



Original Article

Evidence for benefits of argumentation in a Mayan indigenous population

Thomas Castelain^{a,b,c}, Vittorio Girotto^d, Frank Jamet^e, Hugo Mercier^{a,b,*}^a Centre de Sciences Cognitives, Université de Neuchâtel, Neuchâtel, Switzerland^b Institut des Sciences Cognitives–Marc Jeannerod, L2C2, Centre National de la Recherche Scientifique, Lyon, France^c Instituto de Investigaciones Psicológicas, Universidad de Costa Rica, San José, Costa Rica^d Center for Experimental Research in Management and Economics, DCP, University IUAV of Venice, Venice, Italy^e Chart–Cognitions humaine et artificielle, Université de Cergy-Pontoise, Cergy-Pontoise, France

ARTICLE INFO

Article history:

Initial receipt 3 July 2015

Final revision received 12 February 2016

Keywords:

Reasoning

Argumentation

Myside bias

Traditional populations

ABSTRACT

Group discussion improves on individual reasoning performance for a wide variety of tasks. This improvement, however, could be largely specific to members of modern, schooled, affluent Western cultures. In two studies, we observed the same improvement in the members of a traditional population—indigenous Maya from Guatemala. Two features of reasoning can account for this improvement: the myside bias, which precludes individuals from improving their performance on their own, and the ability to soundly evaluate others' arguments, which allows individuals to benefit from group discussions. These two features were observed in the traditional population studied: solitary reasoning performance was marked by the myside bias; individuals were more likely to be convinced by arguments for the correct answer rather than for a wrong answer. Together with previous evidence, the present results strengthen the conclusion that these features are adaptive features of reasoning.

© 2016 Elsevier Inc. All rights reserved.

At least since Descartes, reasoning has generally been understood as a tool of individual cognition: by carefully evaluating and weighing one's reasons, one should arrive at sounder beliefs and better decisions (Descartes, 1637; Kahneman, 2003; Stanovich, 2004). Opposed to this individualistic view, an alternative is that reasoning's main function is social: to find reasons in order to convince others, and to evaluate others' reasons in order to adopt better supported beliefs. The argumentative theory of reasoning (Mercier & Sperber, 2011) is a recent instantiation of this social view of reasoning (for other social views of reasoning, see Baumeister & Masicampo, 2010; Billig, 1996; Gibbard, 1990; Piaget, 1928).

An important result supporting the argumentative theory of reasoning is that, provided minimal conditions are met (e.g. the expression of disagreement, see Janis, 1982), argumentation in the course of group discussion produces sizeable improvements in reasoning performance on a variety of tasks, such as logical, mathematical, and inductive problems (Laughlin, 2011; Moshman & Geil, 1998; Trouche, Sander, & Mercier, 2014), work related tasks (Mercier, 2011c), forecasting (Mellers et al., 2014; Rowe & Wright, 1999), and school tasks (Mercier, 2011b; Slavin, 1995; Smith et al., 2009). The argumentative theory suggests that the gap in performance between individual reasoning and reasoning in discussion stems from a combination of two features of reasoning. The first is the myside bias (or confirmation bias):

individuals overwhelmingly produce reasons for their preferred opinions (Mercier, *in press*; Nickerson, 1998). As a result, reasoning rarely allows the lone reasoner to correct mistaken intuitions. Instead, the myside bias can lead to overconfidence (Koriat, Lichtenstein, & Fischhoff, 1980) and polarization (Tesser & Conlee, 1975). The second feature is reasoning's ability to soundly evaluate others' arguments, rejecting weak arguments and accepting strong enough ones (Hahn & Oaksford, 2007; Petty & Wegener, 1998). In a discussion group members only provide arguments for their side, but they also evaluate each other's arguments. They change their minds when the arguments are good enough, which usually means changing their mind for the best.

These features of reasoning, along with the gap in reasoning performance they seek to explain, could be a peculiarity of WEIRD (Western Educated Industrialized Rich Democratic) cultures. These cultures, in which the research mentioned above has been conducted, differ from other cultures on a number of traits (Henrich, Heine, & Norenzayan, 2010). In particular, participants from WEIRD cultures are sometimes at the far end of the distribution—for instance in terms of individualism (Henrich et al., 2010). Schooling, which is comprehensive in WEIRD cultures but absent from many other cultures, can exert a profound influence on cognition, for instance through the acquisition of literacy (Dehaene, 2009) and numeracy (Dehaene, 1999; Gordon, 2004).

Regarding reasoning and argumentation, WEIRD cultures have a series of traits that, while not necessarily specific to these cultures, might conspire to create the features mentioned above. Compared to many other cultures, members of Western cultures: (1) have long put a higher value on argumentation in their institutions, from science to law or

* Corresponding author. Centre de Sciences Cognitives, Université de Neuchâtel, Rue Pierre à Mazel 7, 2000 Neuchâtel, Switzerland.

E-mail address: hugo.mercier@unine.ch (H. Mercier).

politics (Lloyd, 1996; Peng & Nisbett, 1999; especially compared to Eastern cultures, see Becker, 1986; Nakamura, 1964); (2) put relatively less stress on face-saving and social harmony (Kim & Markus, 1999; Oetzel et al., 2001), which might allow for more confrontational and open debates (Mercier, Deguchi, Van der Henst, & Yama, 2015); (3) adopt a different parenting style in which children tend to argue more with adults: they question more freely their decisions, and receive more explanations—although this difference seems to be restricted to the middle and upper classes (Chouinard, Harris, & Maratsos, 2007; indeed, they talk more with adults generally, Hart & Risley, 1995; Heath, 1983). All of these factors might make argumentation a cognitive skill particularly valued in participants from WEIRD cultures.

The *myside bias* might be a cognitive response to a specific cultural environment in which argumentation is highly valued and it is particularly important to be able to defend one's point of view. Two other traits might help explain the existence of a *myside bias* in WEIRD cultures. First, in WEIRD cultures, individuals are confronted with a variety of choices, values, and worldviews. In such cultures, it makes sense to anticipate having to defend one's choices, since it is likely that one will encounter people who have made different choices (Schwartz, 2004). Second, many of these choices mostly have a symbolic value, so that it arguably matters more to make a decision that is socially acceptable than an intrinsically good decision. For instance, reasoning has been shown to lead customers towards products they enjoy less, but which allow them to be perceived more positively by others (Thompson & Norton, 2008; for review, see Mercier & Sperber, 2011). Even socially consequential choices, such as voting, might mostly have a symbolic value (Sears, Lau, Tyler, & Allen, 1980). The symbolic value of voting, and of political opinions more generally, might help explain why voters in WEIRD cultures have a consistent *myside bias* when reasoning about politics: it matters more that they have arguments to justify their opinions than that they make otherwise sound opinions (Kahan et al., 2012; Taber & Lodge, 2006).

WEIRD cultures thus seem to have a number of traits that favor the development of argumentative skills, traits which in turn might favor the emergence of a *myside bias*. In order to test whether these cultural traits partly or entirely account for the relevant features of reasoning—the efficiency of argumentation and the deficiencies of solitary reasoning—it is necessary to study reasoning and argumentation in a population that does not share these cultural traits. We first argue that a broad type of cultures—which we will call, in an ad-hoc fashion, traditional cultures—differs in many relevant ways from WEIRD cultures. Then we will argue that the population to which the participants recruited here belong, the K'iche' Maya in Guatemala, share these traits.

Here we define traditional cultures as the human groups which are the broad opposite of WEIRD cultures: small-scale groups which are not Western, educated, industrialized, or rich (note that many human groups would thus fall between WEIRD and traditional cultures as so defined). The fact that these cultures are not Western means that they do not share the hypothetical Western values which cast argumentation in a positive light (although they might hold other values to the same effect, see, e.g. Gluckman, 1967). Regarding lack of education (in the sense of formal schooling), one of its most relevant consequences is a reluctance to engage in hypothetical thinking. Experiments conducted in several unschooled populations have revealed that most of their members fail to complete even very simple hypothetical reasoning tasks (Cole, Gay, Glick, & Sharp, 1971; Luria, 1976; Scribner, 1977). This reluctance to engage in hypothetical thinking could hinder argumentation, since argumentation often relies on hypotheticals. Besides the lack of schooling, education also differs in traditional societies in other ways. Of particular relevance here is that parenting in traditional populations conforms to the pattern observed outside of middle and upper middle class WEIRD populations: relatively little talk addressed to children, and in particular a very small number of exchanges requiring reasons and explanations (Gauvain, Munroe, & Beebe, 2013). Because members of traditional cultures, compared to members of WEIRD cultures, might

value argumentation less, be more reluctant to engage in hypothetical thinking, and are not trained to argue early on, they might be less likely to develop the argumentative skills evidenced in the members of WEIRD cultures.

We argued above that the members of WEIRD cultures—particularly in the middle and upper middle classes—are faced with a great variety of choices, and that many of these choices are in large part symbolic. The choices facing members of traditional cultures differ in both respects. As a rule, members of these cultures have far fewer choices: far fewer (if any) products to buy, far fewer (if any) choices of occupation, far fewer (if any) choices of places to live, far fewer (if any) choices of religion, far fewer (if any) choices of people to befriend, etc. (see, Lévi-Strauss, 1966). This relative lack of choice suggests a lighter burden of justification. Members of traditional cultures might thus have less use of a *myside bias* which would help them defend their choices. Moreover, it has been argued that members of traditional cultures make more life and death decisions than members of WEIRD cultures (Diamond, 2012). Contrary to the members of WEIRD cultures, members of traditional cultures do not live only in a human created environment in which many natural dangers are eliminated (e.g. predation) or drastically reduced (e.g. many pathogens). This is reflected in the lower life expectancy in traditional cultures compared to WEIRD cultures (e.g. Gurven & Kaplan, 2007). It is thus plausible that for most decisions, intrinsic value matters more than symbolic value for the members of traditional cultures relative to members of WEIRD cultures. As a result, the *myside bias* could be disadvantageous for the former, as it leads to intrinsically worse decisions, and advantageous for the latter, as it leads to easier to defend symbolic decisions.

On the whole, it is thus plausible that members of traditional cultures, relative to members of WEIRD cultures, reason in a way that is more practical (see Luria, 1976; Medin & Atran, 2004) and more in line with the predictions of the individualistic view of reasoning. Relative to members of WEIRD cultures, members of traditional cultures might suffer less from the *myside bias*, and thus be more efficient solitary reasoners, while benefitting less from argumentation.

In light of the evidence available, the relevant traits of traditional cultures mentioned above appear to be present in the K'iche' Maya who participated in the present experiments. The K'iche' are a preliterate Maya Amerindian group living in rural areas of Guatemala. Men are mainly subsistence farmers and women do household maintenance work. They are obviously not Western, and there is no evidence that their culture puts a particular value on argumentation. The participants in our experiments had received no schooling as children and were in the process of receiving a very moderate amount of schooling as adults. The pattern of interactions with the children seems to follow the pattern typically found outside of middle and upper classes in WEIRD cultures. A study of the interactions between K'iche' adults and children revealed that most of the utterances adults address to children are imperative and very few are questions (Pye, 1986). There is thus little opportunity for argumentative exchanges between children and adults. Overall, the K'iche' do not seem to enjoy any of the cultural traits that might make WEIRD cultures particularly congenial to argumentation.

Like other small-scale societies relying on substance farming, the K'iche' face fewer choices than members of WEIRD cultures. Few products are available for purchase, and they have very little money to purchase them with (UNDP, 2010a, 2010b). The choice of occupation and place to live is extremely restricted (UNDP, 2010a, 2010b, 2012, 2014). Moreover, the environment is harsher than that faced by most members of WEIRD cultures (for instance 94% of the population in the relevant district—Sololá—lacks food security, see INE, 2011, p. 29; UNDP, 2010a; UNFPA, 2014). The risk of disease and early death is much higher (INE, 2011; UNDP, 2010a; UNFPA, 2014). A *myside bias* might thus be less adaptive in such an environment than in that of WEIRD cultures.

The improvement in performance yielded by argumentation, the *myside bias* which explains poor solitary performance, and the

argumentative skills which explain good performance in group discussion are the main pieces of evidence in favor of the argumentative theory of reasoning. This theory posits that reasoning is an evolved adaptation, and that these features, which make it adaptive, should be relatively universal (Mercier & Sperber, 2011). However, these features have mostly been documented among WEIRD cultures. We have argued above that several traits of Western cultures might conspire to explain these features, and that reasoning might exhibit different features in traditional populations such as the K'iche' Maya recruited in the present study. Experiments carried out in such populations can thus provide crucial evidence to understand the function of human reasoning. Accordingly, the goal of the present experiments is twofold. First, to test whether, in this traditional population, discussion improves performance compared to solitary reasoning. Second, if such an improvement was observed, to determine whether it could be explained by the two features of reasoning that explain the gap in reasoning performance in WEIRD populations: myside bias and sound ability to evaluate others' arguments.

1. Methods

1.1. Participants

One hundred and forty preliterate participants were involved in this research, 57 in study 1 (48 females, $M_{age} = 35.12$ years, $SD = 12.16$, range 18–62 years) and 83 in study 2 (78 females, $M_{age} = 43.89$ years, $SD = 12.49$, range 21–76 years). The participants were K'iche' from ten different villages around the city of Nahualá in the department of Sololá, Guatemala. The participants were recruited through an alphabetization program set up by the CONALFA (National Committee for Alphabetization) in which they take part. The program consisted of arithmetic and literacy classes dispensed in K'iche'. The participants were supposed to spend 3 hours per week to learn to read and write, first in their native language and later in Spanish, but few of them were able to attend the classes regularly. Nearly all participants ($n = 127$, 91%) were in their first year of alphabetization, the rest in their second year. Verbal informed consent was obtained from the participants. Each participant received a reward—seeds or sewing threads—that they had requested during the recruitment phase.

1.2. Design

Two studies were conducted and both had the same design (see Supplementary Materials, available on the journal's website at www.ehbonline.org for details). The tasks used in this research—conservation tasks—are designed to assess the understanding of the invariance of physical quantities across physical transformations (Dasen, 1972; Jamet, Baratgin, & Filatova, 2014; Piaget & Inhelder, 1974), and discussion has been shown to improve performance in WEIRD cultures (Doise & Mugny, 1984; Perret-Clermont, 1980). We used weight/volume conservation tasks, in which participants must indicate how the water level in two glasses changes when one object is introduced in each glass (Fig. 1A). The objects differ in some irrelevant dimension(s), such as weight or shape, but not volume; hence the correct answer is that the water rises as much in the two glasses. All the participants first completed two familiarization trials (whose presentation order was counterbalanced): In both tasks, the experimenter provided participants with feedback (no feedback was provided at any other point of the experiment, see Supplementary Materials, section Material and methods, Fig. S1, available on the journal's website at www.ehbonline.org). The participants then completed a series of three (study 1) or two (study 2) volume conservation tasks on their own (pre-test). Groups of three (study 1, $n = 19$, study 2, $n = 25$) or four (study 2, $n = 2$) individuals were then formed randomly. The groups were given the instruction to arrive at a consensual answer for each task (test). Finally, each participant had to solve the task on her own again (post-test) (see Supplementary Materials, Fig. S1, available on the journal's website at

www.ehbonline.org). In study 1, a fourth, new task was presented to the participants in the post-test, in order to test for transfer effects (Task 4).

The tasks used in study 1 (tasks 1, 2, and 3 at all phases, task 4 as a transfer task) were chosen based on results obtained by one of the authors (F.J.) with another traditional population (in Papua New Guinea), in which participants only had to complete the tasks on their own. We chose tasks for which there were no ceiling or floor effects in this other traditional population. The tasks used in study 2 were based on the results of study 1. We chose task 2 because it elicited a lower rate of performance than tasks 1 and 3 at the pre-test of study 1. Task 4 provided a baseline to evaluate the performance in study 1, in which it was used as a transfer task.

In each phase, the tasks were always presented in the same order. The positions of the objects (left and right) were the same during the pre and post-test, and were inverted in the test. During the pre- and the post-test, 114 participants were asked to justify their answers, while 26 were asked instead to think aloud. For all participants, the justifications and think aloud protocols in the pre- and post-test, as well as the discussion during test were recorded, transcribed, translated from K'iche' to Spanish, and coded by two individuals, one of whom was blind to the hypotheses (inter-coder reliability: Kappa = .79, $p < .001$).

1.3. Procedure

Two native female experimenters were trained by one of the authors (T.C.), and they conducted all the experiments. All the experiments were conducted in 2014. The experimenters interacted with the participants in their native language. The translations from Spanish to K'iche' were done by one of the experimenters, the back-translation by the other experimenter, with discrepancies resolved through discussion with one of the authors (T.C.). The experiments took place in the village of each community, in the rooms where the participants follow their alphabetization lessons. In order to ensure that the participants were not discussing the tasks outside of the experimental context, when they were not being tested they had to reintegrate their classroom and to continue to follow the lesson of the day. They were also instructed not to interact until the end of the experiment.

In the familiarization phase, the participants were presented with the material: two glasses and two pieces of clay (Play-Doh) (see Supplementary Materials, Fig. S1, available on the journal's website at www.ehbonline.org). We used the same materials in the following tasks, except for tasks 2 and 4 in which we used one piece of clay and one piece of metal. The experimenter started the familiarization phase by pouring the two glasses with water and invited the participants to observe that their level was equal. Then she presented the first pair of objects in front of the two glasses. In tasks 1 and 3, the transformations—cut the ball in two, turn it into a cylinder—were made in front of the participants, and the differences or similarities in weight or size were notified to them before each trial. The experimenter then asked the participants: “If I put the balls in the two glasses, their level of water will rise, do you think that it will rise more in this glass [The experimenter pointed to one of the glasses]? Do you think it will rise more in this glass [The experimenter pointed to the other glass]? Or it will rise the same in the two glasses?” All the tasks followed the same instructions. During the test, participants were invited to discuss the tasks in groups and to arrive at a consensual answer: “You can discuss now and agree on an answer.” After the discussion, the experimenter asked: “Do you all agree? What is your answer?”

2. Results

At each stage of the studies, we computed the individual scores, assigning 1 to correct answers and 0 to incorrect answers, and the group scores (addition of individual scores), using real groups for the tests, and the equivalent nominal groups for the pre- and post-tests (Fig. 1B, see also Supplementary Materials, section Results for details,

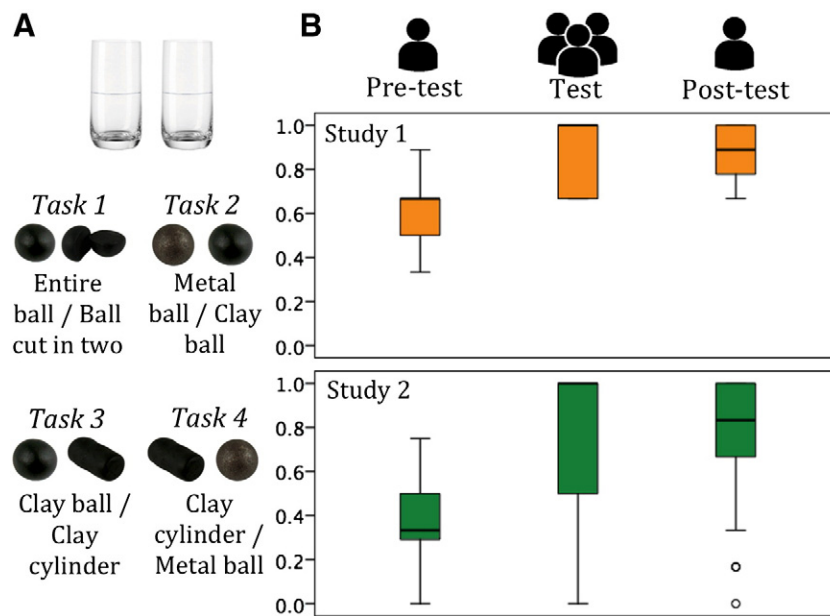


Fig. 1. (A) Illustration of the four tasks. (B) Box plots representing the normalized scores for the groups in study 1 (tasks 1, 2, and 3) and study 2 (tasks 2 and 4). The groups are the real discussion groups in the test and the nominal groups comprising the same participants in the pre- and post-tests. The dark line represents the median value. The boxes represent the middle half of each sample.

available on the journal's website at www.ehbonline.org). Comparing the individual scores, in both studies we observed a significant improvement in performance between the pre-test and the post-test (study 1: $Mdn_{Pre-test} = 2$, $Mdn_{Post-test} = 3$, $Z = 4.79$, $p < .001$, $r = .44$; study 2: $Mdn_{Pre-test} = 0$, $Mdn_{Post-test} = 2$, $Z = 5.33$, $p < .001$, $r = .43$). The pre-test/post-test improvement was also significant when we counted as correct only correct answers accompanied by a correct explanation, e.g. "because the objects have the same size" (study 1: $Mdn_{Pre-test} = 2$, $Mdn_{Post-test} = 3$, $Z = 5.04$, $p < .001$, $r = .45$; study 2: $Mdn_{Pre-test} = 0$, $Mdn_{Post-test} = 2$, $Z = 4.94$, $p < .001$, $r = .36$). Regarding group scores, in both studies we observed a significant improvement in performance between the pre-test and the test (study 1: $Mdn_{Pre-test} = 6$, $Mdn_{Test} = 9$, $Z = 3.64$, $p < .001$, $r = .67$; Study 2: $Mdn_{Pre-test} = 2$, $Mdn_{Test} = 6$, $Z = 3.80$, $p < .001$, $r = .53$) but not between the test and the post-test (study 1: $Mdn_{Test} = 9$, $Mdn_{Post-test} = 8$, $Z = 1.47$, $p = .14$; study 2: $Mdn_{Test} = 6$, $Mdn_{Post-test} = 5$, $Z = .064$, $p = .95$). Thus, group discussion significantly improved the reasoning performance of preliterate individuals.

To test whether the myside bias could explain wrong answers at the pre-test, we analyzed the reasons participants provided. The myside bias should limit the number of reasons that go against the participants' answers. Out of 141 explanations given by participants providing the wrong answer, 1 contained no reason, 134 contained only reasons supporting the participants' answer, and 6 contained reasons against it. This result could also reflect a task demand: participants were asked to explain their answers. Therefore, they might have failed to mention reasons that they had considered but dismissed. To test this possibility, a subset of participants had been asked not to explain their answers, but to think aloud while they were trying to solve the tasks. Out of the 32 explanations given by participants providing wrong answers in think aloud protocols, 3 mentioned reasons going against their final answer, 2 contained no reasons, while 27 contained only reasons supporting the participants' final answer (see Supplementary Materials, section Results for the detail of the coding, available on the journal's website at www.ehbonline.org).

To evaluate the soundness of argument evaluation, we contrasted the effects good and bad arguments had on the outcome of the discussion. An argument could convince other group members not because they think it is good, but because it is defended by the majority. To examine such majority effects, we first analyzed the 43 discussions in which a participant

who had given the correct answer and the correct justification in the pre-test faced group members who all agreed on the same wrong answer. In 33 (77%) of these discussions, the group agreed on the correct answer. Moreover, in 8 out of the 10 remaining discussions, no argument for the correct answer was offered. We also analyzed the 23 discussions in which a participant who had given a wrong answer in the pre-test faced group members who all agreed on the correct answer. In these discussions, the participant with a wrong answer never convinced the group. Thus, a participant who defended an answer on her own was significantly more likely to convince the group if she defended the correct answer than a wrong answer ($p < .001$). As an illustration, Fig. 2 represents the effects of correct and incorrect arguments on the outcome of group discussion for the task 4 of study 2.

If the participants who adopted the correct answer did so because they had understood the good argument, they might be able to use this new understanding to solve a transfer task. In study 1, we used a novel conservation task (task 4) after the post-test. In study 2, task 4 was used in the pre-test. In study 1, participants solved this task correctly more often (63%) than in study 2 (39%; $p = .006$).

3. Conclusion

The results of these studies show that preliterate individuals share some of the most important features of reasoning previously documented in WEIRD populations: reasoning performance significantly improves with discussion; argument production is marked by the myside bias; argument evaluation is efficient, enabling participants to discriminate between good and bad arguments and to be convinced by good arguments.

These results do not demonstrate that these features are universal. We might have considered the wrong traits of WEIRD cultures as being potentially relevant to the shaping of reasoning and argumentation. Other cultural traits, which are shared by WEIRD cultures and the Mayan indigenous group studied here but absent in other cultures, might in fact be determinant. Still, the K'iche' differ from WEIRD cultures in many of the traits that could have plausibly explained why the relevant features of reasoning are observed in WEIRD cultures: the value placed on argumentation, the fact that children and adults exchange many arguments, the fact that people feel they have to justify

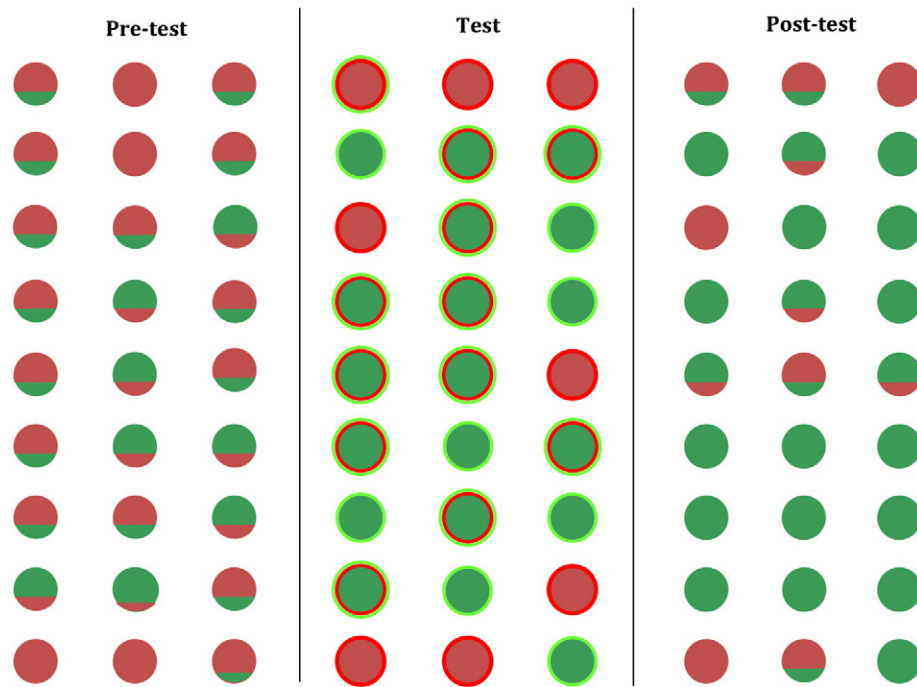


Fig. 2. Details of the performance for task 4 in study 2. For the pre-test and the post-test, each disk represents a nominal group, and the proportion of red (respectively green) in the disks represents the proportion of incorrect (respectively correct) answers. For the test, each disk represents a group and the color represents the outcome of the discussion. A red (respectively green) surrounding circle indicates that an argument for the wrong answer (respectively the correct argument for the correct answer) was uttered at least once during the discussion.

many of their choices, and the relatively low costs of poor decisions. We thus suggest that the present investigation helps make the case that these features are not culturally acquired.

The potential universality of these features of reasoning is relevant to the debates about the function of reasoning. Individualist theories of reasoning do not necessarily predict poor argumentative performance, but it is not clear how they can explain the improvement in performance following group discussion. As for the myside bias, it runs directly against their predictions. By contrast, these features of reasoning support the argumentative theory of reasoning.

The argumentative theory of reasoning suggests that these features are adaptive features of reasoning. They would be adaptive because they enable reasoning to fulfill its main function, argumentation, which in turns greatly facilitates communication (the evolutionary rationale is exposed in Mercier & Sperber, 2011). As expected if these features are adaptive, not only are they extremely robust in adults from WEIRD populations (Mercier, in press; Mercier & Sperber, 2011), but they have now been observed in a culture that differs in many relevant ways from WEIRD cultures, and there is no strong evidence that they are absent from other cultures (Mercier, 2011a; Mercier et al., 2015). Moreover these features seem to develop early (Castelain, Bernard, Van der Henst, & Mercier, in press; Mercier, 2011b).

Although research into cultural variation can yield precious insights about cognition, it is possible that cross-cultural psychology is overly geared towards differences rather than similarities. Focusing on potential universals is also of tremendous theoretical importance, in particular for testing adaptive hypotheses. We hope the present research can be another step in this direction (for another recent example using the same population, see Fontanari, Gonzalez, Vallortigara, & Girotto, 2014).

4. Data availability

All the data are available at this url: <https://sites.google.com/site/hugomercier/Data%20Evidence%20for%20core%20features%20of%20reasoning.zip?attredirects=0>

Acknowledgments

This research was carried out thanks to an Ambizione grant (PZ00P1_142388/1) from the Swiss National Science Foundation to H.M. We thank Jean Baratgin, Laura Fontanari, all the participants of the study, the CONALFA association and their teachers from the Sololá department, Francisca Ixquiactap Lopez for the logistic, and María Guarchaj and Catalina Sac Tambriz, the two experimenters who conducted the studies.

References

- Baumeister, R. F., & Masicampo, E. J. (2010). Conscious thought is for facilitating social and cultural interactions: How mental simulations serve the animal–culture interface. *Psychological Review*, 117(3), 945–971.
- 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.
- Billig, M. (1996). *Arguing and thinking: A rhetorical approach to social psychology*. Cambridge: Cambridge University Press.
- Castelain, T., Bernard, S., Van der Henst, J. B., & Mercier, H. (in press). The influence of power and reason on young Maya children's endorsement of testimony. *Developmental Science*.
- Chouinard, M. M., Harris, P. L., & Maratsos, M. P. (2007). Children's questions: A mechanism for cognitive development. *Monographs of the Society for Research in Child Development*, 1–129.
- Cole, M., Gay, J., Glick, J. A., & Sharp, D. W. (1971). *The cultural context of learning and thinking*. New York: Basic Books.
- Dasen, P. R. (1972). Cross-cultural Piagetian research: A summary. *Journal of Cross-Cultural Psychology*, 3(1), 23–40.
- Dehaene, S. (1999). *The number sense: How the mind creates mathematics*. Oxford: Oxford University Press.
- Dehaene, S. (2009). *Reading in the brain: The new science of how we read*. London: Penguin.
- Descartes, R. (1637). *Discours de la méthode*. Paris: Vrin.
- Diamond, J. (2012). *The world until yesterday: What can we learn from traditional societies?* London: Penguin.
- Doise, W., & Mugny, G. (1984). *The social development of the intellect*. Oxford: Pergamon Press.
- Fontanari, L., Gonzalez, M., Vallortigara, G., & Girotto, V. (2014). Probabilistic cognition in two indigenous Mayan groups. *Proceedings of the National Academy of Sciences*, 111(48), 17075–17080.
- Gauvain, M., Munroe, R. L., & Beebe, H. (2013). Children's questions in cross-cultural perspective a four-culture study. *Journal of Cross-Cultural Psychology*, 44(7), 1148–1165.
- Gibbard, A. (1990). *Wise Choices, Apt Feelings*. Cambridge: Cambridge University Press.

- Gluckman, M. (1967). *The judicial process among the Barotse of Northern Rhodesia (Zambia)*. Manchester: Manchester University Press.
- Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, 306(5695), 496–499.
- Gurven, M., & Kaplan, H. (2007). Longevity among hunter-gatherers: A cross-cultural examination. *Population and Development Review*, 33(2), 321–365.
- Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A bayesian approach to reasoning fallacies. *Psychological Review*, 114(3), 704–732.
- Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: P H Brookes.
- Heath, S. B. (1983). *Ways with words: Language, life and work in communities and classrooms*. Cambridge: Cambridge University Press.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world. *Behavioral and Brain Sciences*, 33(2–3), 61–83.
- INE (2011). *Caracterización República de Guatemala 2008/2011*. Guatemala: Instituto Nacional de Estadística Guatemala (Retrieved from <http://www.ine.gob.gt/sistema/uploads/2014/02/26/L5pNHMXzy5FFWmk9NHCrK9x7E5Qqvvy.pdf>).
- Jamet, F., Baratgin, J., & Filatova, D. (2014). Global warming and the rise of the sea level: A study of intellectual development in preadolescents and adolescents from 11 to 15 years old. *Studia pedagogiczne*, 24, 361–380.
- Janis, I. L. (1982). *Groupthink (2nd Rev)*. Boston: Houghton Mifflin.
- Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, 2(10), 732–735.
- Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9), 697–720.
- 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., Lichtenstein, S., & Fischhoff, B. (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory and Cognition*, 6(2), 107–118.
- Laughlin, P. R. (2011). *Group problem solving*. Princeton: Princeton University Press.
- Lévi-Strauss, C. (1966). The scope of anthropology. *Current Anthropology*, 7(2), 112–123.
- Lloyd, G. E. R. (1996). *Adversaries and authorities*. Cambridge: Cambridge University Press.
- Luria, A. R. (1976). *Cognitive development its cultural and social foundations*. Cambridge, MA: Harvard University Press.
- Medin, D. L., & Atran, S. (2004). The native mind: Biological categorization and reasoning in development and across cultures. *Psychological Review*, 111(4), 960–982.
- Mellers, B., Ungar, L., Baron, J., Ramos, J., Gurey, B., Fincher, K., ... Tetlock, P. (2014). Psychological strategies for winning a geopolitical forecasting tournament. *Psychological Science*, 25(5), 1106–1115.
- Mercier, H. (2011a). On the universality of argumentative reasoning. *Journal of Cognition and Culture*, 11, 85–113.
- Mercier, H. (2011b). Reasoning serves argumentation in children. *Cognitive Development*, 26(3), 177–191.
- Mercier, H. (2011c). When experts argue: Explaining the best and the worst of reasoning. *Argumentation*, 25(3), 313–327.
- Mercier, H. (in pressa). Confirmation (or myside) bias. In R. Pohl (Ed.), *Cognitive Illusions* (2nd Ed.). London: Psychology Press.
- Mercier, H. (in pressb). Reasoning and argumentation. In V. A. Thompson & L. J. Ball (Eds.), *International Handbook of Thinking & Reasoning*. Hove: Psychology Press.
- Mercier, H., Deguchi, M., Van der Henst, J. -B., & Yama, H. (2015). The benefits of argumentation are cross-culturally robust: The case of Japan. *Thinking & Reasoning*, 1–15 (ahead-of-print).
- Mercier, H., & Sperber, D. (2011). Why do humans reason? Arguments for an argumentative theory. *Behavioral and Brain Sciences*, 34(2), 57–74.
- Moshman, D., & Geil, M. (1998). Collaborative reasoning: Evidence for collective rationality. *Thinking and Reasoning*, 4(3), 231–248.
- Nakamura, H. (1964). *Ways of thinking of Eastern peoples: India, China, Tibet, Japan*. Hawaii: University of Hawaii Press.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomena in many guises. *Review of General Psychology*, 2, 175–220.
- Oetzel, J., Ting-Toomey, S., Masumoto, T., Yokochi, Y., Pan, X., Takai, J., & Wilcox, R. (2001). Face and facework in conflict: A cross-cultural comparison of China, Germany, Japan, and the United States. *Communication Monographs*, 68(3), 235–258.
- Peng, K., & Nisbett, R. E. (1999). Culture, dialectics and reasoning about contradiction. *American Psychologist*, 54(9), 741–754.
- Perret-Clermont, A. -N. (1980). *Social interaction and cognitive development in children*. London: Academic Press.
- Petty, R. E., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. In D. T. Gilbert, S. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 323–390). Boston: McGraw-Hill.
- Piaget, J. (1928). *Judgment and reasoning in the child*. London: Routledge and Kegan Paul.
- Piaget, J., & Inhelder, B. (1974). *The child's construction of quantities*. London: Routledge.
- Pye, C. (1986). Quiché Mayan speech to children. *Journal of Child Language*, 13(1), 85–100.
- Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: Issues and analysis. *International Journal of Forecasting*, 15(4), 353–375.
- Schwartz, B. (2004). *The paradox of choice: Why more is less*. New York: Harper Collins.
- 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: Cambridge University Press.
- Sears, D. O., Lau, R. R., Tyler, T. R., & Allen, H. M. (1980). Self-interest vs. symbolic politics in policy attitudes and presidential voting. *American Political Science Review*, 74(03), 670–684.
- Slavin, R. E. (1995). *Cooperative learning: Theory, research, and practice*. Vol. 2nd, London: Allyn and Bacon.
- Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., & Su, T. T. (2009). Why peer discussion improves student performance on in-class concept questions. *Science*, 323(5910), 122–124.
- Stanovich, K. E. (2004). *The robot's rebellion*. Chicago: Chicago University Press.
- Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3), 755–769.
- Tesser, A., & Conlee, M. C. (1975). Some effects of time and thought on attitude polarization. *Journal of Personality and Social Psychology*, 31(2), 262–270.
- Thompson, D. V., & Norton, M. I. (2008). The social utility of feature creep. In A. Lee, & D. Soman (Eds.), *Advances in Consumer Research* (pp. 181–184). Duluth, MN: Association for Consumer Research.
- Trouche, E., Sander, E., & Mercier, H. (2014). Arguments, more than confidence, explain the good performance of reasoning groups. *Journal of Experimental Psychology: General*, 143(5), 1958–1971.
- UNDP (2010a). *Informe Nacional de Desarrollo Humano 2009/2010: Cifras para el desarrollo humano Sololá*. Guatemala: UNDP (Retrieved from <http://desarrollohumano.org.gt/sites/default/files/07%20Fasciculo%20Solola.pdf>).
- UNDP (2010b). *Informe Nacional de Desarrollo Humano 2009/2010: Hacia un Estado para el desarrollo humano*. Guatemala: UNDP (Retrieved from http://hdr.undp.org/sites/default/files/guatemala_indh_2009-10.pdf).
- UNDP (2012). *Informe nacional de desarrollo humano 2011/2012: Guatemala, ¿un país de oportunidades para la juventud?* Guatemala: UNDP (Retrieved from http://desarrollohumano.org.gt/sites/default/files/INDH%202011_2012.pdf).
- UNDP (2014). *Human development report. Sustaining human progress: Reducing vulnerabilities and building resilience*. New York: UNDP.
- UNFPA (2014). *Análisis de la Situación de País 2010/2014*. Guatemala: UNFPA (Retrieved from https://www.unfpa.org/sites/default/files/portaldocument/Guatemala_UNDAF%202010-2014.pdf.pdf).