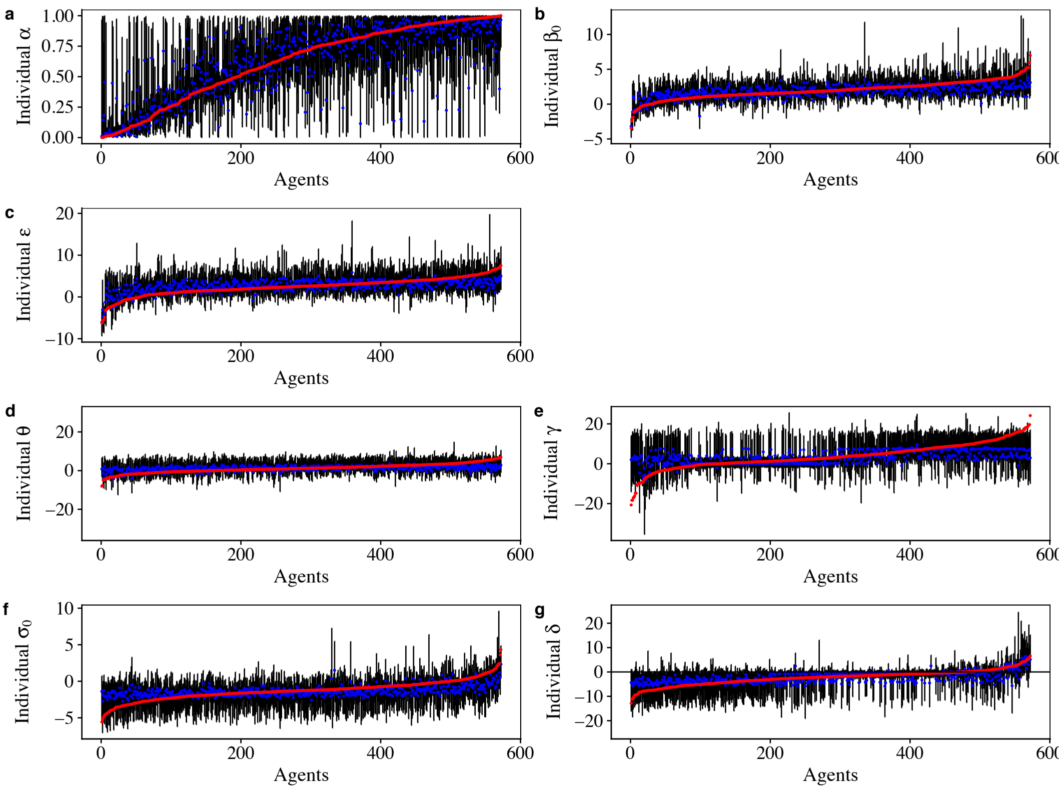

Figure S21. The parameter recovery performance on the individual parameters. (a,b,c,e,g) The red points are the trues individual parameter values, the blue points are the mean posterior fitted values and the black lines are the 95% Bayesian CI. (d) The categorisation of three different strategies based on the 95% Bayesian CI and (e) on the 50% Bayesian CI of individual θ_i values. The red coloured individuals are categorised as the positive frequency-dependent copiers (positive frequency-biased choice), the blue coloured individuals are categorised as the negative frequency-dependent copiers (negative frequency-biased choice) and the grey individual are categorised as the asocial random copiers (nearly random decision-making). The black points are the true θ_i values. The horizontal lines indicate the categorisation threshold where $\theta_i = 0$.

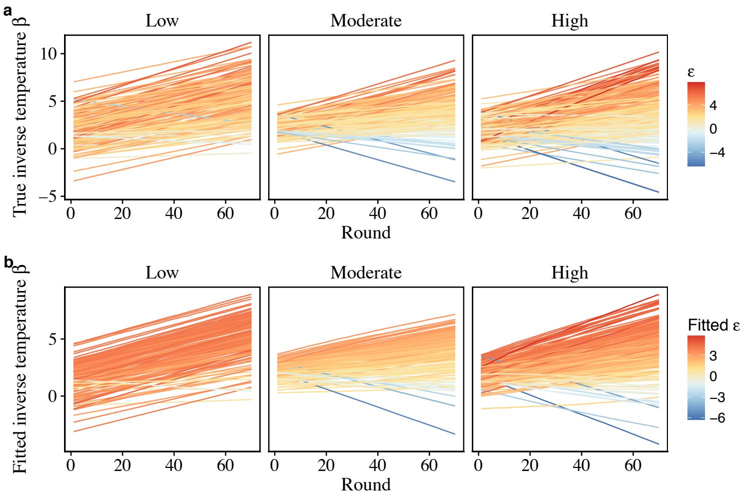

Figure S22. The temporal evolution of (a) the true individual *inverse temperature* $\beta_{i,t}$ parameters and (b) the recovered $\beta_{i,t}$ value. The magnitude of individual slope parameter ϵ_i are shown in different colours.

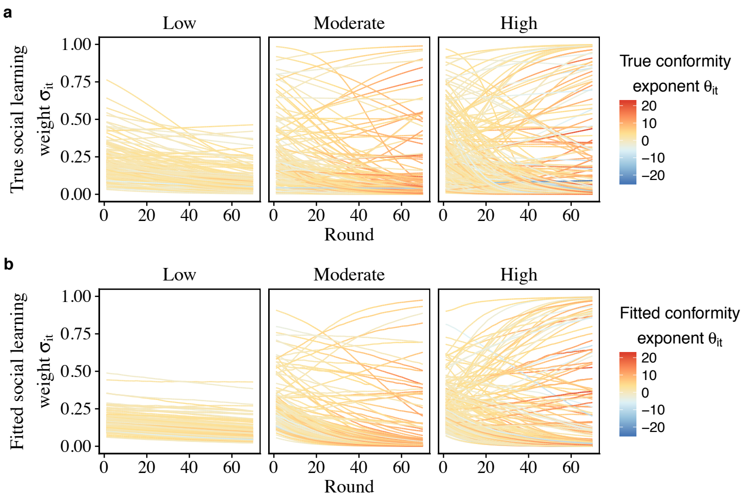

Figure S23. The temporal evolution of (a) the true individual *social learning weight* $\sigma_{i,t}$ parameters and (b) the recovered $\sigma_{i,t}$ value. The magnitude of individual conformity exponent θ_i are shown in different colours.

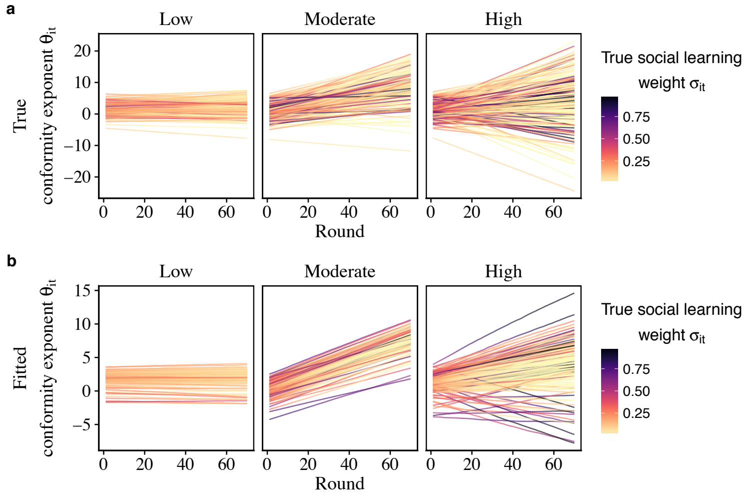

Figure S24. The temporal evolution of (a) the true and (b) the recovered individual *conformity exponent* $\theta_{i,t}$ value. The magnitude of individual social learning weights $\sigma_{i,t}$ are shown in different colours.

ditions decreased rather than accelerated their frequency dependence over time. However, note that the fitting of slope parameter γ_i was relatively unreliable (i.e. only 89% of individual parameters were recovered correctly). Extensive variation in both the social learning weigh $\sigma_{i,t}$ and the conformity exponent $\theta_{i,t}$ found in high-uncertain circumstances are consistent with the main findings (Figure 3g-i, Figure 4a-c).

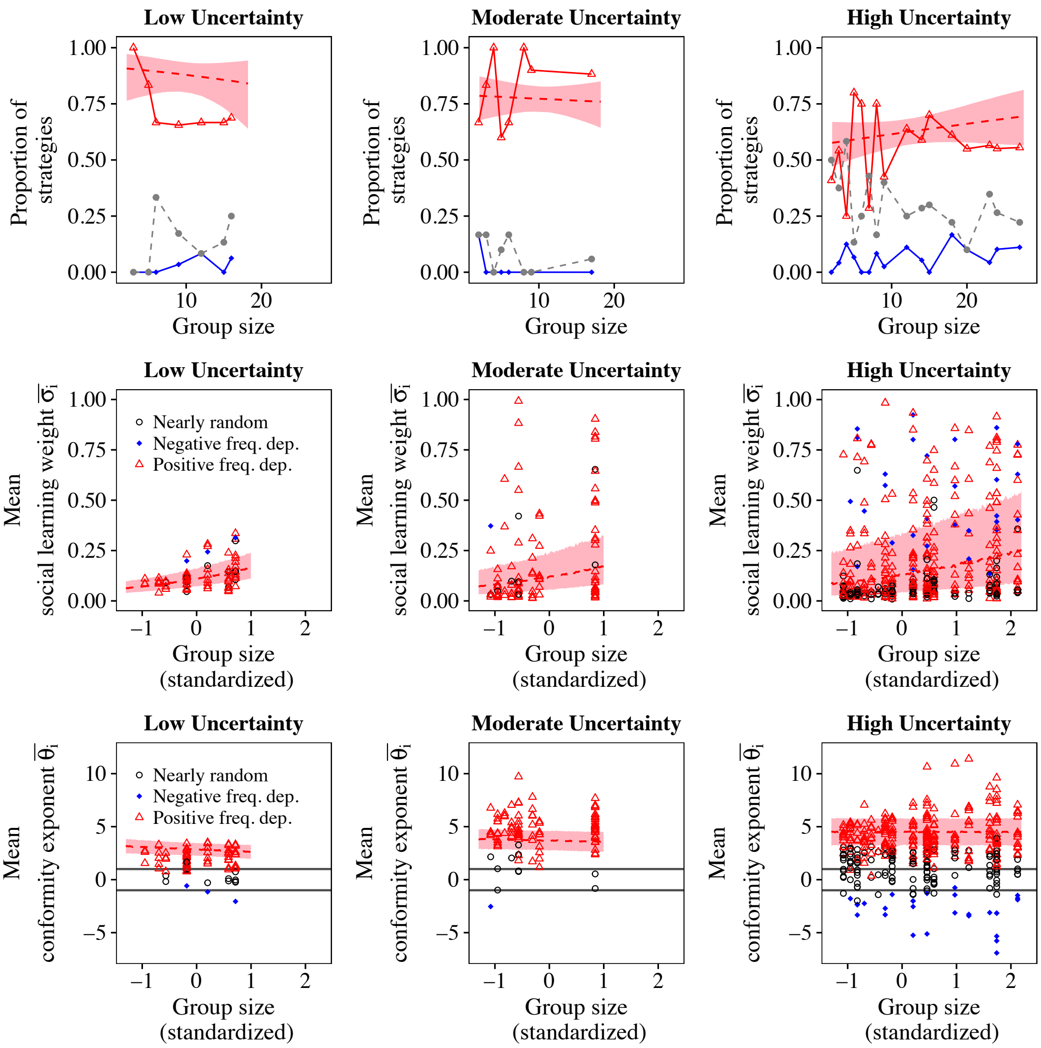

Figure S25. Model fitting for the three different task uncertainty conditions (the Low-, Moderate- and High-uncertainty) and the different group size. Three different learning strategies are shown in different styles (red-triangle: positive frequency-dependent learning, blue-circle: negative frequency-dependent learning; grey-circle: nearly random choice strategy). Note, we averaged individual conformity exponent $\theta_{i,t}$ over time to categorise individual strategies. (Top row) Frequencies of three different learning strategies. (Middle row) Estimated social learning weight, and (Bottom row) estimated mean conformity exponent, for each individual shown for each learning strategy. The 50% Bayesian CIs of the fitted GLMMs are shown by dashed lines and shaded areas. The horizontal lines in (g-i) show a region $-1 < \bar{\theta}_i < 1$.

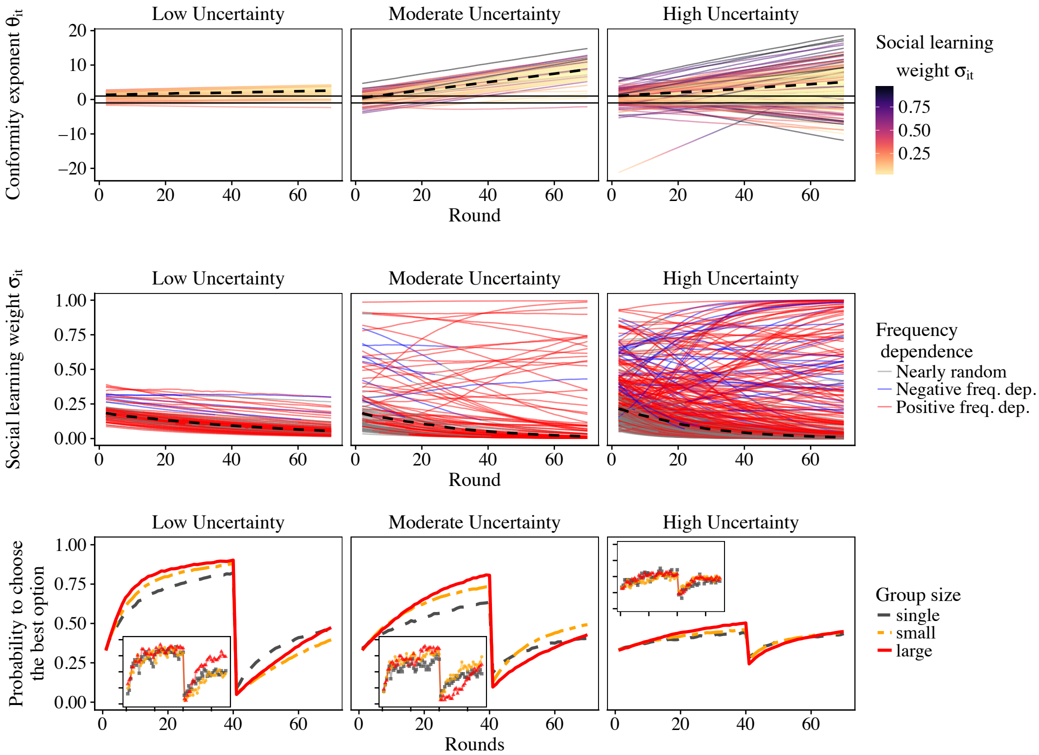

Figure S26. Change in fitted values (i.e. median of the Bayesian posterior distribution) of (Top row) the conformity exponent $\theta_{i,t}$ and (Middle row) the social learning weight $\sigma_{i,t}$ with time for each individual, for each level of task uncertainty. Thick dashed lines are the median values across the subjects for each uncertainty condition. (Bottom row) Change in average decision accuracy of the individual-based post-hoc model simulations using the experimentally fit parameter values of the alternative model (main panels). The inner panels show the average decision accuracies of the experimental participants. Each line indicates different group-size categories (red-solid: large groups, orange-halfdashed: small groups, grey-dashed: lone individuals). All individual performances were averaged within the same size category. The large or small groups were categorised using the median sizes for each experimental condition, i.e. small groups were: $n \leq 9$, $n \leq 6$ and $n \leq 11$ for the Low-, Moderate- and High-uncertain conditions, respectively.

3.5 Statistical analyses for the experimental data

Table S2

Mean and the 95% Bayesian credible intervals of the fixed effects in the GLMM predicting the probability to become a positive frequency-dependent copier. The sized effects whose CI are either below or above zero (i.e. significant) are shown in bold face.

Effect	2.5%	50%	97.5%	Effective sample size	Rhat
β_1 **(intercept)**	1.05	1.71	2.50	667	1.01
β_2 (group size)	-0.94	-0.05	0.87	2744	1.00
β_3 **(uncertainty)**	-1.88	-1.02	-0.25	1548	1.00
β_4 (age)	-0.12	0.43	1.10	925	1.01
β_5 (gender)	-1.06	-0.13	0.84	3154	1.00
β_6 (size*uncrtn)	-0.72	0.24	1.19	2880	1.00
β_7 (size*age)	-0.16	0.08	0.32	1869	1.01
β_8 (size*gndr)	-0.37	0.03	0.44	4875	1.00
β_9 **(uncrtn*age)**	-1.46	-0.73	-0.15	3167	1.00
β_{10} (uncrtn*gndr)	-1.10	0.02	1.09	3661	1.00
β_{11} (age*gndr)	-0.39	-0.02	0.37	2712	1.00

Table S3

Mean and the 95% Bayesian credible intervals of the fixed effects in the GLMM predicting individual parameter values of the social learning weight $\bar{\sigma}_i$. The sized effects whose CI are either below or above zero (i.e. significant) are shown in bold face.

Effect	2.5%	50%	97.5%	nEff	Rhat
β_1 **(intercept)**	-2.32	-2.09	-1.84	4959	1.00
β_2 **(group size)**	0.15	0.52	0.93	5230	1.00
β_3 **(uncertainty)**	-0.98	-0.59	-0.22	4784	1.00
β_4 **(age)**	-0.36	-0.18	-0.02	2126	1.00
β_5 (gender)	-0.45	-0.16	0.13	4513	1.00
β_6 (size*uncrtn)	-0.57	-0.10	0.34	5440	1.00
β_7 (size*age)	-0.19	-0.02	0.14	5359	1.00
β_8 (size*gndr)	-0.32	-0.01	0.30	4127	1.00
β_9 (uncrtn*age)	-0.17	0.07	0.32	4088	1.00
β_{10} (uncrtn*gndr)	-0.37	0.12	0.62	4205	1.00
β_{11} (age*gndr)	-0.09	0.12	0.35	4963	1.00
γ **(uncertainty effct on variance)**	1.11	1.38	1.62	3067	1.00

Table S4

Mean and the 95% Bayesian credible intervals of the fixed effects in the GLMM predicting the social learning weight $\bar{\sigma}_i$ for the positive frequency-biased choice individuals only. The sized effects whose CI are either below or above zero (i.e. significant) are shown in bold face.

Effect	2.5%	50%	97.5%	nEff	Rhat
β_1 **(intercept)**	-2.42	-2.17	-1.91	5601	1.00
β_2 **(group size)**	0.09	0.47	0.90	4509	1.00
β_3 (uncertainty)	-0.75	-0.28	0.17	6011	1.00
β_4 (age)	-0.33	-0.14	0.04	5796	1.00
β_5 (gender)	-0.36	-0.03	0.30	6075	1.00
β_6 (size*uncrtn)	-0.55	-0.01	0.49	5410	1.00
β_7 (size*age)	-0.27	-0.06	0.14	6022	1.00
β_8 (size*gndr)	-0.42	-0.05	0.33	6174	1.00
β_9 (uncrtn*age)	-0.36	-0.04	0.29	6483	1.00
β_{10} (uncrtn*gndr)	-0.75	-0.13	0.49	4746	1.00
β_{11} (age*gndr)	-0.16	0.10	0.36	5927	1.00
γ **(uncertainty effct on variance)**	1.14	1.50	1.80	5729	1.00

Table S5

Mean and the 95% Bayesian credible intervals of the fixed effects in the GLMM predicting individual parameter values of the conformity exponent θ_i. The sized effects whose CI are either below or above zero (i.e. significant) are shown in bold face.

Effect	2.5%	50%	97.5%	nEff	Rhat
β_1 **(intercept)**	1.30	1.64	2.01	2571	1.00
β_2 (group size)	-0.69	-0.17	0.35	5443	1.00
β_3 **(uncertainty)**	0.38	0.90	1.41	2602	1.00
β_4 (age)	-0.19	0.07	0.33	2967	1.00
β_5 (gender)	-0.34	0.10	0.54	3557	1.00
β_6 (size*uncrtn)	-0.40	0.19	0.79	5317	1.00
β_7 (size*age)	-0.27	-0.06	0.14	5172	1.00
β_8 (size*gndr)	-0.24	0.13	0.50	5167	1.00
β_9 (uncrtn*age)	-0.59	-0.26	0.07	3436	1.00
β_{10} (uncrtn*gndr)	-0.86	-0.21	0.45	3509	1.00
β_{11} (age*gndr)	-0.30	-0.02	0.27	4885	1.00
γ **(uncertainty effct on variance)**	1.07	1.31	1.54	4178	1.00

Table S6

Mean and the 95% Bayesian credible intervals of the fixed effects in the GLMM predicting the conformity exponent θ_i for the positive frequency-biased choice individuals only. The sized effects whose CI are either below or above zero (i.e. significant) are shown in bold face.

Effect	2.5%	50%	97.5%	nEff	Rhat
β_1 **(intercept)**	1.74	2.00	2.29	4922	1.00
β_2 (group size)	-0.40	0.03	0.42	5695	1.00
β_3 **(uncertainty)**	1.20	1.64	1.04	4381	1.00
β_4 (age)	-0.32	-0.13	0.05	6046	1.00
β_5 (gender)	-0.40	-0.07	0.26	5988	1.00
β_6 (size*uncrtn)	-0.47	0.00	0.50	4458	1.00
β_7 (size*age)	-0.24	-0.07	0.11	5716	1.00
β_8 (size*gndr)	-0.12	0.19	0.51	5349	1.00
β_9 (uncrtn*age)	-0.15	0.10	0.37	6424	1.00
β_{10} (uncrtn*gndr)	-0.53	-0.01	0.51	5710	1.00
β_{11} (age*gndr)	-0.14	0.09	0.33	6545	1.00
γ **(uncertainty effct on variance)**	0.71	0.91	1.10	6545	1.00

Table S7

Mean and the 95% Bayesian credible intervals of the fixed effects in the GLMM predicting individual parameter values of the learning rate α_i. The sized effects whose CI are either below or above zero (i.e. significant) are shown in bold face.

Effect	2.5%	50%	97.5%	nEff	Rhat
β_1 **(intercept)**	0.12	0.67	1.23	4887	1.00
β_2 (group size)	-0.49	0.24	0.96	5413	1.00
β_3 (uncertainty)	-0.93	-0.27	0.38	4827	1.00
β_4 (age)	-0.48	-0.03	0.40	5794	1.00
β_5 (gender)	-0.40	0.38	1.17	4908	1.00
β_6 (size*uncrtn)	-1.24	-0.48	0.29	5423	1.00
β_7 (size*age)	-0.25	-0.04	0.17	6157	1.00
β_8 (size*gndr)	-0.25	0.15	0.54	6305	1.00
β_9 (uncrtn*age)	-0.09	0.38	0.85	6085	1.00
β_{10} (uncrtn*gndr)	-1.21	-0.29	0.65	4699	1.00
β_{11} (age*gndr)	-0.35	-0.01	0.34	5824	1.00

Table S8

Mean and the 95% Bayesian credible intervals of the fixed effects in the GLMM predicting individual parameter values of the average inverse temperature $\bar{\beta}_i$. The sized effects whose CI are either below or above zero (i.e. significant) are shown in bold face.

Effect	2.5%	50%	97.5%	nEff	Rhat
β_1 **(intercept)**	3.09	3.47	3.85	5906	1.00
β_2 (group size)	-0.48	0.03	0.54	5707	1.00
β_3 (uncertainty)	-0.87	-0.43	0.02	5863	1.00
β_4 (age)	-0.49	-0.21	0.08	4498	1.00
β_5 (gender)	-0.35	0.16	0.67	5845	1.00
β_6 (size*uncrtn)	-0.73	-0.17	0.37	5503	1.00
β_7 (size*age)	-0.20	-0.06	0.08	6454	1.00
β_8 **(size*gndr)**	0.02	0.26	0.50	6492	1.00
β_9 **(uncrtn*age)**	0.01	0.33	0.63	6167	1.00
β_{10} (uncrtn*gndr)	-1.19	-0.58	0.02	5718	1.00
β_{11} (age*gndr)	-0.33	-0.10	0.12	5558	1.00