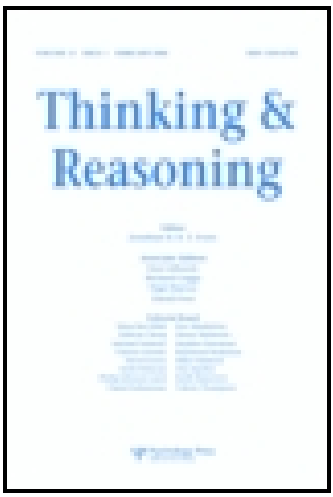

Thinking & Reasoning
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/ptar20

The benefits of argumentation are cross-culturally robust: The case of Japan

H. Mercier[a], M. Deguchi[b], J.-B. Van der Henst[c] & H. Yama[d]

[a] Cognitive Science Center, University of Neuchâtel, Neuchâtel, Switzerland

[b] Faculty of Foreign Studies, Sophia University, Tokyo, Japan

[c] CNRS, Laboratoire Langage, Cerveau et Cognition (L2C2), Université Lyon, Bron, France

[d] Graduate School of Literature and Human Sciences, Osaka City University, Osaka, Japan

Published online: 20 Jan 2015.

To cite this article: H. Mercier, M. Deguchi, J.-B. Van der Henst & H. Yama (2015): The benefits of argumentation are cross-culturally robust: The case of Japan, Thinking & Reasoning, DOI: 10.1080/13546783.2014.1002534

To link to this article: http://dx.doi.org/10.1080/13546783.2014.1002534

Thinking & Reasoning, 2015
http://dx.doi.org/10.1080/13546783.2014.1002534

Routledge
Taylor & Francis Group

The benefits of argumentation are cross-culturally robust: The case of Japan

H. Mercier[1], M. Deguchi[2], J.-B. Van der Henst[3], and H. Yama[4]

[1]Cognitive Science Center, University of Neuchâtel, Neuchâtel, Switzerland
[2]Faculty of Foreign Studies, Sophia University, Tokyo, Japan
[3]CNRS, Laboratoire Langage, Cerveau et Cognition (L2C2), Université Lyon, Bron, France
[4]Graduate School of Literature and Human Sciences, Osaka City University, Osaka, Japan

Thanks to the exchange of arguments, groups outperform individuals on some tasks, such as solving logical problems. However, these results stem from experiments conducted among Westerners and they could be due to cultural particularities such as tolerance of contradiction and approval of public debate. Other cultures, collectivistic cultures in particular, are said to frown on argumentation. Moreover, some influential intellectual movements, such as Confucianism, disapprove of argumentation. In two experiments, the hypothesis that Easterners might not share the benefits of argumentation was tested. In Experiment 1, Japanese participants had to solve a standard logical problem individually and then in groups. They performed significantly better in groups. In Experiment 2, Japanese participants had to estimate the weight of various animals. They did so individually, then after learning of another participant's estimates, then after discussing these estimates with the other participant, and then individually again. While the Japanese participants also benefitted from the discussion, these benefits were only visible when participants provided a final individual estimate. This delay is interpreted as reflecting the pressure to preserve social harmony that would have constrained Japanese participants to yield to their partner even when knowing that this did not improve the accuracy of their answer.

Keywords: Argumentation; Group decision making; Cross-cultural; Japan

Correspondence should be addressed to H. Mercier, Cognitive Science Center, University of Neuchâtel, Espace Louis-Agassiz 1, 2000 Neuchâtel, Switzerland. Email: hugo.mercier@gmail.com

We thank Mike Oaksford and David Moshman for their helpful comments.

No potential conflict of interest was reported by the authors.

We benefited from an Ambizione grant from the Swiss National Science Foundation (to H. Mercier) [grant number PZ00P1_142388/1].

While groups do not make better decisions than individuals in all areas (Kerr & Tindale, 2004), when discussing problems for which there is a demonstrably correct answer, or one that can be supported by strong arguments, groups consistently outperform individuals. This is true of logical (Maciejovsky & Budescu, 2007; Moshman & Geil, 1998), mathematical (Laughlin & Ellis, 1986), inductive (e.g., Laughlin, Bonner, & Miner, 2002) or factual (Liberman, Minson, Bryan, & Ross, 2011; Sniezek & Henry, 1989) problems. Moreover, this is true not only of adults but also of children (e.g., Doise & Mugny, 1984; Miller & Brownell, 1975; Slavin, 1995).

The exchange of arguments plays a critical role in this process. Simply knowing of other participants' answers brings fewer benefits than being able to exchange reasons with them (Minson, Liberman, & Ross, 2011). Experiments have also demonstrated that the relative confidence of the group members cannot account for their convincingness (Trouche, Sander, & Mercier, in press): If some members manage to sway others, it is because they have better arguments (Trognon, 1993). If the group members agree on an answer prior to the discussion, little argumentation takes place and the group does not perform well (e.g., Schulz-Hardt, Brodbeck, Mojzisch, Kerschreiter, & Frey, 2006; Sunstein, 2002).

The experiments demonstrating the power of argumentation to bring groups to better decisions, however, have been conducted in the limited sample of the human population that Henrich, Heine, and Norenzayan (2010) have dubbed Western Educated Industrialised Rich Democratic (WEIRD) people. This population differs from other cultures on a variety of measures. In the broad domain of reasoning, some studies have found significant differences between the members of WEIRD cultures and those of other cultures (e.g., Easterners: Norenzayan, Smith, Kim, & Nisbett, 2002; unschooled participants: Medin & Atran, 2004; Scribner, 1977), while others have failed to observe any difference (e.g., Lee & Johnson-Laird, 2006; compare also Oaksford & Moussakowski, 2004 to Yama, 2001). In the present case, scholars have suggested the existence of wide variations in the status and role played by argumentation in different cultures (for review, see Mercier, 2011).

Argumentation plays a critical role in the most emblematic Western institutions—a heritage that can sometimes be traced as far back as the ancient Greeks. The most important trials are in the hands of juries who must deliberate before reaching a verdict. The people's representatives debate new laws. Science relies on argumentation for a healthy criticism of current ideas. For most of history, rhetoric was a pillar of Western education.

Other cultures put less stress on argumentation in their functioning and their values. In particular, several factors are supposed to conspire to make of East Asian cultures a less hospitable ground for argumentation

(Becker, 1986; Nakamura, 1964; for review, see Mercier, 2011). Collectivistic cultures tend to frown on argumentation as a threat to social harmony. Cultural influences specific to this area—Confucianism, Daoism and Zen Buddhism—also promote a dim view of argumentation. Summarising research about East Asian cultures, Peng and Nisbett note that "there are social, historical, linguistic, and philosophical barriers to the acceptance of argumentation and debate as a method of intellectual discourse or as a strategy for the consideration of new proposals for social or political change" (Peng & Nisbett, 1999, p. 747).

The most extreme claims about East Asian exceptionalism in the domain of argumentation have been proven false. Contrary to Becker (1986, p. 75), there is no complete "lack of argumentation and debate in the far east" (see Branham, 1994; Najita & Koschmann, 1982). Contrary to Nakamura (1964, p. 543), East Asian languages do not constitute a major obstacle to "expressing logical conceptions" (see Harbsmeier, 1998).

On the other hand, some experiments suggest significant cultural variations in several relevant traits. In particular, East Asians seem more tolerant of contradiction than Westerners; as a result, they might be less inclined to start exchanging arguments with each other (Peng & Nisbett, 1999; although see Mercier, Zhang, Qu, Lu, & Van der Henst, in press). It is very plausible that members of East Asian cultures would be less prone than their Western counterparts to argue and challenge each other's views. As a result, discussing groups in East Asia might not outperform individuals, even on problems for which groups consistently outperform individuals in the West.

The question of whether members of East Asian cultures are as apt as Westerners to benefit from argumentation has both theoretical and practical import. The argumentative theory of reasoning (Mercier & Sperber, 2011) hypothesises that reasoning evolved mainly for argumentation: To enable people to find arguments to convince others, and to evaluate others' arguments. For this theory to be plausible, argumentation must generally allow the people with the most accurate beliefs to convince those with less accurate beliefs. The evidence of superior group performance brought about by argumentation validates this prediction. However, if these benefits are specific to Western cultures, they should not count as a support for an evolutionary theory making broader predictions. Accordingly, if a human population were found to be consistently unable to rely on argumentation to improve their beliefs or decisions, the argumentative theory of reasoning would be falsified (although see Norenzayan & Heine, 2005 on the requirements for universals).

The practice of collaborative—or cooperative—learning largely rests on the cognitive benefits that can be derived through argumentation (see, e.g., Nussbaum, 2008). This practice has become firmly established in some Western countries (for the US, see Slavin, 1995) and has also been introduced in many others (e.g., Latin America: Brown & Brown, 1995; Japan: Sugie, 1995; South

Africa: Taylor, 1995). However, if students in East Asian countries are less likely to benefit from argumentation, or if argumentation instead jeopardises relations between students, attempts to use collaborative learning there might prove counterproductive (see, e.g., Nguyen, Terlouw, & Pilot, 2005).

We describe two experiments that test the hypothesis that cultural factors would hinder East Asians from reaping the benefits of argumentation. Each experiment uses a paradigm for which the superiority of groups over individuals has been robustly established among Western participants. The first compares the performance of groups and individuals when solving the Wason selection task. The Wason selection task (Wason, 1966) is a logical task relying on an understanding of conditionals that has proven extremely difficult for individuals to solve, with a rate of correct answers typically found hovering around 10%. By contrast, small groups instructed to reach a consensus on the answer reach an average of 63% of correct answers (Mercier, Trouche, Yama, Heintz, & Girotto, in press).

The second task is one of opinion aggregation by pairs. In this paradigm, participants first have to make a series of individual estimates, such as the amount spent on food by a family of four in a month (Liberman et al., 2011; Minson et al., 2011). Pairs of participants exchange a sheet with their opinions, and form a new opinion based on that of the other pair member. The pairs are then allowed to discuss their answers, attempting to come to an agreement. This manipulation enables a comparison of the effects of merely knowing other people's opinion and being able to use argumentation to calibrate the weight to give these opinions. The results show that discussion improves estimates, compared to not only participants' initial estimates, but also the estimates they had formed after learning of another participant's opinions.

Japanese culture is more collectivistic than most Western cultures (Triandis, 1995; although Japanese collectivism has sometimes been exaggerated, see Hasegawa & Hirose, 2005) and it has been strongly influenced by the intellectual movements mentioned above (Nakamura, 1964). As a result, it constitutes a relevant comparison point with Western cultures for the question in hand.

An important factor that might constrain the benefits of argumentation in Japan is whether one's interlocutors come from one's ingroup or an outgroup. Cross-cultural research has shown that the ingroup–outgroup boundary is more important for members of collectivistic cultures such as Japan (Triandis, 1989). The ingroup or outgroup origin of participants' interlocutors could affect argumentation in several ways. On the one hand, participants might tend to ignore the point of view of an outgroup, even if it is well-argued, reflecting the low level of general trust observed in Japan (Yamagishi & Yamagishi, 1994). On the other hand, a disagreement between ingroups might be perceived as particularly threatening to social harmony, an important value in East Asian cultures (Kim & Markus, 1999; Markus &

Kitayama, 1991). As a result, participants might promptly agree on an average value between their individual answers rather than engage in a thorough discussion of the relative values of their answers, thereby forgoing most of the benefits of the discussion. The ingroup–outgroup factor is manipulated in Experiment 2.

EXPERIMENT 1. GROUP RESOLUTION OF THE WASON SELECTION TASK IN JAPAN

Methods

Participants. The participants were 74 undergraduates from a Japanese university. Participants were part of two classes and took part in the experiment during class.

Materials. The task used was the abstract version of the Wason selection task. Participants were presented with the representation of four cards on a sheet of paper, with the following symbols, one on each card: A, 7, D and 4. They were told that each card had a letter on one side and a number on the other. Below the letters was the following rule: "If a card has a vowel on one side then it has an even number on the other side." Students were asked to test the rule by selecting the card(s) they would need to turn over to determine whether the rule was true or false for the cards shown. The correct answer is to select the A and the 7 cards. A similar Japanese translation of the task had already been used in other experiments (Yama, 2001). Half of the participants were also asked to estimate typical performance in this task. This question did not affect performance and its results will be analysed in another publication.

Design. All participants solved the task individually and then in groups. Previous research shows that the group effects obtained with this design are similar to those obtained with a between-participants design (Moshman & Geil, 1998).

Procedure. Participants were asked if they agreed to take part in an experiment investigating reasoning and problem solving (all agreed). They received a sheet of paper describing the problem and were instructed to solve it on their own. Participants were given at least 10 minutes to do so.

After the first sheets were gathered, 17 groups of three to six participants were formed at random. A new sheet displaying the same problem was distributed. Participants were asked to solve the problem again, this time discussing it with the group, trying to reach a consensus. Up to 20 minutes were provided before the sheets were gathered.

Results and discussion

In the individual condition, 15 out of 74 participants provided the correct answer (20%). The modal wrong answers typically observed in the previous literature were also the modal wrong answers here (A, 15 participants; A and 4, 15 participants) (see Table 1). One group failed to reach a unanimous answer in the allotted time. Counting only the unanimous groups, 10 out of 16 provided the correct answer (62%). Counting a majority of correct answers as correct, 11 out of 17 groups provided the correct answer (65%). In both cases the difference in performance with the individual condition is highly significant (respectively $\chi^2(1, N = 90) = 13.16$, $p = .0003$ and $\chi^2(1, N = 91) = 15.16$, $p = .0001$). It is important to note that in no group had the majority of the members found the right answer individually. In 12 groups, a minority of the members had found the correct answer individually. Eight of these groups gave the correct answer. In four groups the members held all the pieces of the correct answer from their individual answers (i.e., at least one of the members had done at least one of each of the following: Selected A, selected 7, not selected D, not selected 4). Three of these groups provided the correct answer. In the only non-unanimous group, the correct answer was initially present neither in complete nor in distributed form, yet three out of four group members gave it after discussion.

Only one difference between the effects of group discussion in the West and in Japan emerged from these results. In experiments conducted in the West, groups are more likely to solve the task if one of the members has found the correct answer individually than if no one has (11 out of 13 vs. 7 out of 27, $\chi^2(1, N = 40) = 3.01$, $p = .0079$, based on the data of Maciejovsky & Budescu, 2007; Moshman & Geil, 1998). This difference was not found in Japan (8 out of 12 vs. 3 out of 5). While the small number of groups in the Japanese sample forbids drawing any strong conclusion, this tentative difference suggests that a single individual with the right answer might be more likely to yield to the group in Japan, but that Japanese participants might be better able to put together the correct answer when pieces of it are distributed among the group members.

TABLE 1
Percentages of answers in Experiment 1, for individual and group problem solving

	A and 7	*A*	*A and 4*	*7*	*4*	*All four cards*	*A, 4 and 7*	*Other patterns*
Individual answer	20%	20%	20%	15%	8%	8%	0%	8%
Answer in group	62%	18%	7%	0%	0%	5%	8%	0%

In any case, the results show that Japanese participants are able to use group discussion to let the members with the best answer—whether it is the whole answer or part of it—convince other members, yielding an increase in group performance compared to individual performance.

EXPERIMENT 2. OPINION AGGREGATION THROUGH DISCUSSION IN JAPAN

To test the robustness of the conclusion drawn from Experiment 1, in Experiment 2 Japanese participants were asked to discuss not the solution of a logical problem but numerical estimates—the weight of various animals. Contrary to the Wason selection task, this type of problem does not admit of an accessible and demonstrable correct answer: Participants can argue for the greater plausibility of their estimates, rather than for the logical correctness of their answer.

In addition, in order to check whether the benefits of argumentation are robust to the origin—ingroup or outgroup—of the interlocutors, we used a minimal grouping manipulation and made participants discuss their estimates with participants from either their group or another group. Minimal grouping manipulations have been shown to affect a variety of outcomes for Japanese participants (Yamagishi, Jin, & Miller, 1998).

Methods

Participants. The participants were 90 undergraduates from a Japanese university. The experiment took part during class.

Materials. Participants were asked to estimate the weight of nine animals. Pictures of nine zoo animals that had been previously weighted by the keepers were taken for the purpose of the experiment.

Design. Following the minimal grouping manipulation, approximately half of the participants were paired with participants from their teams, and the other half paired with participants from the other team.

Each pair of participants then carried out the four rounds of the experiment.

- Round 1: Initial individual estimates.
- Round 2: Estimates informed by the estimates of the other member of the pair, without discussion.
- Round 3: Estimates formed during discussion with the other member of the pair, with instruction to reach an agreement.
- Round 4: Final individual estimates.

Procedure. First, the minimal grouping manipulation was performed. When participants entered the classroom, pairs were formed and participants found on their seats packets with the materials of the experiment and a red or a blue tag. The tags were laid out in such a way that about half the pairs had the same colour and half the pairs had different colours. Participants were asked to wear their tags and told that they would be used in a second experiment in which two teams would be formed based on the colour of the tags. Participants were told that in this second experiment members of each team would collaborate with each other and compete with the members of the other team. The blue and red teams were called in turn and the participants from each team were asked to raise their hands, so that each participant's allegiance was made clear and salient. Before starting the first experiment, participants were told that the second experiment would take place straight after the first, and so they should wear their tags to be ready to proceed as soon as possible.

Participants were then introduced to the first—the actual—experiment. They were told that their task was to estimate the weight of various animals based on their pictures, and that they would be asked to do so four times. The different phases of the experiment, as outlined in the design section here, were explained to them. They were also informed of the reward structure: At each round, they could win ¥3000, but ¥100 would be taken off for each percentage point they are off on a question. After participants had completed the four rounds, they were debriefed.

Results

All estimates were normalised so that the correct answer corresponds to 1 (i.e., each answer was divided by the correct answer). Without this procedure estimates of the weight of heavier animals would overly affect the average error. The error was then computed by taking the absolute value of the difference between the normalised estimate and the correct answer. Round 1 estimates that were over two standard deviations from the mean were treated as outliers. Each time an outlier occurred, the whole series of estimates (Rounds 1–4) was discounted, along with the symmetrical series from the other member of the pair. Twenty-three outliers were identified so that 46 sets of four answers were deleted (out of 810 sets, or 5.7%).

Since they had to be analysed at the level of the pair for the Round 3 estimates that were reached in agreement, all results were analysed at that same level. For each pair and for each round, an average error was computed. Figure 1 plots the evolution of the average error as a function of condition: Ingroup (participants were paired with a member of the same team) or Outgroup (participants were paired with a member of the other team).

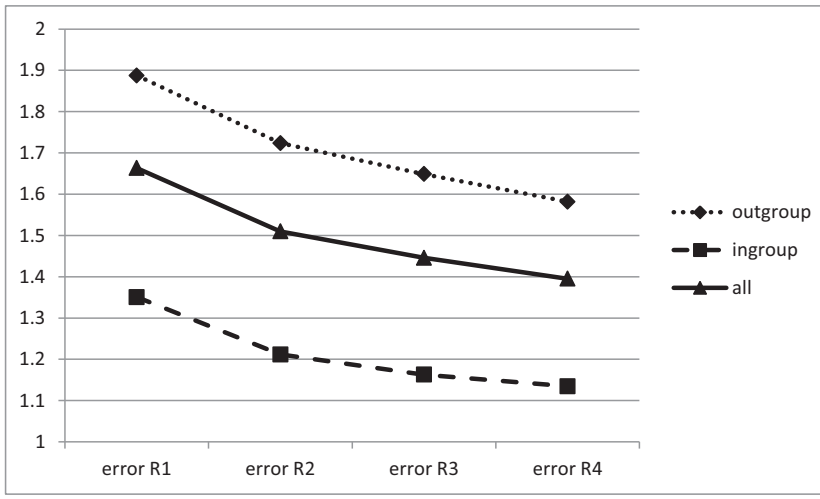

Figure 1. Evolution of the average error from round to round in both conditions and for all the participants.

Visual inspection of Figure 1 suggests that the minimal grouping manipulation only affected the general error level rather than the way the partner's estimates were evaluated, whether when presented on their own or during discussion. At all four rounds participants from the Ingroup condition were more accurate than participants from the Outgroup condition (all four $t(44) > 2.35$, $p < .023$). We suggest potential explanations for this finding in the discussion. By contrast, regarding the percentage of improvement from R1 to R2, to R3 or to R4, no difference emerges between the two conditions (all three $t(44) < 0.51$, $p > .60$). Given that only round to round evolution is presently relevant, all subsequent analyses combine the results of the two conditions.

There was a significant improvement from Round 1 to Round 2 (from $M = 1.66$ to $M = 1.51$, $t(89) = 4.31$, $p < .0001$) and from Round 3 to Round 4 (from $M = 1.45$ to $M = 1.40$, $t(89) = 2.39$, $p = .019$). Statistically, the improvement from Round 2 to Round 3 was only a trend ($t(89) = 1.82$, $p < .072$), but that from Round 2 to Round 4 was significant ($t(89) = 2.20$, $p = .002$).

To explore further the cause of the Round 3 to Round 4 improvement we looked at whether, on questions showing such an improvement, participants were more likely to revert towards their Round 1 estimate or to move further away from their Round 1 estimate. For instance, a participant who had answered 12 kg in Round 1 and 8 kg in Round 3 could improve on her Round 3 estimate either by reverting towards her Round 1 position—with a Round 4 estimate of 10 kg for instance. For another animal, a participant in

a similar situation could improve on her Round 3 estimate by moving further away from her Round 1 estimate—with a Round 4 estimate of 6 kg for instance. There were 148 occurrences of Round 3 to Round 4 improvements. In 16 of these cases participants had given the same answer in Round 1 and Round 3, so that they could neither revert back nor move away from their Round 1 estimate. Out of the remaining 132 occurrences, 103 consisted in reversions towards Round 1 estimate and 29 of moves away from Round 1 estimate. On average, pairs had 13% chances of improving from Round 3 to Round 4 by reverting towards their Round 1 estimate and 4% chances of improving from Round 3 to Round 4 by moving away from their Round 1 estimate, a highly significant difference ($t(44) = 7.31$, $p > .00001$).

Discussion

Replicating previous results gathered both in the West (for review, see Bonaccio & Dalal, 2006) and Japan (Mercier, Yama, Kawasaki, Adachi, & Van der Henst, 2012), Japanese participants proved able to take others' opinions into account to improve their estimates. The presentation of another participant's estimates enabled the participants to improve their estimates (Round 1 to Round 2). Since the only novel information available between Round 2 and Round 4 was acquired during discussion, the Round 2 to Round 4 improvement demonstrates a positive effect of the discussion, even though the Round 2 to Round 3 improvement was only a trend.

These results suggest, however, that Japanese participants derive benefits from discussion in a different way than Westerners. In a previous experiment using the same methodology but conducted in Israel (Minson et al., 2011), the benefits from discussion were immediately visible at Round 3 and there was little difference between the outcome of the conversation (Round 3) and the final individual answer (Round 4). This suggests that Israeli participants were genuinely endorsing the outcome of the group discussion and felt no need to change their mind when answering privately.

Japanese participants derived fewer immediate benefits from the discussion, with only a trend in improvement from Round 2 to Round 3. By contrast, they enjoyed solid benefits from Round 4, which was significantly more accurate than either Round 2 or Round 3. This could be interpreted in two ways. Participants could be too willing to cede ground during the discussion, yielding for social reasons rather than epistemic ones: For instance, even though they only face weak arguments, they want to preserve harmony. Recognising having yielded for reasons unrelated to accuracy, they would privately revert towards their initial estimates. Alternatively, one of the participants could refuse to budge during the discussion, again for social reasons rather than epistemic ones: For instance, even though they face strong arguments, they want to affirm their status. In this case, recognising that

they have failed to budge for reasons unrelated to accuracy, they would privately yield and move away from their initial estimates.

The first interpretation is more in line with the pressure to preserve harmony that supposedly pervades collectivistic cultures. It is also bolstered by the fact that when an estimate improved from Round 3 to Round 4, it was more often because of a reversion reverted towards Round 1 estimate rather than moved away from it. This suggests that during the discussion participants were yielding for social reasons but that they had enough awareness of this process to revert back to their original estimates when appropriate.

The only significant correlate of the minimal grouping manipulation was the baseline accuracy. There could be two explanations for this finding. One is that the participants who expected to interact with an ingroup were more motivated to provide an accurate answer. Another was that participants who expected to work with an outgroup might have been disrupted by the implicit social competition. The literature on the effect of working with an ingroup vs. an outgroup has yielded conflicting results, depending on the difficulty of the task (e.g., Lount & Phillips, 2007). Moreover, none of these results were acquired in an Eastern culture. As a result, the present effect of the minimal grouping manipulation seems difficult to interpret. Given that this effect was not the main focus of our investigation, we will not dwell on it further.

CONCLUSION

So-called WEIRD people derive solid benefits from argumentation. In a variety of tasks, groups consistently outperform individuals. One explanation is that humans are endowed with reasoning mechanisms that are tailored to such a task: To express arguments defending one's point of view and to evaluate others' arguments to change one's mind when appropriate. Another explanation is that these WEIRD people belong to cultures that value argumentation and train their members to make the best of it. Other cultures, by contrast, might avoid the conflicts necessary for argumentation to emerge and find its practice contrary to widely accepted cultural values. East Asians in particular have often been thought to frown on argumentation.

To test whether East Asians were able to reap the benefits of argumentation, we replicated previous demonstrations of the efficiency of argumentation among Japanese participants.

Experiment 1 showed that Japanese participants perform much better on a classic logical problem (the Wason selection task) when solving the task in small groups than individually. Experiment 2 showed that Japanese participants use discussion in pairs to arrive at more accurate numerical estimates. In both cases the overall efficacy of group discussion was similar to that observed in experiments conducted in the West.

We thus observed broadly similar improvements to those observed in WEIRD cultures. In WEIRD cultures, substantial evidence shows that these improvements stem from argumentation (for tasks similar to that used in Experiment 1, see Mercier & Sperber, 2011; Trouche et al., in press; for tasks similar to that used in Experiment 2, see Rowe & Wright, 1996). An economical interpretation is that similar patterns of performance in both cultures are driven by the same processes. It could be argued, however, that in fact the Japanese participants relied on different mechanisms.

Although the current experiments cannot conclusively show that argumentation was at play, some results suggest that it is likely to have been the case. In Experiment 1, there was a tendency for groups to adopt the correct answer, rather than any of the other answers originally present in the group. This was true even when the correct answer was only defended by one individual. It is not clear to us how that individual could have managed to get her answer adopted by the others except through argumentation: The Wason selection task is not a eureka type problem, and simply being presented with what happens to be the correct selection has not been shown, to the extent of our knowledge, to change anyone's mind.

In Experiment 2, the fact that discussion (ultimately) led to improved performance beyond simply learning about another participant's answer shows that the discussion played a role. It is possible, however, that the discussion improved performance only by allowing participants to realise which participant was the most confident, assuming that confidence and accuracy were correlated (see Koriat, 2012). More evidence will have to be gathered to completely rule out this hypothesis.

Both experiments also pointed to interesting cultural differences. The results of Experiment 1 weakly suggest that a Japanese participant who alone has the correct answer might be less likely to convince her peers than a WEIRD participant in a similar situation, but that Japanese groups are more likely to find the correct answer on the basis of pieces of good answer distributed among the group members. If confirmed, these effects would be consistent with a greater role of consensus and social harmony in collectivistic cultures: Japanese participants would strive harder to arrive at a satisfying consensus, whether this means accepting the opinion of the majority or taking everyone's opinion into account when forging a consensus (as in the practice of *nemawashi* for instance, see Saito, 1982).

In Experiment 2, Japanese participants provided estimates during discussion that failed to make the best of the discussion. However, when they could give private answers after the discussion, their estimates significantly improved. This improvement was chiefly due to frequent reversions towards the participants' initial estimates. A likely interpretation of these findings is that Japanese participants sometimes yielded to their partner for reasons they recognised as being irrelevant to the accuracy of their answers, such as

preserving social harmony. When they were offered to give final, private estimates, these participants could express what they thought was the best estimate, which typically meant reverting towards their own initial estimates.

This research bolsters the prospect for collaborative learning and other uses of argumentation to reach better outcomes in Japan and other East Asian cultures. It also suggests that some precautions should be taken both to measure the effects of argumentation and to make the best of argumentation in East Asian cultures. Experiment 2 shows that Japanese participants might hold a private opinion that is more accurate than the opinion they are willing to express in front of a peer. To get at this opinion, the use of a private, post-discussion question is advisable, whether it is a repeat of the question asked during the discussion or a transfer problem. Both experiments, in line with previous cross-cultural research, also suggest that Japanese participants put more weight than Westerners on the preservation of social harmony during group discussion. On the one hand, by precluding members from expressing minority views, this could reduce the benefits of group discussion. On the other hand, if these attempts are effective, they could also guarantee that group members are able to maintain positive relations for longer, a critical achievement if the same people are asked to interact repeatedly, for instance throughout a school year.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

Manuscript received 28 October 2014
Revised manuscript received 11 December 2014
Revised manuscript accepted 21 December 2014
First published online 20 January 2015

REFERENCES

Becker, C. B. (1986). Reasons for the lack of argumentation and debate in the Far East. *International Journal of Intercultural Relations, 10*(1), 75–92.

Bonaccio, S., & Dalal, R. S. (2006). Advice taking and decision-making: An integrative literature review, and implications for the organizational sciences. *Organizational Behavior and Human Decision Processes, 101*(2), 127–151.

Branham, R. J. (1994). Debate and dissent in late Tokugawa and Meiji Japan. *Argumentation and Advocacy, 30*(3), 131–149.

Brown, W., & Brown, P. B. (1995). Cooperative learning in Latin America: A perspective from the alternative education experience. *International Journal of Educational Research, 23*(3), 255–265.

Doise, W., & Mugny, G. (1984). The *s*ocial *d*evelopment of the *i*ntellect. Oxford: Pergamon Press.

Harbsmeier, C. (1998). *Language and logic*. Cambridge, MA: Cambridge University Press.

Hasegawa, Y., & Hirose, Y. (2005). What the Japanese language tells us about the alleged Japanese relational self. *Australian Journal of Linguistics, 25*(2), 219–251.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world. *Behavioral and Brain Sciences, 33*(2–3), 61–83.

Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology, 55*, 623–655.

Kim, H., &Markus, H. R. (1999). Deviance or uniqueness, harmony or conformity? A cultural analysis. *Journal of Personality and Social Psychology, 77*(4), 785–800.

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

Laughlin, P. R., Bonner, B. L., & Miner, A. G. (2002). Groups perform better than the best individuals on letters-to-numbers problems. *Organizational Behavior and Human Decision Processes, 88*, 605–620.

Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and social combination processes on mathematical intellective tasks. *Journal of Experimental Social Psychology, 22*, 177–189.

Lee, N. Y. L., & Johnson-Laird, P. N. (2006). Are there cross-cultural differences in reasoning? In R. Sun (Ed.), *Proceedings of the 28th Annual Cognitive Science Conference* (pp. 459–464). Mahwah, NJ: Lawrence Erlbaum Associates.

Liberman, V., Minson, J. A., Bryan, C. J., & Ross, L. (2011). Naïve realism and capturing the "wisdom of dyads." *Journal of Experimental Social Psychology, 48*(2), 507–512.

Lount, R. B., Jr., & Phillips, K. W. (2007). Working harder with the out-group: The impact of social category diversity on motivation gains. *Organizational Behavior and Human Decision Processes, 103*(2), 214–224.

Maciejovsky, B., & Budescu, D. V. (2007). Collective induction without cooperation? Learning and knowledge transfer in cooperative groups and competitive auctions. *Journal of Personality and Social Psychology, 92*(5), 854–870.

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*(2), 224–253.

Medin, D. L., & Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. *Psychological Review, 111*(4), 960–982.

Mercier, H. (2011). On the universality of argumentative reasoning. *Journal of Cognition and Culture, 11*, 85–113.

Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences, 34*(2), 57–74.

Mercier, H., Trouche, E., Yama, H., Heintz, C., & Girotto, V. (in press). Experts and laymen grossly underestimate the benefits of argumentation for reasoning. *Thinking & Reasoning*.

Mercier, H., Yama, H., Kawasaki, Y., Adachi, K., & Van der Henst, J.-B. (2012). Is the use of averaging in advice taking modulated by culture? *Journal of Cognition and Culture, 12*(1–2), 1–16.

Mercier, H., Zhang, J., Qu, Y., Lu, P., & Van der Henst, J.-B. (in press). Do Easterners and Westerners treat contradiction differently? *Journal of Cognition and Culture*.

Miller, S. A., & Brownell, C. A. (1975). Peers, persuasion, and Piaget: Dyadic interaction between conservers and nonconservers. *Child Development, 46*(4), 992–997.

Minson, J. A., Liberman, V., & Ross, L. (2011). Two to tango. *Personality and Social Psychology Bulletin, 37*(10), 1325–1338.

Moshman, D., & Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Thinking & Reasoning, 4*(3), 231–248.

Najita, T., & Koschmann, J. V. (1982). *Conflict in modern Japanese history: The neglected tradition*. Princeton, NJ: Princeton University Press.

Nakamura, H. (1964). *Ways of thinking of Eastern peoples: India, China, Tibet, Japan*. Honolulu, HI: University of Hawaii Press.

Nguyen, P. M., Terlouw, C., & Pilot, A. (2005). Cooperative learning vs Confucian heritage culture's collectivism: Confrontation to reveal some cultural conflicts and mismatch. *Asia Europe Journal, 3*(3), 403–419.

Norenzayan, A., & Heine, S. J. (2005). Psychological universals: What are they and how can we know. *Psychological Bulletin, 131*(5), 763–784.

Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive Science, 26*(5), 653–684.

Nussbaum, E. M. (2008). Collaborative discourse, argumentation, and learning: Preface and literature review. *Contemporary Educational Psychology, 33*(3), 345–359.

Oaksford, M., & Moussakowski, M. (2004). Negations and natural sampling in data selection: Ecological versus heuristic explanations of matching bias. *Memory & Cognition, 32*(4), 570–581.

Peng, K., & Nisbett, R. E. (1999). Culture, dialectics and reasoning about contradiction. *American Psychologist, 54*(9), 741–754.

Rowe, G., & Wright, G. (1996). The impact of task characteristics on the performance of structured group forecasting techniques. *International Journal of Forecasting, 12*(1), 73–89.

Saito, M. (1982). Nemawashi: A Japanese form of interpersonal communication. *ETC: A Review of General Semantics, 39*(3), 205–214.

Schulz-Hardt, S., Brodbeck, F. C., Mojzisch, A., Kerschreiter, R., & Frey, D. (2006). Group decision making in hidden profile situations: Dissent as a facilitator for decision quality. *Journal of Personality and Social Psychology, 91*(6), 1080–1093.

Scribner, S. (1977). Modes of thinking and ways of speaking: Culture and logic reconsidered. In P. N. Johnson-Laird & P. C. Wason (Eds.), *Thinking: Readings in cognitive science* (pp. 483–500). New York, NY: Cambridge University Press.

Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice* (Vol. *2*). London: Allyn & Bacon.

Sniezek, J. A., & Henry, R. A. (1989). Accuracy and confidence in group judgment. *Organizational Behavior and Human Decision Processes, 43*(1), 1–28.

Sugie, S. (1995). Cooperative learning in Japan. *International Journal of Educational Research, 23*(3), 213–225.

Sunstein, C. R. (2002). The law of group polarization. *Journal of Political Philosophy, 10*(2), 175–195.

Taylor, C. A. (1995). Cooperative learning in an African context. *International Journal of Educational Research, 23*(3), 239–253.

Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review, 96*(3), 506–520.

Triandis, H. C. (1995). *Individualism & collectivism*. Boulder, CO: Westview Press. Retrieved from http://psycnet.apa.org/psycinfo/1995-97791-000

Trognon, A. (1993). How does the process of interaction work when two interlocutors try to resolve a logical problem? *Cognition and Instruction, 11*(3&4), 325–345.

Trouche, E., Sander, E., & Mercier, H. (in press). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General.*

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology: I* (pp. 106–137). Harmandsworth: Penguin.

Yamagishi, T., Jin, N., & Miller, A. S. (1998). In-group bias and culture of collectivism. *Asian Journal of Social Psychology, 1*(3), 315–328.

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion, 18*(2), 129–166.

Yama, H. (2001). Matching versus optimal data selection in the Wason selection task. *Thinking & Reasoning, 7*(3), 295–311.