



ARTICLE

DOI: 10.1057/s41599-018-0161-1

OPEN

# Diversity of inference strategies can enhance the 'wisdom-of-crowds' effect

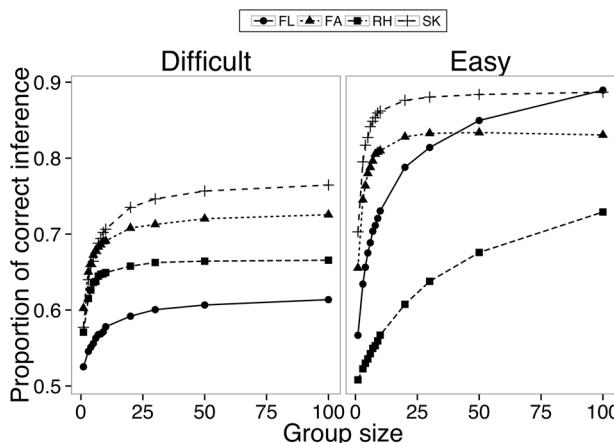
Itsuki Fujisaki<sup>1</sup>, Hidehito Honda<sup>2</sup> & Kazuhiro Ueda<sup>1</sup>

**ABSTRACT** Studies on inference have shown that people use a variety of inference strategies depending on the situation. Despite a great deal of discussion on the use of these strategies at an individual level, very little research has examined how the strategies people use affect group performance. To address this issue, we conducted two computer simulation studies on group decision-making. Our focus was primarily the diversity of strategies used in groups, as previous studies have suggested that diversity plays a critical role in the wisdom of crowds. Therefore, we systematically manipulated the diversity of inference strategies among group members and examined the effect on group performance. In Study 1, we conducted computer simulations using behavioural data from a previous study and found that diversity of strategies could improve group performance. That is, the group whose members used diverse strategies had higher accuracy than groups where all members used an identical strategy. We also investigated how such a phenomenon emerged. In Study 2, we created multiple hypothetical environmental settings and examined the effect. The environmental settings in Study 1 was limited to the 'kind' setting, in which correct inferences could be made for most problems by using a certain strategy, and the results of Study 2 showed that the findings of Study 1 could be generalized to other settings. For example, diversity could improve group performance in the 'wicked' environment where an inference strategy tends to lead an individual to the wrong answer. We also identified conditions in which the diversity enhanced group performance in each environment. Finally, for Study 1, we conducted additional simulations and discussed the conditions in which diversity would improve group performance more. The contributions to the research on the wisdom of crowds and human inference are discussed.

<sup>1</sup> Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan. <sup>2</sup> Department of Psychology, Yasuda Women's University, Hiroshima, Japan. Correspondence and requests for materials should be addressed to I.F. (email: [bpmx3ngj@gmail.com](mailto:bpmx3ngj@gmail.com)) or to K.U. (email: [ueda@gregorio.c.u-tokyo.ac.jp](mailto:ueda@gregorio.c.u-tokyo.ac.jp))

## Introduction

In the field of judgment and decision making, many studies have examined how individuals make inferences. There is growing evidence that people adaptively use different inference strategies depending on the situation (Kahneman and Frederick, 2005; Marewski and Schooler, 2011; Honda et al., 2017). For example, Honda et al. (2017) showed that the subjective difficulty of a problem affects strategy selection: when people feel that a problem is difficult, they tend to rely on simple heuristics, such as the recognition heuristic (Goldstein and Gigerenzer, 2002; Gigerenzer and Brighton, 2009; Brighton and Gigerenzer, 2011; Gigerenzer and Goldstein, 2011), and when they deem a problem easy, they tend to make an inference on the basis of available knowledge (Oppenheimer, 2003; Newell and Shanks, 2004; Hilbig and Richter, 2011). Despite a great deal of discussion on the use of these strategies at an individual level, only a few previous studies (Reimer and Katsikopoulos, 2004; Kämmer et al., 2014) have examined how such different inference strategies affect decision making at the collective level. Particularly from the perspective of wisdom-of-crowds research (Hertwig, 2012; Prelec et al., 2017; Jayles et al., 2017), the key question is 'Which inference strategy will lead to a successful wisdom-of-crowds effect?' It seems natural that a group can achieve high accuracy when its members use a strategy that demonstrates high accuracy at the individual level. However, Fujisaki et al. (2017) have shown that this is not always true. The results of their study are shown in Fig. 1. Using behavioural data (Honda et al., 2017), they first simulated individual binary choice problems using four inference strategy models examined in Honda et al. (2017) study. Second, collective decision making, wherein all group members used an identical inference strategy (e.g., the recognition heuristic), was simulated, and the accuracy levels of the strategies were compared. They found that the inference strategy that had the highest accuracy could change depending on whether it was used by an individual or the group. For example, in 'Easy' (right panel, Fig. 1), which is composed of 105 subjectively easy problems, subjective knowledge-based inference (SK) showed the highest accuracy among the strategies. However, when the group size was large (e.g., 100), the fluency heuristic (FL; Hertwig et al., 2008) showed the highest accuracy, although it had the second-lowest accuracy for the inference of an individual.

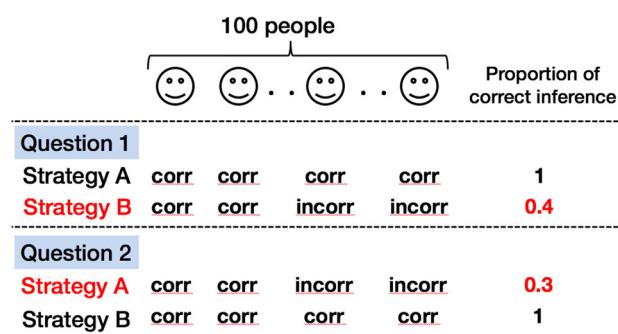


**Fig. 1** Results of Fujisaki et al. (2017) study. In this simulation study, all group members used an identical inference strategy. The line indicates the performance of each inference strategy (FL fluency heuristic, FA familiarity heuristic, RH recognition heuristic, SK subjective knowledge-based inference) in collective decision making and an individual inference. There were some cases where the strategy that showed the highest accuracy differed between the individual and group

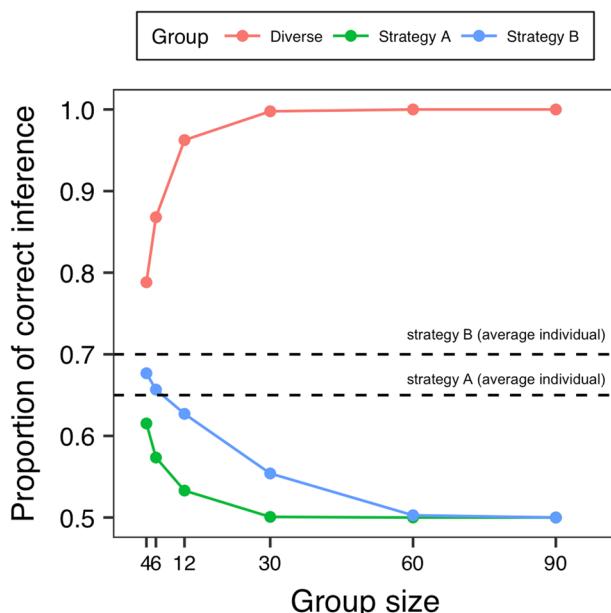
The underlying mechanisms are as follows. All the inference strategies used in the simulation study were adaptive (Gigerenzer et al., 1999) in that they led to accurate inferences as a whole. In other words, the proportion of correct inferences was mostly above the level of chance (i.e., 0.5). However, there were certain problems for which the proportion of correct inferences was below 0.5. Hereafter, we regard the strategy as having a 'bias' (Kahneman, 2011) for these problems. In collective decision making, which is based on the majority decision rule (Hastie and Kameda, 2005) for a binary choice problem, bias plays a critical role (Galesic et al., 2018). Namely, when the strategy had a bias, the accuracy of collective decision making in a group decreased to 0 as the group size increased. In contrast, it increased to 1 when the strategy did not have a bias (Condorcet Jury Theorem; Condorcet, 1785). As for FL, the mean accuracy of an individual's inference across problems was relatively low; however, the bias was the weakest. Consequently, for most of the problems, the accuracy increased to 1 as the group size increased. Thus, FL led to the most successful wisdom-of-crowd effect (see 'Detailed descriptions of Fujisaki et al., 2017' and Fig. S1 of SI for more details).

Fujisaki et al. (2017) examined only a situation in which all group members used an identical inference strategy. However, recent studies have shown that diversity can be a critical factor of successful group performance in terms of collective behaviour (Perc and Szolnoki, 2008; Santos et al., 2008; Yang et al., 2009; Yun et al., 2011; Santos et al., 2012; Analytis et al., 2017) and wisdom-of-crowds or collective intelligence in a wide range of tasks, such as decision-making (Galton, 1907; Lorenz et al., 2011; Krause et al., 2011; Luan et al., 2012; Conradt et al., 2013; Mavrodiev et al., 2013; Jönsson et al., 2015; Tump et al., 2018), problem-solving (Hong and Page, 2004; Liker and Bokony, 2009; Page, 2014), game theory (Mann and Helbing, 2017), and prediction of preferences (Müller-Trede et al., 2017). Thus, the diversity of inference strategies may play a critical role in the wisdom-of-crowds effect. However, to the best of our knowledge, no previous studies have examined this issue (see the Discussion section for a detailed review).

Figure 2 shows a hypothetical situation where the diversity of strategies leads to a successful wisdom-of-crowds effect. In this situation, a group consisting of 100 people made inferences for two binary choice problems. We set two inference strategies: Strategy A and Strategy B. Both strategies had a bias for one problem and always led to the correct answer for the other. Specifically, Strategy A (Strategy B) had a bias for Question 2



**Fig. 2** Illustration of hypothetical situation. Here, 'corr' means that a person infers correctly using a strategy and 'incorr' means that a person infers incorrectly using a strategy. The rightmost column indicates the proportion of correct inference across people. Red text indicates that the strategy has a bias for the question (i.e., the proportion of correct inference was below 0.5)



**Fig. 3** Results of the hypothetical situation. The dotted line indicates the proportion of correct inference by an individual using the respective inference strategy. In collective decision making, the Diverse group showed higher accuracy than both the Non-diverse groups. The simulation procedure was essentially the same as the following main simulations: group members were randomly selected from all people and they first made inferences individually using an allocated strategy depending on the group. Then, collective decision making was performed in accordance with the outcome of the inference by an individual under the simple majority rule

(Question 1) and always led to the correct answer in Question 1 (Question 2). Now consider the following three groups: all members use Strategy A or Strategy B (Non-diverse groups) or half the members rely on Strategy A while the others use Strategy B (Diverse group). Using computer simulations, we calculated the accuracy of collective decision making for each group by changing the group size (Fig. 3). We found that the proportion of correct inferences in the Non-diverse group converged to 0.5 as the group size increased because both strategies had the bias for one of the two questions. In contrast, in the Diverse group, the accuracy increased to 1.0 as the group size increased. For both questions, the expected values of accuracy by an individual were above 0.5 (Question 1:  $[1 + 0.4] / 2 = 0.7$ , Question 2:  $[0.3 + 1] / 2 = 0.65$ ). In other words, in this situation, diverse strategies were superior to non-diverse strategies because the bias in a single strategy was cancelled out by another strategy.

In two simulation studies, we examined the relationship between the diversity of inference strategies and group performance. We particularly focused on whether diversity could enhance the wisdom-of-crowds effect. In these studies, each group member first completed a series of binary choice problems individually by using the allocated inference strategy, and group decision making was then performed under the simple majority rule, known as the commonly used aggregation rule (Sorkin et al., 2001; Hastie and Kameda, 2005). Here, we systematically manipulated the diversity of inference strategies among group members to examine the effect on group performance. For the inference of an individual, in Study 1, we harnessed behavioural data from Honda et al. (2017) study, and in Study 2, we created multiple environmental settings on a computer to examine this from a broader perspective.

## Results

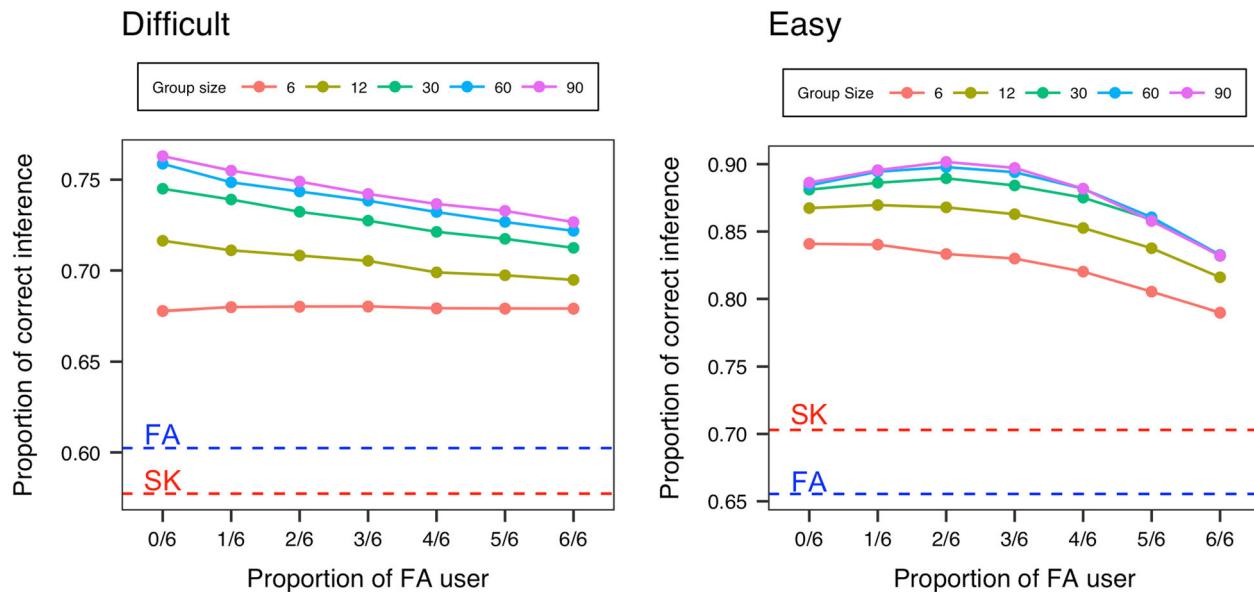
**Study 1.** In this study, we simulated the inference of an individual using the behavioural data from Honda et al. (2017) study. We assumed that the group members made inferences using one of two strategies: the familiarity heuristic (FA; Honda et al. 2011) or subjective knowledge-based inference (SK). We chose these two strategies because Honda et al. (2017) showed that FA and SK well-explained people's inference through competitive tests on inference models. In the set of simulations, group members were selected randomly from among all participants and they simulated a binary choice task using FA or SK. The binary choice task consisted of Easy and Difficult lists. Then, the collective decision making was performed (see the Materials and Methods section for more details). In order to manipulate the diversity of inference strategies, we set seven levels of group type: the proportion of group members who used FA (hereafter, FA users) to those who used SK (hereafter, SK users) was 0/6 to 6/6, 1/6 to 5/6, 2/6 to 4/6, 3/6 to 3/6, 4/6 to 2/6, 5/6 to 1/6, and 6/6 to 0/6. We also set five levels of group size: 6, 12, 30, 60, and 90. For example, when the group size was 90 and the proportion of FA users to SK users was 2/6 to 4/6, 30 members used FA, and the remaining 60 used SK. The proportion of correct inferences was separately calculated for each group type, group size, and list.

Figure 4 shows the relationship between the group types and the proportion of correct inferences for each group size. It is clear that the diversity of inference strategies enhanced the wisdom-of-crowds effect in the Easy list: when the group sizes were large (i.e., 30, 60 and 90), the accuracy reached its peak when the group members relied on diverse strategies (i.e., accuracy lines as a function of the proportion of FA users were inverse U-shaped). In contrast, in the Difficult list, the accuracy of collective decision making decreased as the proportion of FA users increased when the group sizes were relatively large (i.e., 30, 60 and 90).

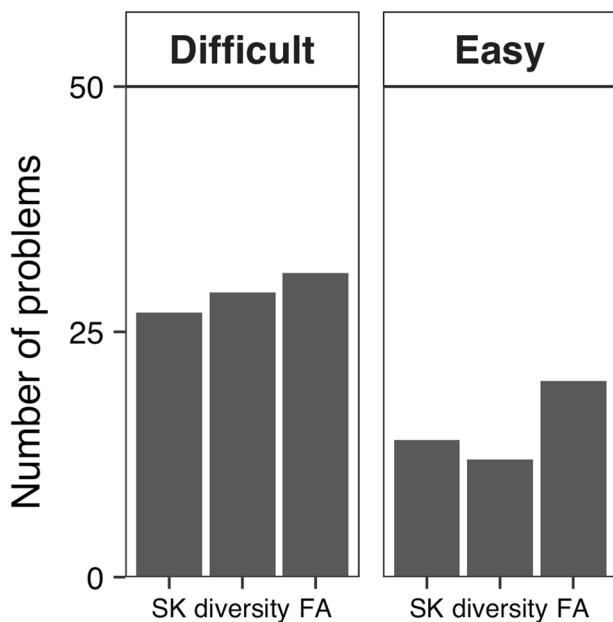
This different pattern of results is explained in Fig. 5, which shows the number of problems where the proportion of correct inferences by an individual was below 0.5 for the use of single and diverse strategies (diversity). In the Easy list, the number of such problems was smaller for the use of diverse strategies than for the use of both FA and SK since the bias in a single strategy was partly cancelled out by another strategy, as described earlier (see also 'Scatter plot of all problems in Study 1' and Fig. S2 of SI for more details). Thus, the group in which the members used diverse strategies could attain a higher accuracy than the groups whose members relied on FA or SK only. In contrast, in the Difficult list, the number of such problems was higher for the use of diverse strategies than for the use of SK because the bias when using SK was barely cancelled out by using FA. Therefore, with a large group size, the accuracy of collective decision making reached its peak when all group members relied on SK.

We conducted two additional simulations based on different assumptions concerning inference strategy: a group member made inferences (1) using SK or FL (the fluency heuristic; Hertwig et al. 2008) instead of FA and (2) using one of FL, FA, or SK (see 'Different assumptions on Study 1: Inference strategies', Figs. S3, 4 of SI).

**Study 2.** The environmental setting in Study 1 was 'kind' (Koriat, 2012; Hertwig, 2012) in that an inference strategy led to a correct answer as a whole; in other words, the mean accuracy of an inference strategy by an individual was above the level of chance. In other environmental settings, such as 'wicked', where an inference strategy tends to lead an individual to a wrong answer, how does the diversity of strategies affect group performance? To investigate this, we created various hypothetical environments



**Fig. 4** Results of Study 1. The blue and red dashed lines indicate the average accuracy of inference by an individual using the FA and SK. In the Easy list, when group size was large (e.g., 90), the proportion of correct inference reached its peak when group members used diverse strategies (e.g., proportion of FA users was 2/6). In contrast, in the Difficult list, such cases were not observed



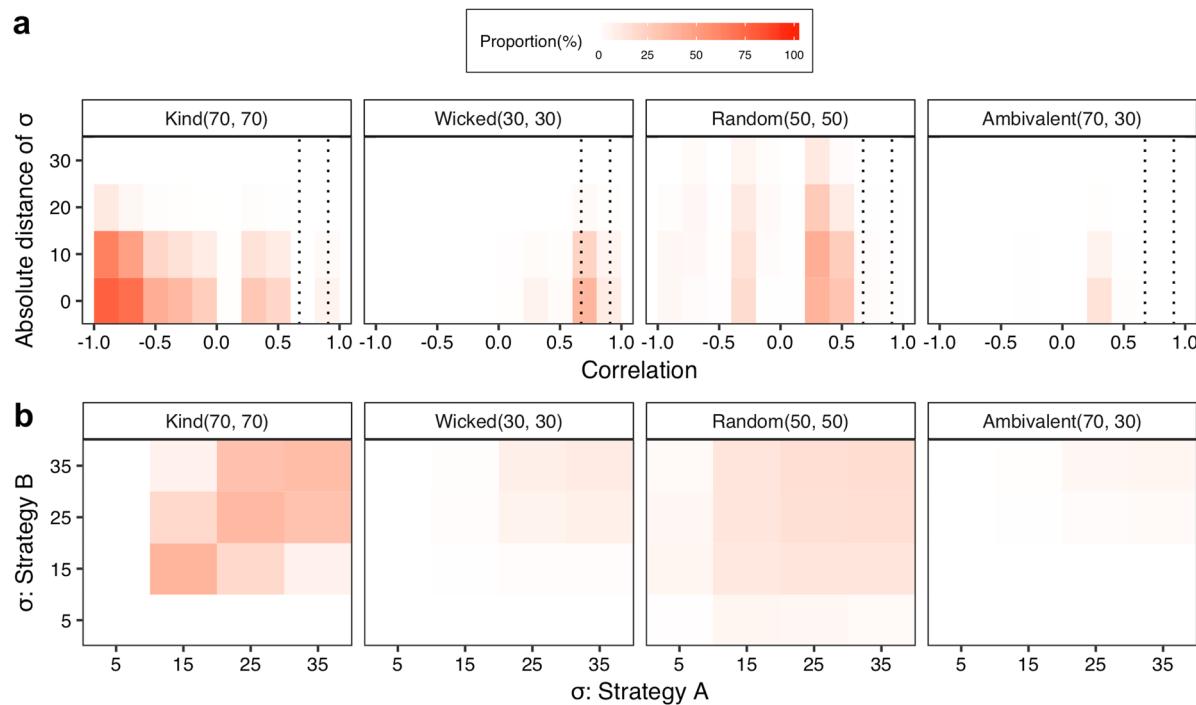
**Fig. 5** Number of problems where proportion of correct inference by an individual was below 0.5. For simplicity, in 'diversity,' we calculated the expected value of the accuracy for each problem, for which half of the participants used FA and the remaining half used SK. Then, we counted the number of problems where these expected values were below 0.5

and examined their effect on group performance using computer simulations. We particularly explored the cases where the diversity of inference strategies strongly enhanced the wisdom-of-crowds effect.

The settings of the computer simulations were as follows. We created four environments: kind, wicked, random, and ambivalent. All environments consisted of 100 individuals, 100 problems, and two inference strategies (hereafter, strategy A and strategy B). The accuracy of the two strategies differed among the environments. The mean percentages of correct answers of an individual across the problems were: kind = [60, 60], wicked = [40, 40],

random = [50, 50], and ambivalent = [60, 40]. We assumed the accuracy of an inference strategy followed a normal distribution across the problems. In this set of simulations, we manipulated the variance ( $\sigma$ ) as far as it followed a normal distribution. In the following analysis, we exclude a few trials in which the accuracy did not follow a normal distribution (see 'Excluded trials in Study 2 and additional simulations in Study 1' and Tabs. S1, 2 of SI for more details). We set four levels of  $\sigma$  for each strategy: 5, 15, 25 and 35. On the basis of this distribution, the percentages of correct answers of 100 problems were generated, and these percentages were then used to generate the outcomes of inference for an individual. To vary the correlation between the accuracy of the two strategies, we set 24 patterns for the order of percentages of correct answers (see the Materials and Methods section and Tab. 1 for more details). Then, the collective decision making was performed. For simplicity, in this study, we set the group size to 100 (i.e., all individuals joined the collective decision making) and set three group types: all members used strategy A or strategy B (Non-diverse groups) or half of the members relied on strategy A while the others used strategy B (Diverse group).

Figure 6a summarizes the results of the simulations. In this analysis, we categorized the absolute distance of  $\sigma$  and the correlation values. We counted the number of trials in each category and then calculated the proportion of cases in which the Diverse group showed higher accuracy (especially 5%) than both the Non-diverse groups. We found that, in addition to the kind environment, there were certain cases in the random, wicked, and ambivalent environments in which diversity improved the group performance. In the ambivalent environment, there were only a few such cases, indicating that in order for the diversity to enhance the wisdom-of-crowds effect, the accuracy of two strategies should be close. Regarding the correlation, there was a wide range of values in the kind and random environments. In contrast, in the wicked environment, the diversity improved the group performance only when the correlation was positive. We analysed the behavioural data in Honda et al. (2017) study and found a strong positive correlation between FA and SK for both Easy and Difficult lists (dotted lines in Fig. 6a). Therefore, we speculate that even for the wicked environment, the diversity of strategies could enhance the wisdom-of-crowds effect. As for the



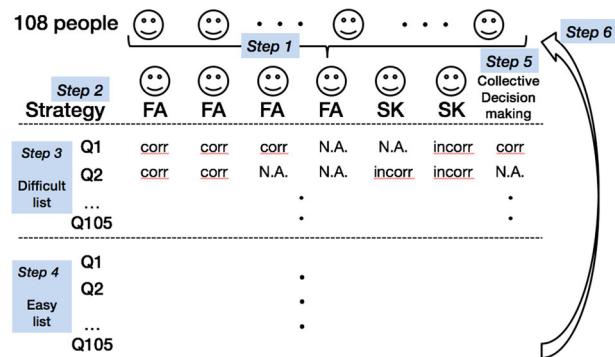
**Fig. 6** Results of Study 2. The coloured part indicates the proportion of cases where the Diverse group showed a higher accuracy than both Non-diverse groups (i.e., >5%) in each category. **a** We categorized the absolute distance of  $\sigma$ s into four levels from 0 to 30 (in increments of 10%) and the correlation values into 10 levels from -1 to 1 (in increments of 0.2). Then we counted the number of trials in each category defined by the absolute distance of  $\sigma$  and the correlation. The black dotted lines indicate the correlation values in the behavioural data (Honda et al., 2017; Left line: Easy list, Right line: Difficult list). In addition to kind, the wicked, random, and ambivalent environments had cases where diversity improved the group performance. **b** We categorized the magnitude of  $\sigma$ s into four levels from 5 to 35 (in increments of 10%). This analysis provides more detailed conditions wherein the Diverse group could show higher accuracy

absolute distance of  $\sigma$ , there was a wide range of values in the kind environment, while in the other environmental settings, the diversity improved the accuracy only when the absolute distance of  $\sigma$  was relatively small.

Figure 6b represents  $\sigma$ s in the two strategies. In this analysis, we categorized the magnitude of  $\sigma$ s using the same analysis procedure as in Fig. 6a. This figure shows more details of the conditions in which the Diverse group showed higher accuracy. For example, in the wicked environment, both  $\sigma$ s should be relatively large (e.g.,  $\sigma$ s = 25). When the  $\sigma$  was large, the accuracy of the strategy for some problems was above the level of chance, and therefore the strategy could cancel out the bias in another strategy. In the kind environment, the closer the values of the two  $\sigma$ s, the higher the proportion of cases where the diversity improved group performance. In the random environment,  $\sigma$  could be a wide range of values (e.g.,  $\sigma$  = 5) since the bias in a single strategy could be cancelled out by another strategy regardless of the magnitude of  $\sigma$  at the theoretical level.

## Materials and methods

**Materials in Study 1.** One hundred and eight participants took part in the experiments in Honda et al. (2017) study. They performed a binary choice inference task on population size, such as 'Which city has the larger population, Osaka or Kyoto?' Pairs of cities for the task consisted of two types of lists: Difficult and Easy. Each list comprised 15 Japanese cities, and 105 pairs ( $=15 \times 14/2$ ) were made for the task. The difference in the population size was generally much higher for the pairs in the Easy list than for those in the Difficult list, suggesting that the former was easier for the participants to make inferences about than the latter. This was confirmed by the behavioural experiment in Honda et al. (2017) study.



**Fig. 7** Simulation procedure in Study 1. Step 1: Group members are randomly selected from among 108 participants. Step 2: An inference strategy is allocated to each member. Steps 3 and 4: Each group member makes inferences individually using the allocated strategy for all 105 problems on the Difficult and Easy lists. Step 5: Using the simulated inferences of individuals, collective decision making is conducted under the simple majority rule. Step 6: Return to Step 1. Steps 1–5 are repeated 1,000 times. This figure shows an example whose group size is 6 with a proportion of FA users to SK users of 4/6 to 2/6

Using this behavioural data, we can simulate the inference of an individual when they use FA and SK.

**Simulation procedure for Study 1.** For all combinations of group type and group size, we performed the following simulation (Fig. 7).

Step 1: Group members are randomly selected from among the 108 participants.

**Table 1 Patterns of the order of percentage of correct answers in strategy A**

Category	1	2	3	4
Pattern 1	1, 2, ...25	26, 27, ...50	51, 52, ...75	76, 77, ...100
Pattern 2	1, 2, ...25	26, 27, ...50	76, 77, ...100	51, 52, ...75
Pattern 23	76, 77, ...100	51, 52, ...75	1, 2, ...25	26, 27, ...50
Pattern 24	76, 77, ...100	51, 52, ...75	26, 27, ...50	1, 2, ...25

Numbers are allocated to percentages of correct answers in decreasing order of values. For simplicity, only Patterns 1, 2, 23 and 24 are shown

Step 2: An inference strategy is allocated to each member.

Steps 3 and 4: Each group member makes inferences individually using the allocated strategy for all 105 problems in the Difficult and Easy lists.

Step 5: Using the simulated inferences of individuals, collective decision making is performed under the simple majority rule.

Step 6: Return to Step 1. Steps 1–5 are repeated 1000 times.

**Assumptions in Study 1.** In Steps 3 and 4, the inferences made by an individual were classified into three categories: ‘correct’, ‘incorrect’, and ‘cannot use strategy’ (in Fig. 7, denoted as ‘corr’, ‘incorr’, and ‘N.A.’, respectively). The category ‘cannot use strategy’ was used when a simulated group member could not make an inference based on the strategy since the inference model did not make any specific prediction. As with Fujisaki et al.’s study (Fujisaki et al., 2017), we assumed that the group members did not join the collective decision making because a previous study (Reimer and Katsikopoulos, 2004) reported that in group decision making, people who could not use an inference strategy (e.g., the recognition heuristic) were less influential than those who could. Note that we conducted additional simulations based on a different assumption: a group member who could not make an inference based on the strategy made a random choice and joined the collective decision making. We found that even with this assumption, the diversity of strategies could enhance the wisdom-of-crowds effect (see ‘Different assumption on Study 1: A group member cannot use the allocated inference strategy’ and Fig. S5 of SI).

In Step 5, there was one case where a group could not make a decision: half of the group members who joined the collective decision making inferred correctly while the other half did so incorrectly (see Q2 in Fig. 7, represented by ‘N.A.’). In this case, we assume that the group made a random choice. That is, the group made an accurate decision with 50% probability.

**Simulation procedure for Study 2.** In each environment, we conducted the following simulation procedure for all 16 combinations of  $\sigma$ s in two strategies ( $=4 \times 4$ ).

Step 1: Percentages of correct answers of 100 problems are generated on the basis of distributions of strategies A and B.

Step 2: The order of percentages of correct answers of 100 problems in two strategies is manipulated.

Step 3: On the basis of percentages of correct answers, the outcomes of inference for 100 group members using a strategy is generated.

Step 4: Each group member is allocated an inference strategy and makes inferences individually for 100 problems.

Step 5: Collective decision making is performed under the simple majority rule.

Step 6: Return to Step 4. Steps 4–5 are repeated 100 times.

Step 7: Return to Step 2. Steps 2–6 are repeated until all patterns of manipulation are examined.

Step 8: Return to Step 1. Steps 1–7 are repeated 1000 times.

**Table 2 Behavioural data in Honda et al. (2017)**

	Difficult list		Easy list	
	FA	SK	FA	SK
<i>M</i>	60.24	57.73	65.55	70.29
$\sigma$	17.43	10.47	13.98	14.14
Correlation	0.91		0.67	

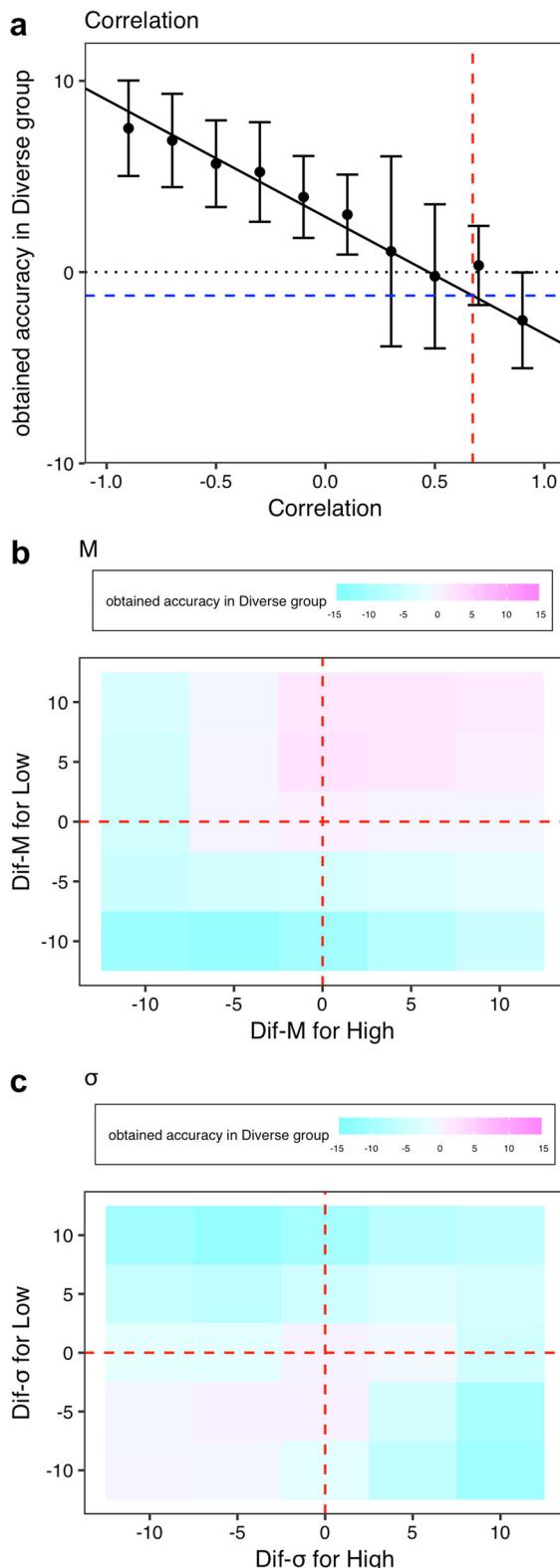
In Step 2, we manipulated the order of the percentages of correct answers in two strategies: (1) we sorted them in ascending order, and (2) for strategy A, we additionally manipulated the order according to the order pattern (Table 1) and recorded the correlations (Spearman’s rank correlation coefficient). In Steps 4–5, the accuracy of all group types (Diverse group and Non-diverse groups) was recorded.

**Assumptions in Study 2.** In order to conduct the simulations, in Step 1, if the percentage of correct answers was over 100 or under 0, we converted it to 100 or 0, respectively (see ‘Information excluded by the assumptions in Study 2 and additional simulations in Study 1’ and Figs. S6, 7 of SI for more details). For simplicity, in Step 4, we assumed that the outcome of an inference by an individual was classified as correct or incorrect. In Step 5, if half of the group members inferred correctly while the other half did so incorrectly, we assumed that the group made a random choice, as in Study 1.

This study included neither ethical approval nor informed consent since only computer simulations were conducted and no behavioural experiments were carried out.

## Discussion

In the present study, we examined how the diversity of inference strategies used by group members affects the group performance. Through two simulation studies, we found that the diversity of strategies could enhance the wisdom-of-crowds effect when the bias in a single strategy was partly cancelled out by another strategy. Previous studies on inference have mainly focused on individuals and discussed topics such as bias (Kahneman, 2011) as well as the adaptive nature (Gigerenzer et al., 1999) of heuristics. In contrast, the present study and that of Fujisaki et al. (2017) clarify that investigation at the group level can also offer insights into the nature of human inference. Moreover, this study makes a novel contribution to the literature on the wisdom-of-crowds in that it demonstrates the efficacy of the diversity of strategies on the wisdom-of-crowds. As mentioned in the Introduction, many previous studies have also pointed out the important role of diversity in the wisdom-of-crowds, but the definitions of diversity were apparently different from in our study, focusing as they did on the diversity of estimated values among group members for numerical estimation tasks (Galton,



**Fig. 8** Results of additional simulation in Study 1. Each panel shows the relationship between the obtained accuracy in the Diverse group and **a** the correlation, **b** Dif- $M$  (the differences in  $M$  from the default value in the Easy list), and **(c)** Dif- $\sigma$  (the differences in  $\sigma$  from the default value in the Easy list). The red dashed lines indicate the value in the Easy list. For factors that were not the focus of analysis, values were fixed to those in the Easy list. In **b** and **c**, 'High' (Low) indicates the strategy whose accuracy was high (low) at the individual level, and for the correlation, we considered trials in which the correlation value was within 0.1 from the correlation value in the Easy list ( $=0.67$ ). **a** For simplicity, the correlation values were categorized into 10 levels from -1 to 1 into 10 (in increments of 0.2). Blue dashed line indicates the values of the obtained accuracy in the Diverse group when the correlation value corresponded to the default value in the Easy list (indicated by the red line). The lower the correlation, the higher the obtained accuracy in the Diverse group. **b** The obtained accuracy of the Diverse group reached its peak when Dif- $M$  for Low was 5 and that for High was 0. **c** The obtained accuracy in the Diverse group became higher—albeit slightly—when Dif- $\sigma$  for Low became -5 and Dif- $\sigma$  for High was -5 or 0

strategies could play a critical role in the wisdom-of-crowds. Moreover, our study is one of the few studies on the wisdom-of-crowds that has examined the cognitive process, in contrast to many previous studies that focused on its output (e.g., estimated values).

**Implications for real-world collective decision making.** As discussed in the Introduction, Honda et al. (2017) showed that people use different inference strategies depending on the situation: they tend to use simple heuristics (especially FA) for difficult problems and knowledge-based inference (SK) for easy problems. However, Honda et al. (2017) also reported that people have a certain degree of diversity in inference strategies. In the Difficult list, approximately 25% of the participants used SK, while in the Easy list, approximately 25% used a simple heuristic (especially the familiarity heuristic, FA). What implications does this moderate diversity have in real-world collective decision making? We now compare the results reported in Honda et al. (2017) study with the results of Study 1. In the Difficult list, it seems these minority SK users may play a role in improving the accuracy of collective decision making, as the accuracy increases when the proportion of SK users increases. Similarly, in the Easy list, the minority FA users may play a role in enhancing the wisdom-of-crowds effect because the accuracy reached its peak for large groups when the proportion of FA users was moderate (i.e., 2/6), so we speculate that our moderate diversity of inference strategies may work adaptively in real-world collective decision making.

**Additional simulations in Study 1.** In Study 1, the diversity of inference strategies improved the group accuracy by up to 2% in the Easy list: when the group size was 90, the accuracy for 2/6 FA users reached its peak, which was about 0.90 (heterogeneous strategy), while when there were no FA users (homogeneous strategy), the accuracy was about 0.88. Then, in which conditions would the diversity further improve the group performance? In this additional study, we analysed the effect on inference strategy of three factors: (a) correlation between the accuracy of the strategies, (b) the mean percentage of correct answer rates across the problems for an individual ( $M$ ), and (c)  $\sigma$ . For this study, we modified the simulation settings in Study 2 concerning  $M$  and  $\sigma$ . First, we set two strategies whose  $M$  and  $\sigma$  corresponded to those in the Easy list (see Table 2). We defined the strategy whose accuracy was high (low) at the individual level as 'High' (Low). Second, for the set of simulations, we manipulated the factor that was the focus of analysis so that its value differed from the default

1907; Lorenz et al., 2011; Krause et al., 2011; Mavrodiev et al., 2013; Jönsson et al., 2015), the diversity of range of information searched by group members (Luan et al., 2012), the diversity of goals across individuals (Conradt et al., 2013), and the diversity in cue beliefs (Tump et al., 2018). To the best of our knowledge, the present study is the first to show that the diversity of inference

one in the Easy list: for (b), the differences in  $M$  from the default value (Dif- $M$ ) were set to  $-10$ ,  $-5$ ,  $\pm 0$ ,  $+5$ , and  $+10$  and for (c), the differences in  $\sigma$  from the default value (Dif- $\sigma$ ) were set to  $-10$ ,  $-5$ ,  $\pm 0$ ,  $+5$ , and  $+10$ . For factors that were not the focus of analysis, values were fixed to those in the Easy list. For example, in (b),  $\sigma$ s from behavioural data  $= \pm 0$ . That is, we conducted the simulation procedure in Study 2 for all 25 combinations ( $= 5 \times 5$ ) of  $M$ s in (b) and  $\sigma$ s in (c). Note that in (a), both Dif- $M$ s and Dif- $\sigma$ s  $= \pm 0$ . As an index of how much the diversity improved the group performance, we considered 'obtained accuracy of the Diverse group', which was calculated by subtracting the accuracy of the Non-diverse group (in particular, the more accurate group of the two) from that of the Diverse group.

Figure 8 shows the results of these simulations. Red dashed lines indicate the value in the Easy list. On the basis of this, the condition in which the diversity should improve the accuracy more than the results of Study 1 can be predicted. In (a), we can see that for further improvement, the correlation value should be lower. In (b), the obtained accuracy in the Diverse group tended to be higher when Dif- $M$  for Low became higher. In particular, the obtained accuracy of the Diverse group reached its peak when Dif- $M$  for Low (the strategy with low accuracy) was 5 and that for High (the strategy with high accuracy) was 0. In (c), although the effect was slight, the obtained accuracy of the Diverse group became higher when Dif- $\sigma$  for Low became  $-5$  and Dif- $\sigma$  for High was  $-5$  or 0.

**Diversity of inference strategies in other environments.** In Study 2, we examined multiple environments (including those from Study 1) and found that the diversity of inference strategies could also enhance the wisdom-of-crowds effect in other environments: for example, in the wicked environment, where individuals tend to give the wrong answer (Laan et al., 2017). Moreover, we clarified the condition in which diversity improves the group performance: in the wicked environment, the correlation of the accuracy between the two strategies should be strongly positive and both strategies need to have a high  $\sigma$ . Thus, the findings of Study 2 help predict cases in which the diversity of strategies can improve the accuracy beyond the environment in Study 1. Further experimental studies are needed to clarify this.

**Related studies.** We can connect the findings of the present study to those in other fields such as evolutionary game theory, although in this work we did not consider any interaction between agents. First, experimental (Isaac et al., 1994) and theoretical (Szolnoki and Perc, 2011) studies have shown that the evolution of cooperation is promoted in large groups, especially those above 40 (Group-size effects). In this respect, the present study also reported that in large groups (e.g., above 30 in Study 1 on the Easy list), diversity plays a positive role in the group. Thus, the present study sheds light on the intriguing nature of large groups. Second, some previous studies (Zhao et al., 2011; Szolnoki et al., 2012; Szolnoki and Perc, 2015) have examined the relationship between group performance and the diversity of behavioural strategies. These studies have defined the degree of diversity as the proportion of agents who do not imitate others' behaviours, such as cooperation, and examined the effect of the degree of diversity on group performance. In this respect, the present study addressed this issue from the perspective of inference strategies.

**Future studies.** We suggest two directions for future studies. One path is to examine the effect of strategy selection. In the present study, an individual was allocated a single strategy across all problems. However, previous studies have shown that there is

variability in selecting inference strategies (Rieskamp and Otto, 2006) and people adapt their strategies to the structure of the environment (Payne et al., 1986; Pachur et al., 2011). How the diversity of strategies is moderated should be further studied by considering strategy selection and how the group performance consequently changes. Another path is to examine whether one can experimentally control the diversity of inference strategies and leverage it to improve the group performance. We presume that the two strategies examined in this study, familiarity heuristic and knowledge-based inference, essentially correspond to Systems 1 and 2 (Kahneman, 2011), respectively. Therefore, one could manipulate the inference strategies people use by providing them with an instruction to think quickly or slowly. Consequently, a group provided with diverse instructions (e.g., to think quickly for half of the members and to think slowly for the rest) may perform better than groups provided with homogeneous instructions (i.e., to think either quickly or slowly for all members). Experimental studies that examine the relationship between thinking styles and group performance at the collective level are needed.

Received: 12 March 2018 Accepted: 6 August 2018

Published online: 04 September 2018

## References

Analytis P, Stojic H, Gelastopoulos A (2017) Diversity of preferences can increase collective welfare in sequential exploration problems. <https://arxiv.org/abs/1703.10970>.

Brighton H, Gigerenzer G (2011) Towards competitive instead of biased testing of heuristics: a reply to Hilbig and Richter (2011). *Top Cogn Sci* 3(1):197–205

Condorcet M (1785) *Essai su l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. (trans. McLean L, Hewitt F) Aldershot. Edward Elgar, Hants, 1994

Conradt L, List C, Roper TJ (2013) Swarm intelligence: when uncertainty meets conflict. *Am Nat* 182(5):592–610

Fujisaki I, Honda H, Ueda K (2017) Wisdom of crowd from diversity: theoretical analyses from collective decision making. *Cogn Stud Bull Jpn Cogn Sci Soc* 24 (3):284–299

Galesic M, Barkoczi D, Katsikopoulos K (2018) Smaller crowds outperform larger crowds and individuals in realistic task conditions. *Decision* 5(1):1–15

Galton F (1907) Vox populi (The wisdom of crowds). *Nature* 75(7):450–451

Gigerenzer G, Brighton H (2009) Homo heuristics: Why biased minds make better inferences. *Top Cogn Sci* 1(1):107–143

Gigerenzer G, Goldstein D (2011) The recognition heuristic: a decade of research. *Judgm Decis Mak* 6(1):100–121

Gigerenzer G, Todd P, the ABC Research Group (1999) Simple heuristics that make us smart. Oxford University Press: New York

Goldstein DG, Gigerenzer G (2002) Models of ecological rationality: the recognition heuristic. *Psychol Rev* 109(1):75–90

Hastie R, Kameda T (2005) The robust beauty of majority rules in group decisions. *Psychol Rev* 112(2):494–508

Hertwig R, Herzog SM, Schooler LJ, Reimer T (2008) Fluency heuristic: a model of how the mind exploits a by-product of information retrieval. *J Exp Psychol Learn Mem Cogn* 34(5):1191–1206

Hertwig R (2012) Tapping into the wisdom of the crowd—with confidence. *Science* 336(6079):303–304

Hilbig B, Richter T (2011) Homo heuristics outnumbered: comment on Gigerenzer and Brighton (2009). *Top Cogn Sci* 3(1):187–196

Honda H, Abe K, Matsuka T, Yamagishi K (2011) The role of familiarity in binary choice inferences. *Mem Cogn* 39(5):851–863

Honda H, Matsuka T, Ueda K (2017) Memory-based simple heuristics as attribute substitution: competitive tests of binary choice inference models. *Cogn Sci* 41 (S5):1093–1118

Hong L, Page SE (2004) Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proc Natl Acad Sci* 101(46):16385–16389

Isaac RM, Walker JM, Williams AW (1994) Group size and the voluntary of public goods. *J Public Econ* 54(1):1–36

Jayles B, Kim H, Escobedo R, Cezera S, Blanchet A, Kameda T et al. (2017) How social information can improve estimation accuracy in human groups. *Proc Natl Acad Sci* 114:12620–12625

Jönsson ML, Hahn U, Olsson EJ (2015) The kind of group you want to belong to: Effects of group structure on group accuracy. *Cognition* 142:191–204

Kahneman D (2011) Thinking, fast and slow. Macmillan: New York

Kahneman D, Frederick S (2005) A model of heuristic judgment. In: Holyoak J, Morrison RG (eds) *The Cambridge handbook of thinking and reasoning*. Cambridge University Press: New York, pp 267–293

Kämmer JE, Gaissmaier W, Reimer T, Schermuly CC (2014) The adaptive use of recognition in group decision making. *Cogn Sci* 38(5):911–942

Koriat A (2012) When are two heads better than one and why? *Science* 336 (6079):360–362

Krause S, James R, Faria JJ, Ruxton GD, Krause J (2011) Swarm intelligence in humans: diversity can trump ability. *Anim Behav* 81(5):941–948

Laan A, Madirolas G, de Polavieja GG (2017) Rescuing collective wisdom when the average group opinion is wrong. *Front Robot AI* 4:56. <https://doi.org/10.3389/frobt.2017.00056>

Liker A, Bokony V (2009) Larger groups are more successful in innovative problem solving in house sparrows. *Proc Natl Acad Sci* 106(19):7893–7898

Lorenz J, Rauhut H, Schweitzer F, Helbing D (2011) How social influence can undermine the wisdom of crowd effect. *Proc Natl Acad Sci* 108 (22):9020–9025

Luan S, Katsikopoulos KV, Reimer T (2012) When does diversity trump ability (and vice versa) in group decision making? a simulation study. *PLoS One* 7 (2):e31043. <https://doi.org/10.1371/journal.pone.0031043>

Mann R, Helbing D (2017) Optimal incentives for collective intelligence. *Proc Natl Acad Sci* 114(20):5077–5082

Marewski JN, Schooler LJ (2011) Cognitive niches: an ecological model of strategy selection. *Psychol Rev* 118(3):393–437

Mavrodiev P, Tessone CJ, Schweitzer F (2013) Quantifying the effects of social influence. *Sci Rep* 3(1):1360

Müller-Trede J, Choshen-Hillel S, Barneron M, Yaniv I (2017) The wisdom of crowds in matters of taste. *Manag Sci* 64(4):1779–1803

Newell B, Shanks D (2004) On the role of recognition in decision making. *J Exp Psychol Learn Mem Cogn* 30(4):923–945

Oppenheimer D (2003) Not so fast!(and not so frugal!): rethinking the recognition heuristic. *Cognition* 90(1):B1–B9

Pachur T, Todd PM, Gigerenzer G, Schooler LJ, Goldstein DG (2011) The recognition heuristic: a review of theory and tests. *Front Psychol* 2:147. <https://doi.org/10.3389/fpsyg.2011.00147>

Page SE (2014) Where diversity comes from and why it matters? *Eur J Soc Psychol* 44(4):267–279

Payne JW, Bettman JR, Johnson EJ (1986) Adaptive strategy selection in decision making. *J Exp Psychol Learn Mem Cogn* 14(3):534–552

Perc M, Szolnoki A (2008) Social diversity and promotion of cooperation in the spatial prisoner's dilemma game. *Phys Rev E* 77(1):011904

Prelec D, Seung HS, McCoy J (2017) A solution to the single-question crowd wisdom problem. *Nature* 541(7638):532–535

Reimer T, Katsikopoulos KV (2004) The use of recognition in group decision-making. *Cogn Sci* 28(6):1009–1029

Rieskamp J, Otto PE (2006) SSL: a theory of how people learn to select strategies. *J Exp Psychol Learn Mem Gen* 135(2):207–236

Santos FC, Pinheiro FL, Lenaerts T, Pacheco JM (2012) The role of diversity in the evolution of cooperation. *J Theor Biol* 299:88–96

Santos FC, Santos MD, Pacheco JM (2008) Social diversity promotes the emergence of cooperation in public goods games. *Nature* 454(7201):213–216

Sorkin R, Hays C, West R (2001) Signal-detection analysis of group decision making. *Psychol Rev* 108(1):183–203

Szolnoki A, Perc M (2011) Group-size effects on the evolution of cooperation in the spatial public goods game. *Phys Rev E* 84:047102

Szolnoki A, Perc M (2015) Conformity enhances network reciprocity in evolutionary social dilemmas. *J R Soc Interface* 12(103):20141299

Szolnoki A, Wang Z, Perc M (2012) Wisdom of groups promotes cooperation in evolutionary social dilemmas. *Sci Rep* 2:576. <https://doi.org/10.1038/srep00576>

Tump AN, Wolf M, Krause J, Kurvers RHJM (2018) Individuals fail to reap the collective benefits of diversity because of over-reliance on personal information. *J R Soc Interface* 15(142):20180155

Yang HX, Wu ZX, Zhou C, Zhou T, Wang BH (2009) Effects of social diversity on the emergence of global consensus in opinion dynamics. *Phys Rev E* 80:046108

Yun CK, Masuda N, Kahng B (2011) Diversity and critical behavior in prisoner's dilemma game. *Phys Rev E* 83(5):057102

Zhao L, Yang G, Wang W, Chen Y, Huang JP, Ohashi H et al. (2011) Herd behavior in a complex adaptive system. *Proc Natl Acad Sci* 108 (37):15058–15063

## Data availability

The datasets generated during and/or analysed during the current study are not publicly available due to using data from Honda et al. (2017) study; but they are available from the corresponding author upon reasonable request.

## Acknowledgements

This study was supported by JSPS KAKENHI Grant Number JP16K16070 and JP18H03501, for the second author, as well as JSPS KAKENHI Grant Number JP16H01725 and MEXT KAKENHI Grant Number JP26118002, for the third author.

## Additional information

**Supplementary information:** The online version of this article (<https://doi.org/10.1057/s41599-018-0161-1>) contains supplementary material, which is available to authorized users.

**Competing interests:** The authors declare no competing interests.

**Reprints and permission** information is available online at <http://www.nature.com/reprints>

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018