Collective Intelligence Under a Volatile Task Environment: a Behavioral Experiment Using Social Networks and Computer Simulations

Tatsuya Kameda (tkameda@l.u-tokyo.ac.jp)
University of Tokyo

Aoi Naito
University of Tokyo

Naoki Masuda
University at Buffalo, State University of New York

Research Article

Keywords:

Posted Date: January 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1230730/v1

Collective intelligence under a volatile task environment: A behavioral experiment using social networks and computer simulations

Aoi Naito[1,2], Naoki Masuda[3,4] and Tatsuya Kameda[1,5,6*]

[1]Department of Social Psychology, The University of Tokyo, Tokyo, Japan

[2]Japan Society for the Promotion of Science, Tokyo, Japan

[3]Department of Mathematics, University at Buffalo, State University of New York, NY, USA

[4]Computational and Data-Enabled Sciences and Engineering Program, University at Buffalo, State University of New York, NY, USA

[5]Brain Science Institute, Tamagawa University, Tokyo, Japan

[6]Center for Experimental Research in Social Sciences, Hokkaido University, Hokkaido, Japan

*Corresponding author at Department of Social Psychology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan. Email address: tkameda@l.u-tokyo.ac.jp (T. Kameda)

Abstract

Collective intelligence in our highly-connected world is a topic of interdisciplinary interest. Previous research has demonstrated that social network structures can affect collective intelligence, but the potential network impact is unknown when the task environment is volatile (i.e., optimal behavioral options can change over time), a common situation in modern societies. Here, we report a laboratory experiment in which a total of 250 participants performed a "restless" two-armed bandit task either alone, or collectively in a centralized or decentralized network. Although both network conditions outperformed the solo condition, no sizable performance difference was detected between the centralized and decentralized networks. To understand the absence of network effects, we analyzed participants' behavior parametrically using an individual choice model. We then conducted exhaustive agent-based simulations to examine how different choice strategies may underlie collective performance in centralized or decentralized networks under volatile or stationary task environments. We found that, compared to the stationary environment, the difference in network structure had a much weaker impact on collective performance under the volatile environment across broad parametric variations. These results suggest that structural impacts of networks on collective intelligence may be constrained by the degree of environmental volatility.

Introduction

Collective intelligence is an emergent phenomenon whereby a group of individuals provide a better solution to a cognitive problem than any single individual can[1,2]. Since Francis Galton's demonstration that a simple statistical aggregation of lay people's numeric estimates outperformed experts' individual estimates[3], collective intelligence has attracted broad interest across social[4–6], biological[7–9] and information sciences[10,11].

Research on human group decision-making has revealed several key factors underlying collective intelligence[1]. One critical factor is the degree of mutual independence between members. Previous laboratory studies using the Galtonian paradigm have often asked participants to independently provide estimates of a target quantity (e.g., the number of marbles in a jar) first, then aggregated them statistically into a collective estimate to compare its accuracy with the best individual's estimate. Aggregation via mode[12,13], mean[14,15], or median[16] typically outperforms the best individual in the group. Theories explaining mechanisms of such collective intelligence, including the Condorcet jury theorem in social sciences[13,17], the many wrongs principle[18], and the information center hypothesis[19] in biology, all agree that cancellation of individual informational errors via aggregation underpins improved decision-making accuracy at the group level. In real life, however, it is implausible to assume that people's judgements are independent from each other, particularly in modern societies where people are connected closely and influence each other through various information and communication technologies. Indeed, available evidence remains equivocal about when and how social influence through interaction can yield collective intelligence above and beyond the mere aggregation of independent inputs, with some studies reporting no or negative impacts[20–22], while others report positive impacts[6,23].

Becker and colleagues pointed out that social network structure may serve as an important mediator of the emergence of collective intelligence in the real world[24]. In their experiments, differences in distributions of members' interaction opportunities between centralized and decentralized networks had a large impact on the average accuracy of individual estimates. In the decentralized network, individuals generally improved their estimates by observing the estimates of direct neighbors. On the other hand, in the centralized network, the average accuracy of individual estimates depended critically on the performance of the person who occupied the single node (which we will call the central node), which had far more connections than the other nodes (which we will call the peripheral nodes). Other recent studies also demonstrated that social network structure can affect collective intelligence by constraining patterns of social influence[25–27]. However, it still remains unclear how such network effects are underpinned by micro-level decision processes of the individuals inhabiting the network, who may adjust their strategies depending on the task[28].

To shed light on such computational processes that may underlie individual decision making in the network, we focus on a situation where people work collectively on a temporally volatile task. Here, temporal volatility means that the decision environment and optimal choice behavior are not stationary but can change over time, as is often the case in political, economic, and natural situations. We conjecture that, in a volatile environment, decision makers must use social information more carefully than in a stationary environment[29]. In a stationary decision environment (e.g., estimating the weight of an ox as investigated in Galton, 1907), group-level aggregation often cancels out random noise in individual estimates[16,18]. Hence, using social information generally helps to refine an individual's estimate, particularly in a large group[5,13]. On the other hand, this is not

necessarily the case in a volatile environment, where social information can become outdated due to subsequent changes in the environment. Thus, individuals must flexibly adjust their decision strategy according to the degree of volatility. Specifically, in volatile environments, they may need to decrease reliance on social information and increase the weight given to individual learning[29–33].

In this study, we examined the emergence of collective intelligence in different social network structures in a volatile task setting. We conjectured that decision makers would decrease their reliance on social information in the volatile environment as compared to the static environment, and that, as a result, the impacts of different social network structures on collective intelligence, which have been repeatedly observed in the previous research using static tasks, may be lessened. Here, we tested this conjecture by combining a behavioral experiment and agent-based simulations.

First, we examined the effect of network structure using a laboratory experiment in which participants worked collectively on a modified version of the two-armed bandit task ("restless" two-armed bandit[33,34]). In this laboratory task, participants repeatedly chose between two "slot machines" and obtained rewards from their individual choices (Fig. 1a). We introduced volatility by changing the mean reward of one slot machine several times over the trials (Fig. 1b). As in the previous study[24], we implemented two networks composed of 10 nodes in the laboratory — centralized and decentralized, in terms of the degree distribution of the nodes (i.e., the number of neighbors that each node has). Participants also performed the task individually, which provided a baseline performance (Fig. 1c). Besides comparing performance across the three conditions, we also introduced a decision model inspired by a standard reinforcement learning model[33,35–37]. This modeling approach allowed

us to assess the decision strategies of each participant in each network under the volatile environment.

Next, we conducted a series of agent-based simulations to explore the robustness of our behavioral results for different task settings. We compared the performance patterns observed under the volatile environment (the restless two-armed bandit task) and those under the stationary environment (the regular two-armed bandit task). We explored possible interaction of the network structure (centralized vs. decentralized) and the task environment (volatile vs. stationary) using simulations in which we systematically varied the parameter values of the individual decision model.

Results

Behavioral experiment

In each laboratory session, ten participants worked on a restless two-armed bandit task. Participants chose between two “slot machines” and obtained rewards from their individual choices repeatedly over a total of 80 trials (Fig. 1a). The total number of trials was unknown to the participants. As shown in Fig. 1b, the mean reward from one “slot machine” (hereafter option) changed during the 80 trials, whereas the mean reward from the other option was held constant throughout the 80 trials. Participants were initially informed that the relative monetary rewards of the two slot machines would change over trials but were not informed about how often or when. Accordingly, participants had to repeatedly consider which slot machine would be better during the 80 trials to maximize their total reward.

In each laboratory session, we ran three within-participant conditions: the solo, centralized network, and decentralized network conditions (Fig. 1c). We counterbalanced the order of the three conditions across sessions. In the solo condition, the participants performed the task alone without information about others’ choices. In the centralized and decentralized network conditions, the participants were randomly assigned to the 10 nodes of each network. In each trial except the first, the participants in the network conditions were provided with social information in the form of the count of their immediate neighbors who had chosen each of the two options in the preceding trial. The participants received this social information before making choices in the current trial (Fig. 1a). The participants had no bird’s-eye knowledge about the network structure and remained anonymous to each other throughout the experiment. The total number of edges was equal between the centralized and decentralized networks. In the decentralized network, the number of neighbors was identical

(= 3) for all 10 nodes. In the centralized network, one central node had 9 neighbors, and the other nine peripheral nodes had either 2 or 3 neighbors (Fig. 1c; see Methods).

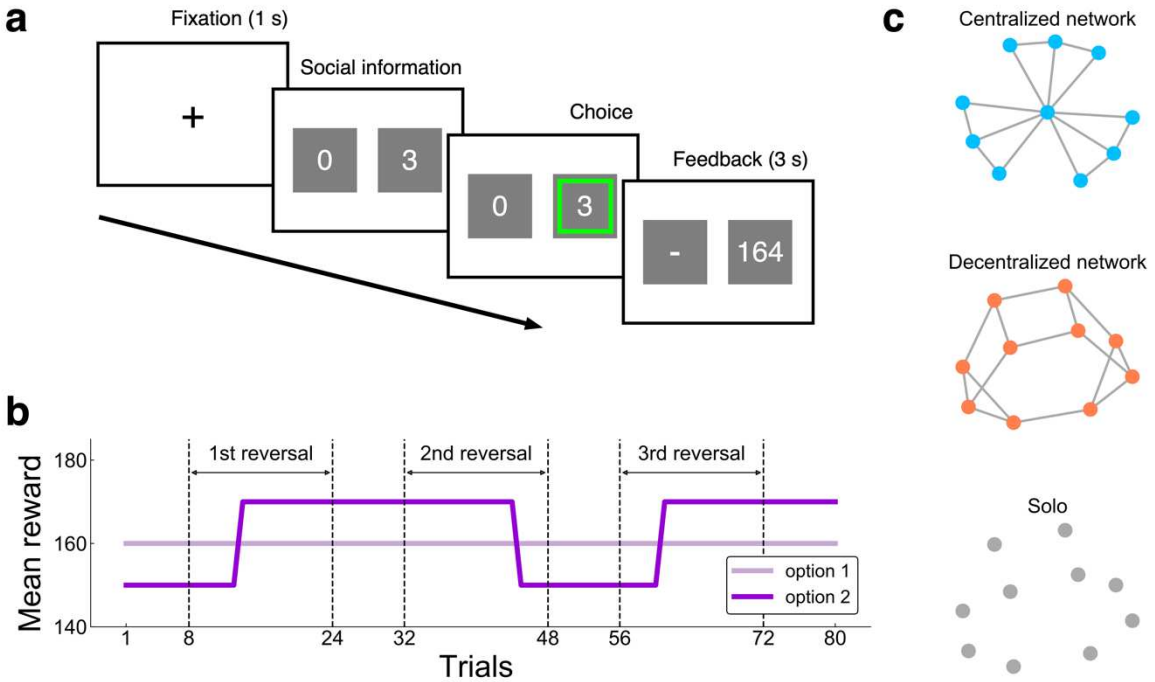

Fig. 1 | Design of the behavioral experiment. (a) Flow within each trial. The two shaded boxes on the display represent two choice options ("slot machines"). In the two network conditions (panel **(c)**), participants were first shown the frequencies of their immediate neighbors who had chosen each option in the last trial. In this example, the participant had three neighbors, all of whom had chosen option 2. No social information was provided in the solo condition. Each participant then proceeded to the choice stage, made their choice, and the chosen option was highlighted with the green frame. In the present example, the participant chose option 2 in the current trial, and received a reward (164 points) in the feedback stage. This process was repeated for a total of 80 trials in each condition. **(b)** Time course of the mean reward of the two options. In this example, the mean reward from option 1 was constant, whereas that from option 2 changed three times during the 80 trials. In option 2, we generated the timings of the change randomly from a uniform distribution, with the restriction that the first reversal in the relative payoffs of the two options occurred between

the 8th and the 24th trials, the second reversal between the 32nd and the 48th trials, and the third reversal between the 56th and 72nd trials. **(c)** Three conditions in the experiment. The order of the three conditions was counterbalanced across laboratory sessions.

Overall performance. We defined a participant's overall performance in each condition as the number of trials (out of 80) in which the participant selected the option with the larger mean reward. We first compared the participants' performance in the three conditions using a binomial GLMM (see Methods for details). Figure 2a indicates that the participants performed better in the two network conditions than in the solo condition on average ($z = 2.97$, $p < 0.05$ for the centralized network; $z = 3.99$, $p < 0.05$ for the decentralized network). This result confirms that opportunities for social learning improved collective performance in the volatile environment, which is in line with the results of previous studies using stationary environments[24,26,38]. However, no significant difference was detected between the centralized and decentralized network conditions ($z = 1.02$, $p = 0.31$). Figure 2b further indicates that the number of neighbors of a participant in the centralized network (2, 3, or 9; see Fig. 1c) had no significant impact on performance ($z = 0.03$, $p = 0.57$). These behavioral results suggest that social-learning opportunities improved collective performance under the volatile environment as well, but that its effect was not dependent on the network structure or the number of neighbors (or position in the network) for each participant.

Individual participants' decision strategies. To examine more closely how social information affected participants' decisions, we analyzed their trial-by-trial behaviors using the following choice model[33–35,39]. We assume that participant i chooses option 1 and option 2 at trial t with probability $P_{i,1,t}$ and $1 - P_{i,1,t}$, respectively. We also assume that $P_{i,1,t}$ is a weighted sum of two conditional choice probabilities, one using social information, and the

other using participant i's own past experience (i.e., reinforcement learning implemented as Q-learning):

$$P_{i,1,t} = \gamma\, P_{i,1,t}^{soc} + (1-\gamma)\, P_{i,1,t}^{rein}, \tag{1}$$

where $\gamma \in [0, 1]$. In Eq. (1), we assume that $P_{i,1,t}^{soc}$ (i.e., the choice probability conditioned on the participant using social information) increases monotonically with the number of the neighbors of i choosing option 1 in the (t-1)th trial (see Eqs. (2) and (3) in Methods). This part of the model represents a social learning mechanism. We also assume that $P_{i,1,t}^{rein}$ (i.e., the choice probability conditioned on the participant using their own past experience) follows a standard reinforcement learning model[40] (see Eqs. (4) and (5) in Methods).

Under these assumptions, the model has four parameters, denoted by α, β, and θ, besides the aforementioned γ that regulates the weighting between social information and past experience (see Methods for details). In the social learning part of the model (Eqs. (2) and (3)), θ represents conformity bias[41] such that a participant with a larger θ favors the option chosen by the majority in the last trial disproportionately more than expected from its proportion among neighbors. In the reinforcement-learning part of the model (Eqs. (4) and (5)), α specifies the learning rate, i.e., the speed at which the participant updates the value of the selected option[42], and β reflects the amount of noise when the participant makes a choice, also known as the "inverse temperature" parameter[43].

We first compared our model (which we refer to as the full model hereafter) and four sub-models with reduced sets of parameters to identify the best model that predicted the behavioral data (see Supplementary Information for details). We fitted all candidate models

using a hierarchical Bayesian model-fitting method (HBM), with which we estimated the parameter values for the individual participants by drawing them from a group-level distribution that constrained the range of each parameter for the individual participants. This procedure allowed us to simultaneously estimate the group-level estimates of the parameters and the estimates for the individual participants. We compared the five models by their widely applicable information criterion (WAIC) values. In line with the previous literature[33,34], we found that the full model provided the best fit to the behavioral data in terms of the WAIC (see Supplementary Table S2). To assess the fit of the full model, we endowed all agents with the group-level means of the four estimated parameters and ran agent-based simulations under the same setup as in the behavioral experiment. We found that the two network conditions yielded similar mean performance values and that the mean performance under these two conditions was larger than under the solo condition (see Supplementary Fig. S1), qualitatively replicating the behavioral results. Therefore, we used the full model in the following analyses and agent-based simulations.

The HBM allowed us to directly compare parameter values between the two network conditions through the estimation of the group-level parameters in each condition. For each of the four group-level parameters, the distribution was not significantly different between the centralized and decentralized network conditions, as summarized in Table 1. To further support this result, Fig. 2c shows the estimate of the four parameters for each individual and their posterior distributions in each of the two network conditions. We also found that, for the individual participants, all four parameters were significantly correlated between the two network conditions, although the correlation values were small except for α (α: $r = 0.61$, $p = 0.00$; β: $r = 0.27$, $p = 0.00$; γ: $r = 0.16$, $p = 0.01$; θ: $r = 0.21$, $p = 0.00$; the Pearson's

correlation coefficient and its p value are denoted by r and p, respectively; see Supplementary Fig. S1). This result implies that the participants adopted significantly similar decision strategies under the two network conditions.

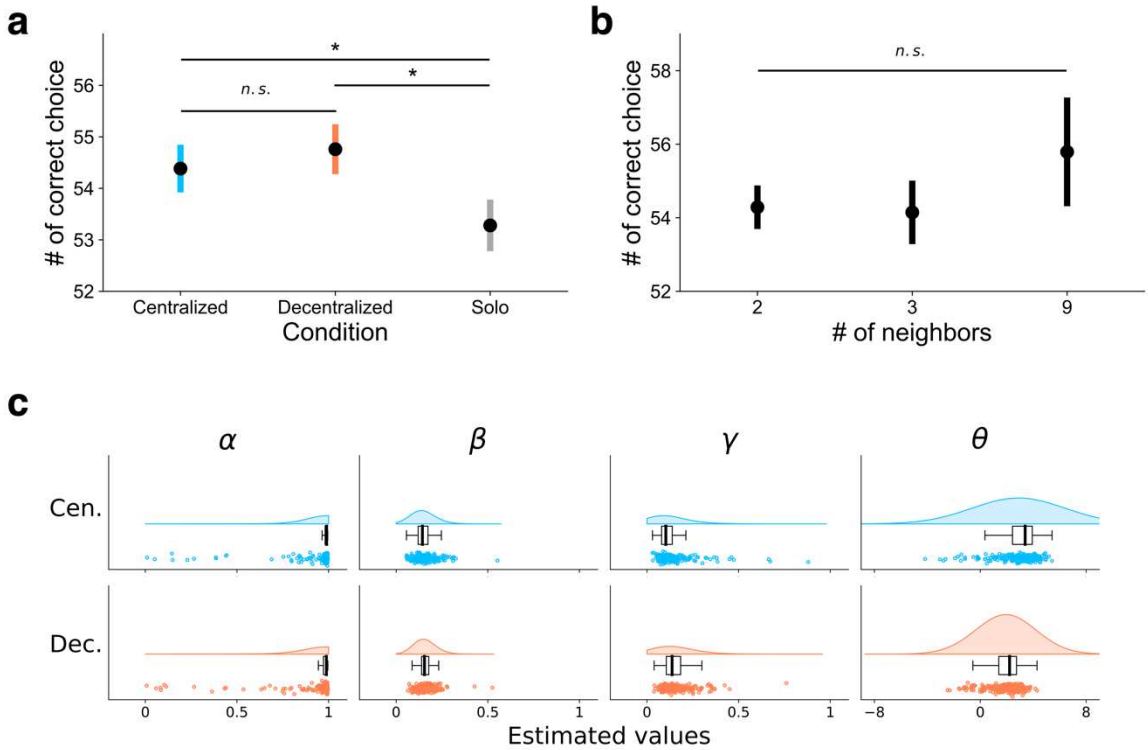

Fig. 2 | Results from the behavioral experiment. (a) Mean performance (i.e., the number of trials in which a participant chose the option with the larger mean reward) in the three conditions. The filled circle indicates the mean. The error bar indicates the standard error of the mean. * $p < 0.05$. **(b)** Mean performance of the participant conditioned on the number of neighbors (2, 3, or 9) in the centralized network condition. **(c)** Results of parameter estimations. The probability density functions are those of the group-level parameters estimated by the HBM. Each circle represents the estimate for one participant. Box plots show five-number summaries of the distributions; these quantities are the first quartile (Q_1), the median, the third quartile (Q_3), the minimum without outliers ($Q_1 - 1.5 \times \mathrm{IQR}$), and the maximum without outliers ($Q_3 + 1.5 \times \mathrm{IQR}$), where $\mathrm{IQR} = Q_3 - Q_1$. We abbreviated the centralized and decentralized networks as Cen. and Dec., respectively.

Table 1 | Mean and 95% confidence intervals of the posterior distributions of group-level parameter values. The upper rows in each cell show the mean values, and the lower rows show the 95% confidence intervals (CIs). When the interval does not cross zero, the parameter is significantly different between the two conditions.

	Network		Difference
	Centralized	Decentralized	
α (i.e., learning rate)	0.996 [0.993, 0.999]	0.996 [0.992, 0.999]	0.000 [-0.007, 0.008]
β (i.e., inverse temperature)	0.139 [0.129, 0.150]	0.151 [0.140, 0.161]	-0.011 [-0.025, 0.005]
γ (i.e., social learning weight)	0.094 [0.071, 0.117]	0.128 [0.103, 0.152]	-0.034 [-0.066, 0.0003]
θ (i.e., conformity coefficient)	3.08 [1.472, 4.696]	1.95 [1.336, 2.558]	1.137 [-0.634, 3.068]

Agent-based simulations

Thus far, the behavioral experiment and the model analyses have yielded two major results. First, social learning opportunities improved collective performance. Second, the two network conditions (centralized and decentralized) revealed no sizable difference in terms of the collective performance and the values of the estimated decision parameters under the volatile task. The latter result is in stark contrast to the results of previous studies of stationary tasks[24,26]. In these previous studies, participants who worked collectively on stationary tasks (e.g., estimating a target numerical quantity) exhibited different group performance depending on the network structure (centralization[24]; efficiency[26]).

To further investigate the effects of the network structure, task environment, and the individual participant's choice strategies (i.e., how to combine social information and past experience) on collective performance, we carried out a series of agent-based simulations.

Comparing the volatile and stationary environments for the effect of network centralization on collective performance. We assessed how the participants' behavior in the volatile task in our experiment may relate to behavior when the task is stationary. We note that a previous experimental study with a stationary task used the same centralized and decentralized networks as those in the present study[24]. We implemented the stationary task using a two-armed bandit task with a fixed mean reward for each option: 150 ± 10 (mean ± std) for option 1 and 160 ± 10 for option 2. In the simulations of the volatile task, we used the same task settings as in the behavioral experiment (Fig. 1b). Each agent was assumed to independently select either option with the same probability (i.e., $P_{i,1,1} = P_{i,2,1} = 1/2$) in the first trial, and then obeyed the full model in the remaining 79 trials.

We compared the mean performance of the agents between the centralized and decentralized networks (Fig. 1c), under the stationary and volatile environments and for various parameter values of the individual choice model (Eqs. (1)-(5)). Of the four parameters of our full model, we focused on α (the rate of learning from the agent's own experience) and γ (reliance on social learning). To recall, when $\alpha \in [0, 1]$ is large, an agent updates the Q-value for each option rapidly (Eq. (4)). When $\gamma \in [0, 1]$ is large, an agent tends to rely on other agents' decisions in the preceding round (Eq. (1)). In the social learning literature, people are known to adjust these two parameters in response to whether the task is volatile or stationary[29–32]. Accordingly, we scanned the full range of both parameters, i.e., from 0 to 1 with a step size of 0.1, yielding a total of 121 (= 11×11) combinations of the parameter values. We assigned the same α and γ values to all agents. We ran 1,000 simulations for each of the 121 combinations of the parameter values, each network (centralized and decentralized), and each task environment (volatile and stationary). Then, we calculated the average of the collective performance over the 1,000 simulations for each combination of the α and γ values. For the other two parameters (i.e., β and θ), we used the population mean of the estimates observed from the experiment (Table 1), which we averaged over the two network conditions. We also ran a full factorial (11×11×11×11) simulation by varying the values of all four parameters to check the robustness of the following results (see the Additional Simulations section in Supplementary Information).

The mean performance of the 10 agents in the volatile and stationary environments are shown in Figs. 3a and 3b, respectively. In each figure, the left panel shows the results for the centralized network, the middle panel for the decentralized network, and the right panel

for the performance ratio defined by the performance of the decentralized network divided by that of the centralized network. We make the following three observations.

First, not surprisingly, for both networks, the mean performances under the volatile environment (see the left and middle panels in Fig. 3a) were generally lower than those under the stationary environment (see the left and middle panels in Fig. 3b; see also Supplementary Fig. S2).

Second, in both volatile and stationary environments, and in both centralized and decentralized networks, the mean performance was large when α was large and γ was intermediate. This result suggests that the combination of relatively rapid value-updating (i.e., a large α value) and a roughly equal frequency of using social learning and individual learning (i.e., an intermediate γ value) enhances collective performance. In the behavioral experiment with the volatile task, we observed that the group-level mean of α was approximately equal to 0.996 for both centralized and decentralized networks and that the group-level means of γ were 0.094 and 0.128 for the centralized and decentralized network, respectively (see Table 1). Therefore, according to our simulation results, participants in the experiment could have improved their performances slightly if they had relied more on social information, both in the centralized and decentralized networks. Nonetheless, we point out that their actual performances were fairly good, if not optimal, as shown in the comparison with the full-factorial simulation; see Supplementary Fig. S2.

Third, the simulations revealed different magnitudes of the network effect between the volatile and stationary tasks, as shown in the right panels of Figs. 3a and 3b. The figures show the ratio of the performance in the decentralized network to that in the centralized network. A ratio far from 1 implies that the impact of the network structure on collective

performance is large. We find that the network structure has a much weaker impact on collective performance (i.e., the ratio is closer to 1) under the volatile environment (see the right panel in Fig. 3a) than the stationary environment (see the right panel in Fig. 3b) in a broad parameter region. See also Supplementary Fig. S2 for the same observation in the full-factorial simulation. Furthermore, in the volatile environment, the impact of the network structure is negligible for small γ values. This numerical result is consistent with the results of our behavioral experiment, in which most participants relied little on social information, yielding small γ values (see the distribution of γ in Fig. 2c), and the collective performance was negligibly different between the centralized and decentralized networks.

Taken together, these simulation results may provide a possible explanation for the lack of sizable difference in terms of collective performance between the centralized and decentralized networks in the behavioral experiment under the volatile task (Fig. 2a), in contrast to the sizable network effect reported in the previous study[24] using a stationary task.

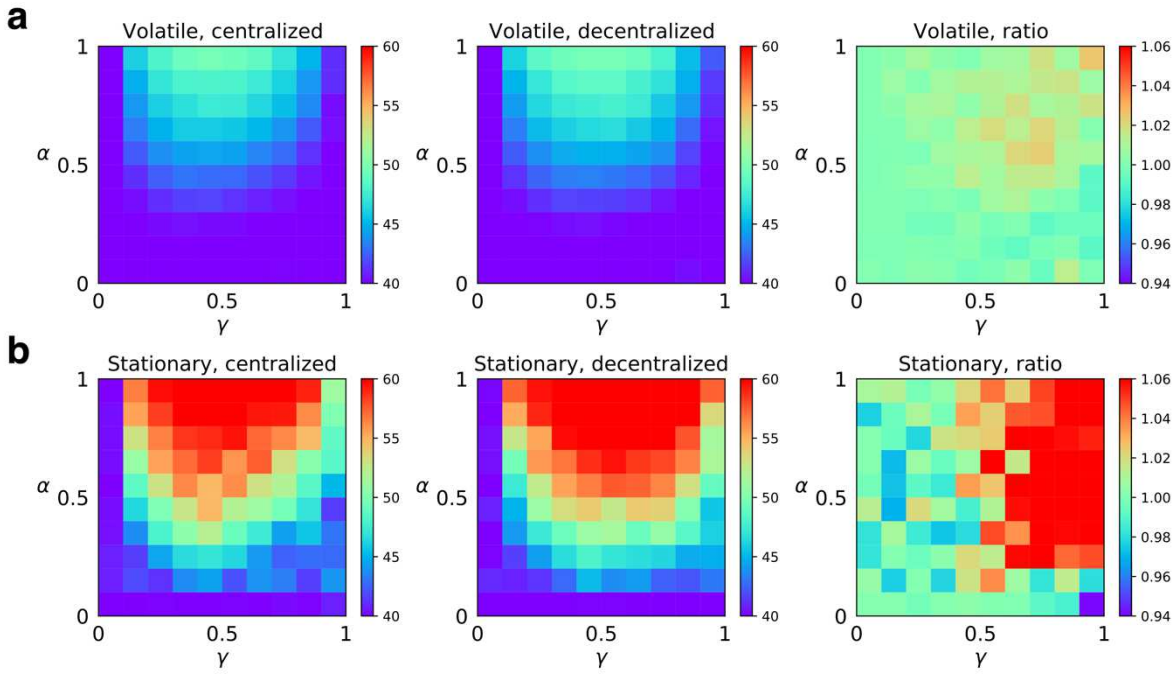

Fig. 3 | Results of agent-based simulations. Mean performance for the two network conditions in **(a)** the volatile task environment and **(b)** the stationary task environment. In both task environments, we ran 1,000 simulations for each value of α and γ. The left panel shows the mean performance in the centralized network, the middle panel shows that in the decentralized network, and the right panel shows the ratio of the performance in the decentralized network to that in the centralized network.

Discussion

A volatile task environment, in which changes occur often and the optimal behavior "yesterday" may not be even appropriate anymore, is a characteristic feature of many social, economic, and political domains. At the same time, rapid advancement of communication technologies has been modifying people's connectivity in various networks. Understanding how collective intelligence may emerge under such volatile environments in relation to social network structures is thus a practical concern. Here, we approached this question using a laboratory behavioral experiment, statistical modeling of behavioral data, and agent-based simulations.

In the behavioral experiment, participants worked on a restless two-armed bandit task either alone, or collectively in a centralized or decentralized network. We found that, although the benefits of collective intelligence emerged in the networks compared to the solo situation, no sizable performance difference was detected between the centralized and decentralized networks (Fig. 2a). Using an individual decision model inspired by a standard reinforcement learning model[33,35,37], we observed that distributions of the behavioral parameters were indistinguishable between the centralized and decentralized network conditions (Fig. 2c). This absence of network effect is in stark contrast to the results of previous studies in which participants working on stationary tasks showed different group performance depending on the network structure (on centralization[24], and on efficiency[26]).

To understand this discrepancy between volatile and stationary task environments, we conducted a series of agent-based simulations. In the numerical simulations, we focused on possible interactions of the environment (volatile vs. stationary), network structure (centralized vs. decentralized), and individual choice strategies (how to use social information and past experience). The simulations led to three main observations (Figs. 3a

and 3b). Firstly, irrespective of network centralization, collective performance was generally better in the stationary case than in the volatile case. Secondly, again irrespective of network centralization, the combination of relatively rapid value-updating through the agent's own experience (i.e., a large α value) and roughly equally weighted mixtures of social learning and individual learning (i.e., an intermediate γ value) yields superior collective performance. Lastly, the difference in centralization of the network has a much weaker impact on collective performance in the volatile environment (Fig. 3a right) than in the stationary environment (Fig. 3b right) across a broad parameter region (see also Supplementary Fig. S2).

The past behavioral studies that employed the "restless" multi-armed bandit task[33,34,44] have consistently shown that groups can perform better than solo individuals. Emergence of such collective intelligence through interaction was replicated consistently in the present behavioral experiment as well as in the agent-based simulations. These consistent patterns confirm that opportunities for social learning are indeed beneficial not only in the stationary task but also in the volatile task. However, our results suggest that the known effect of network structure (e.g., centralization) on collective performance may be quite limited in volatile task environments.

Several questions remain unanswered in the current study.

Firstly, we tested collective performance only under a volatile task environment in the behavioral experiment. While our simulation results suggested that the difference in network centralization could impact collective performance more under the stationary task than the volatile task, we did not provide empirical evidence of this finding. Although the impacts of network structure in stationary tasks have been documented[24,26], it is desirable to replicate

such a network effect directly in an experiment using the two-armed bandit task with the fixed mean rewards as used in the simulation.

Secondly, our reinforcement learning model is an arbitrary choice. It could be replaced with other models such as Bayesian learning[45]. Although we conjecture that different statistical modeling would yield similar observations in both behavioral and numerical experiments, it would be worthwhile to conduct a robustness check with other learning algorithms.

Lastly, it remains unknown whether and how collective intelligence may be improved if individuals are allowed to modify their connections to other individuals. To our knowledge, empirical research on the relationship between collective intelligence and dynamic formation of social networks has begun only recently. For example, using a perceptual decision-making task, Almaatouq *et al.* (2020) has shown that individuals strengthened their connections with competent individuals and improved their performances[46]. As seen in Almaatouq *et al.* (2020), understanding how implicit or explicit coordination may facilitate efficient aggregation of inputs from members with different skills or expertise is becoming a key agenda in research on group decision making. Yet, the ways in which dynamic formation of networks in collective decision-making tasks[46–49], including networking with artificial bots in human-machine ensembles[50,51], may interact with adjustments of individual-level decision strategies remains largely unexplored. Given the high fluidity of communication patterns in modern networks through mobility and advanced technologies[52–54], addressing these questions should be important not only theoretically but also to help address urgent societal problems, including the spread of infectious disease, destructive political propaganda, and unfounded or malicious rumors, all of which are impacting people's welfare in the digital era.

Methods

Participants

A total of 250 undergraduate students (183 males and 67 females; mean ± s.d. age: 20.76 ± 2.12 years) participated in the behavioral experiment. We recruited 130 participants at random from the subject pool of Hokkaido University (Hokkaido, Japan) and 120 participants from the University of Tokyo (Tokyo, Japan). The experiment was approved by the Ethics Committee of Hokkaido University and the Ethics Committee of the University of Tokyo. Written informed consent was obtained from each participant before the experiment.

Experimental procedure

There were a total of 25 laboratory sessions with 10 participants each. Upon arrival, participants were seated in a semi-private cubicle with a computer terminal connected through a local area network. All instructions were displayed on a computer monitor and were simultaneously read aloud by the Japanese voice synthesizer software "Voice Sommelier Neo" (Hitachi Solutions Create, Ltd.). No direct communication was allowed between participants, who remained completely anonymous to each other throughout the experiment. A laboratory session lasted about 90 minutes. We developed the code to run the experiments using Python and the Psychopy[55] package.

In each laboratory session, there were three within-participant conditions：the solo, centralized network, and decentralized network conditions (see Fig. 1c for a schematic illustration). Ten participants worked on the restless two-armed bandit task over 80 trials under each of the three conditions, for a total of 240 trials per session. We counterbalanced

435 the order of the three conditions according to a 3 × 3 Latin square design (see Supplementary
436 Table S1). Participants were asked to choose between two options ("slot machines") to
437 maximize their individual cumulative reward over the 80 trials. The exact number of trials in
438 each condition was not known to the participants in advance.

439 As shown in Fig. 1a, each trial started with the presentation of a crossbar on the
440 display for 1 second. Next, two boxes representing the two options appeared on the display,
441 one on the left side and the other on the right side. Participants chose either box by pressing a
442 corresponding key on a keyboard. The selected option was highlighted by a green frame,
443 followed by the reward from the choice whose value appeared on the display for 3 seconds,
444 after which the next trial started.

445 In each trial, we generated a reward for each participant by independently drawing a
446 value from a normal distribution, rounding it up to an integer value if it was positive, and
447 truncating it to 0 if negative. The two options had different means but the same standard
448 deviation (= 10 points). For one option, which we refer to as option 1, we fixed its mean
449 reward at 160 points., while the mean reward for the other option, which we refer to as option
450 2, switched between 150 and 170 points three times during the 80 trials. At the start of the
451 experiment under each condition, we randomly determined which of the two options was
452 displayed to the left or right on the display with equal probability (i.e., 1/2), which remained
453 the same for all participants throughout the 80 trials. For option 2, the first switch occurred in
454 a trial that we uniformly randomly selected between the 8th and the 24th trial (i.e.,
455 probability 1/17 each), the second switch between the 32nd and the 48th trial, and the third
456 switch between the 56th and 72nd trial (see Fig. 1b). We told the participants in advance that
457 the relative average profitability of the two options would be reversed several times over the

trials. However, they were not informed of the total number of trials, the distributions of the reward, when and how often the reversals would occur, or the exchange rate of the accumulated points for compensation.

For compensation, we used different schemes between Hokkaido University and the University of Tokyo due to regional hourly wage differences. Participants from Hokkaido University were paid 1,182 JPY on average (max: 2,000; min: 1,000) and participants from the University of Tokyo were paid 1,763 JPY on average (max: 2,500; min: 1,500).

Network structure

We assigned 10 participants in each laboratory session uniformly at random to the 10 nodes of the network for the centralized and decentralized network conditions. We did not give the participants bird's eye knowledge about their network locations. In each trial, they were informed about the number of their immediate neighbors who had chosen each option in the preceding trial (Fig. 1a).

For the decentralized network condition, we used the configuration model to uniformly randomly wire edges such that each of the 10 nodes had exactly three edges. This network is known as a regular random graph. For the centralized network condition, we used a star-like network, which consisted of a single node, denoted by v, that was directly connected to all other nine nodes; each of the nine nodes had two or three neighbors including v. We uniformly randomly wired the nine nodes except for the edges formed with v a la mode de the configuration model. Accordingly, the centralized and decentralized networks had the same numbers of nodes and edges, implying that the average number of neighbors per node was the same between the two networks.

Statistical analysis

We analyzed participant performance with a binomial generalized linear mixed model (binomial-GLMM). The dependent variable was the overall performance, i.e., the number of trials in which the participant selected the currently superior option. Fixed effects included those of the decentralized network condition (yes = 1 / no = 0) and the centralized network condition (yes = 1 / no = 0) against the solo (control) condition, and the number of neighbors in the network. We entered the participant ID and the group ID as random effects.

Computational model

We fitted the following choice model to each participant's behavioral data, separately for the two network conditions[27,28,31]. We denote the probability that participant *i* selects option 1 at trial *t* by $P_{i,1,t}$. We assume that $P_{i,1,t}$ is a weighted average of two conditional choice probabilities, i.e.,

$$P_{i,1,t} = \gamma_i \, P_{i,1,t}^{soc} + (1 - \gamma_i) \, P_{i,1,t}^{rein} \tag{1}$$

where $\gamma_i \in [0, 1]$. Variables $P_{i,1,t}^{soc}$ and $P_{i,1,t}^{rein}$ represent the probability with which participant *i* selects option 1 in the *t*th trial when using social information and when using past experience, respectively.

The participant was assumed to adopt a frequency-dependent copying strategy when using social information[33–35], with which the participant stochastically copies its neighbors' choices. Specifically, we set

$$P_{i,1,t}^{soc} = \frac{(F_{i,1,t-1}+0.1)^{\theta_i}}{(F_{i,1,t-1}+0.1)^{\theta_i}+(F_{i,2,t-1}+0.1)^{\theta_i}}, \tag{2}$$

where $F_{i,k,t-1}$ (with $k = 1, 2$) is the number of participant i's neighbors choosing option k in the (t - 1)th trial ($t \geq 2$). The probability that the ith participant using social information selects option 2 in trial t, denoted by $P_{i,2,t}^{soc}$, is given by

$$P_{i,2,t}^{soc} = \frac{(F_{i,2,t-1}+0.1)^{\theta_i}}{(F_{i,1,t-1}+0.1)^{\theta_i}+(F_{i,2,t-1}+0.1)^{\theta_i}}. \tag{3}$$

It should be noted that $P_{i,1,t}^{soc} + P_{i,2,t}^{soc} = 1$. Parameter $\theta_i \in (-\infty, \infty)$ quantifies participant i's conformity bias, which is also known as positive frequency dependence[29]. When $\theta_i > 0$, the option chosen by more neighbors has a disproportionately greater influence on i's choice. In contrast, when $\theta_i < 0$, the option supported by fewer neighbors has a greater influence on i's choice. We added 0.1 to $F_{i,1,t-1}$ and $F_{i,2,t-1}$ in Eqs. (2) and (3) to allow the participant to choose an option chosen by nobody in the preceding trial.

When using past experience, the participant is assumed to follow standard reinforcement learning according to the Rescola-Wagner[32] rule. Let us denote by $Q_{i,1,t}$ and $Q_{i,2,t}$ the values of options 1 and 2 that the ith participant perceives at the tth trial,

respectively. When the *i*th participant selects option 1 in the (*t*-1)th trial, the value of option 1 is updated by

$$Q_{i,1,t} = (1 - \alpha_i)Q_{i,1,t-1} + \alpha_i R_{i,1,t-1}, \tag{4}$$

where $\alpha_i \in [0, 1]$ is the learning rate (i.e., the weight given to new experience) for the *i*th participant, and $R_{i,1,t-1}$ is the reward that the *i*th participant obtained by selecting option 1 in the (*t* -1)th trial. In this case, we do not update $Q_{i,2,t}$ (i.e., $Q_{i,2,t} = Q_{i,2,t-1}$). When the participant selects option 2 in the (*t* -1)th trial, we update $Q_{i,2,t}$ in the same manner as in Eq. (4), and $Q_{i,1,t}$ remains the same (i.e., $Q_{i,1,t} = Q_{i,1,t-1}$). We assumed that all participants had no prior preference for either option at the outset (i.e., $Q_{i,1,0} = 0$ and $Q_{i,2,0} = 0$).

When the participant uses the reinforcement learning rule, the probability of choosing option 1 at trial *t* is assumed to follow the softmax rule given by

$$P_{i,1,t}^{rein} = \frac{\exp(\beta_i Q_{i,1,t})}{\exp(\beta_i Q_{i,1,t}) + \exp(\beta_i Q_{i,2,t})}, \tag{5}$$

where $\beta_i \in (0, \infty)$, called the inverse temperature or noise parameter, determines the amount of stochasticity in the choice on the basis of $Q_{i,1,t}$ and $Q_{i,2,t}$. As $\beta_i \to 0$, the choice is independent of $Q_{i,1,t}$ and $Q_{i,2,t}$ such that the participant is maximally exploratory. As $\beta_i \to \infty$, participant *i* always selects the option with the higher value such that the participant is maximally exploitative. Therefore, β_i regulates the individual's exploration tendency[34].

Hierarchical Bayesian model-fitting (HBM)

Following examples in the literature[33,56], we fit all candidate models using a hierarchical Bayesian model-fitting method (HBM). With HBM, we can estimate the parameter values for the individual participants by assuming that they are drawn from a common group-level distribution[57,58]. Therefore, this procedure allows us to simultaneously estimate the parameter values for the group and the individuals. The concept of the group-level distribution also enables us to directly compare the parameter estimates in the two network conditions.

As explained in the results section, we adopted the full model. For each condition, c (centralized or decentralized network condition), we obtained $R_{i,1,t-1}$ (the reward given to participant i on the (t -1)th trial), $Ch_{i,t}$ (the choice of participant i on the tth trial), and $F_{i,1k,t-1}$ (the number of i's neighbors choosing option k in the (t - 1)th trial) from the behavioral data. For the sake of simplicity, we have omitted the subscript c and will continue to do so in the following text. There are four parameters per participant, namely, α_i, β_i, γ_i, and θ_i. We estimate α_i^* ($-\infty < \alpha_i^* < \infty$), γ_i^* ($-\infty < \gamma_i^* < \infty$), and β_i^* ($-\infty < \beta_i^* < \infty$), which are defined by

$$\alpha_i = \frac{1}{1 + \exp\left(-\alpha_i^*\right)}, \tag{6}$$

$$\gamma_i = \frac{1}{1 + \exp\left(-\gamma_i^*\right)}, \tag{7}$$

and

$$\beta_i = \exp(\beta_i^*), \tag{8}$$

respectively, instead of α_i $(0 < \alpha_i < 1)$, γ_i $(0 < \gamma_i < 1)$, and $\beta_i (0 < \beta_i < \infty)$ to stabilize estimation of the parameters.

To estimate $\alpha_i^*, \beta_i^*, \gamma_i^*$, and θ_i, we assumed Student's *t*-distributions for individual random effects to accommodate outliers, because Student's *t*-distribution has a longer tail than a normal distribution[33]. To estimate α_i^*, for example, we used the following reparameterization:

$$\alpha_i^* = \mu_{\alpha_i^*} + \nu_{\alpha_i^*} \varepsilon_i, \tag{9}$$

where $\mu_{\alpha_i^*}$ is a group-level mean of α_i^*, and $\nu_{\alpha_i^*}$ is a group-level scale parameter[33,56]; we estimated $\mu_{\alpha_i^*}$ and $\nu_{\alpha_i^*}$ separately for each network condition. In Eq. (9), ε_i is a standardized individual-level random variable, which we drew from a Student's *t*-distribution with degrees of freedom equal to 4, location parameter equal to 0, and scale parameter equal to 1. We drew $\mu_{\alpha_i^*}$ from a normal prior distribution with mean 0 and standard deviation 5, and $\nu_{\alpha_i^*}$ from a half-normal prior distribution with mean 0 and standard deviation 3. We estimated β_i^*, γ_i^*, and θ_i in the same way. Therefore, there are eight group-level parameters (i.e., 4 μs and 4 νs) for each network condition.

Data availability

The data used in the current study is available from the corresponding author on reasonable request.

References

1. Krause, J., Ruxton, G. D. & Krause, S. Swarm intelligence in animals and humans. *Trends Ecol. Evol.* **25**, 28–34 (2010).
2. Michelucci, P. & Dickinson, J. L. HUMAN COMPUTATION. The power of crowds. *Science* **351**, 32–33 (2016).
3. Galton, F. Vox populi. *Nature* **75**, 450–451 (1907).
4. Surowiecki, J. *The wisdom of crowds*. (Doubleday Books, 2004).
5. Kameda, T., Tsukasaki, T., Hastie, R. & Berg, N. Democracy under uncertainty: the wisdom of crowds and the free-rider problem in group decision making. *Psychol. Rev.* **118**, 76–96 (2011).
6. Jayles, B. *et al.* How social information can improve estimation accuracy in human groups. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 12620–12625 (2017).
7. Liker, A. & Bókony, V. Larger groups are more successful in innovative problem solving in house sparrows. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 7893–7898 (2009).
8. Couzin, I. D. *et al.* Uninformed individuals promote democratic consensus in animal groups. *Science* **334**, 1578–1580 (2011).
9. Morand-Ferron, J. & Quinn, J. L. Larger groups of passerines are more efficient problem solvers in the wild. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 15898–15903 (2011).

10. He, F., Pan, Y., Lin, Q., Miao, X. & Chen, Z. Collective Intelligence: A Taxonomy and Survey. *IEEE Access* **7**, 170213–170225 (2019).

11. Garattoni, L. & Birattari, M. Autonomous task sequencing in a robot swarm. *Sci. Robot.* **3**, eaat0430 (2018).

12. Kurvers, R. H. J. M. *et al.* Boosting medical diagnostics by pooling independent judgments. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 8777–8782 (2016).

13. Hastie, R. & Kameda, T. The robust beauty of majority rules in group decisions. *Psychol. Rev.* **112**, 494–508 (2005).

14. Hong, L. & Page, S. E. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 16385–16389 (2004).

15. Stock, J. H. & Watson, M. W. Combination forecasts of output growth in a seven-country data set. *J. Forecast.* **23**, 405–430 (2004).

16. Krause, S., James, R., Faria, J. J., Ruxton, G. D. & Krause, J. Swarm intelligence in humans: diversity can trump ability. *Anim. Behav.* **81**, 941–948 (2011).

17. List, C. & Goodin, R. E. Epistemic democracy: Generalizing the Condorcet jury theorem. *J. Polit. Philos.* **9**, 277–306 (2001).

18. Simons, A. M. Many wrongs: the advantage of group navigation. *Trends Ecol. Evol.* **19**, 453–455 (2004).

19. Ward, P. & Zahavi, A. The importance of certain assemblages of birds as "information-centres" for food-finding. *Ibis* **115**, 517–534 (2008).

20. Lorenz, J., Rauhut, H., Schweitzer, F. & Helbing, D. How social influence can undermine the wisdom of crowd effect. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 9020–9025 (2011).

21. Mahmoodi, A. *et al.* Equality bias impairs collective decision-making across cultures. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 3835–3840 (2015).

22. Mahmoodi, A., Bahrami, B. & Mehring, C. Reciprocity of social influence. *Nat. Commun.* **9**, 1–9 (2018).

23. Jayles, B. *et al.* The impact of incorrect social information on collective wisdom in human groups. *J. R. Soc. Interface* **17**, 20200496 (2020).

24. Becker, J., Brackbill, D. & Centola, D. Network dynamics of social influence in the wisdom of crowds. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E5070–E5076 (2017).

25. Lazer, D. & Friedman, A. The network structure of exploration and exploitation. *Adm. Sci. Q.* **52**, 667–694 (2007).

26. Mason, W. & Watts, D. J. Collaborative learning in networks. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 764–769 (2012).

27. Derex, M. & Boyd, R. Partial connectivity increases cultural accumulation within groups. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 2982–2987 (2016).

28. Barkoczi, D. & Galesic, M. Social learning strategies modify the effect of network structure on group performance. *Nat. Commun.* **7**, 13109 (2016).

29. Boyd, R. & Richerson, P. J. *Culture and the Evolutionary Process*. (University of Chicago Press, 1988).

30. Kameda, T. & Nakanishi, D. Cost–benefit analysis of social/cultural learning in a nonstationary uncertain environment. *Evol. Hum. Behav.* **23**, 373–393 (2002).

31. Kameda, T. & Nakanishi, D. Does social/cultural learning increase human adaptability?: Rogers's question revisited. *Evol. Hum. Behav.* **24**, 242–260 (2003).

32. Laland, K. N. Social learning strategies. *Learn. Behav.* **32**, 4–14 (2004).

33. Toyokawa, W., Whalen, A. & Laland, K. N. Social learning strategies regulate the wisdom and madness of interactive crowds. *Nat Hum Behav* **3**, 183–193 (2019).

34. Toyokawa, W., Saito, Y. & Kameda, T. Individual differences in learning behaviours in humans: Asocial exploration tendency does not predict reliance on social learning. *Evol. Hum. Behav.* **38**, 325–333 (2017).

35. McElreath, R. *et al.* Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 3515–3528 (2008).

36. Horita, Y., Takezawa, M., Inukai, K., Kita, T. & Masuda, N. Reinforcement learning accounts for moody conditional cooperation behavior: experimental results. *Sci. Rep.* **7**, 39275 (2017).

37. Sutton, R. S. & Barto, A. G. *Reinforcement Learning, second edition: An Introduction*. (MIT Press, 2018).

38. Toyokawa, W., Kim, H.-R. & Kameda, T. Human collective intelligence under dual exploration-exploitation dilemmas. *PLoS One* **9**, e95789 (2014).

39. Cohen, J. D., McClure, S. M. & Yu, A. J. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362**, 933–942 (2007).

40. Trimmer, P. C., McNamara, J. M., Houston, A. I. & Marshall, J. A. R. Does natural selection favour the Rescorla-Wagner rule? *J. Theor. Biol.* **302**, 39–52 (2012).

41. Henrich, J. & Boyd, R. The evolution of conformist transmission and the emergence of between-group differences. *Evol. Hum. Behav.* **19**, 215–241 (1998).

42. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).

43. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).

44. Rendell, L. *et al.* Why copy others? Insights from the social learning strategies tournament. *Science* **328**, 208–213 (2010).

45. Payzan-LeNestour, E. & Bossaerts, P. Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Comput. Biol.* **7**, e1001048 (2011).

46. Almaatouq, A. *et al.* Adaptive social networks promote the wisdom of crowds. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 11379–11386 (2020).

47. Moussaïd, M., Noriega Campero, A. & Almaatouq, A. Dynamical networks of influence in small group discussions. *PLoS One* **13**, e0190541 (2018).

48. Rand, D. G., Arbesman, S. & Christakis, N. A. Dynamic social networks promote cooperation in experiments with humans. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 19193–19198 (2011).

49. Harrell, A., Melamed, D. & Simpson, B. The strength of dynamic ties: The ability to alter some ties promotes cooperation in those that cannot be altered. *Science Advances* **4**, eaau9109 (2018).

50. Miyoshi, T. & Matsubara, S. Dynamically Forming a Group of Human Forecasters and Machine Forecaster for Forecasting Economic Indicators. *IJCAI* (2018).

51. Shirado, H. & Christakis, N. A. Locally noisy autonomous agents improve global human coordination in network experiments. *Nature* **545**, 370–374 (2017).

52. Holme, P. Modern temporal network theory: a colloquium. *Eur. Phys. J. B* **88**, 234 (2015).

53. Holme, P. & Saramäki, J. *Temporal Network Theory*. (Springer, Cham, 2019). doi:10.1007/978-3-030-23495-9.

54. Masuda, N. & Lambiotte, R. *A Guide to Temporal Networks*. (WORLD SCIENTIFIC (EUROPE), 2020). doi:10.1142/q0268.

55. Peirce, J. W. PsychoPy--psychophysics software in Python. *J. Neurosci. Methods* **162**, 8–13 (2007).

56. Van Slooten, J. C., Jahfari, S., Knapen, T. & Theeuwes, J. How pupil responses track value-based decision-making during and after reinforcement learning. *PLoS Comput. Biol.* **14**, e1006632 (2018).

57. Katahira, K. How hierarchical models improve point estimates of model parameters at the individual level. *J. Math. Psychol.* **73**, 37–58 (2016).

58. Piray, P., Dezfouli, A., Heskes, T., Frank, M. J. & Daw, N. D. Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies. *PLoS Comput. Biol.* **15**, e1007043 (2019).

Acknowledgements

We thank W. Toyokawa for helpful feedback about the computational model analysis. This study was supported by the Japan Society for the Promotion of Science (20J20900 to A.N. and JP16H0632 to T.K.), AFOSR European Office (FA9550-19-1-7024 to N.M.), and Japan Science and Technology Agency CREST (JPMJCR17A4 (17941861) to T.K.). We

appreciate support from the Center for Experimental Social Sciences at Hokkaido University and UTokyo Center for Integrative Science of Human Behavior.

Author contributions

A.N., N.M., and T.K. designed the study, discussed the methods and results, and wrote the manuscript. A.N. collected and analyzed the data. T.K. supervised the entire process of the study.

Additional Information

Competing interests

The authors declare no competing interests.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- NaitoNetworkExpSI01052022.pdf