Memory Monitoring and Control in Japanese and German Preschoolers

Sunae Kim, Atsushi Senju, Beate Sodian, Markus Paulus, Shoji Itakura, Akiko Okuno, Mika Ueno & Joelle Proust

Abstract

Prior studies explored the early development of memory monitoring and control. However, little work has examined cross-cultural similarities and differences in metacognitive development in early childhood. In the present research, we investigated a total of 100 Japanese and German preschool-aged children's memory monitoring and control in a visual perception task. After seeing picture items, some of which were repeated, children were presented with picture pairs, one of which had been presented earlier and the other was a novel item. They then were asked to identify the previously presented picture. Children were also asked to evaluate their confidence about their selection, and to sort the responses to be used for being awarded with a prize at the end of the test. Both groups similarly expressed more confidence in the accurately remembered items than in the inaccurately remembered items, and their sorting decision was based on their subjective confidence. Japanese children's sorting more closely corresponded to memory accuracy than German children's sorting, however. These findings were further confirmed by a hierarchical Bayesian estimation of metacognitive efficiency. The present findings therefore suggest that early memory monitoring and control have both culturally similar and diverse aspects. The findings are discussed in light of broader sociocultural influences on metacognition.

Keywords: Metacognition; cross-cultural; preschoolers; memory monitoring; memory control

Memory Monitoring and Control in Japanese and German Preschoolers

Metacognition – the ability to monitor one's own mental states and processes, and guide one's behavior accordingly – is fundamentally involved in our everyday decisions. Memory-monitoring and control behaviors are crucial for adaptive human learning. It allows us to decide, for example, when to stop putting effort into learning, to employ learning strategies, or to seek help (e.g., Baars et al., 2020; Gönül et al., 2021). Moreover, some scholars argue that explicit metacognition in humans may have underpinned human cooperation and culture (e.g., Heyes et al., 2020; Shea et al., 2014).

Developmentally, metacognition is firmly present in elementary years (e.g., Koriat & Ackerman, 2010; Schneider & Lockl, 2008), and develops further into adolescence (e.g., Paulus et al., 2014; Schneider, 2008). In recent years, studies have begun to investigate younger children's nascent metacognitive abilities. Children as young as 3 years old display uncertainty monitoring and control behaviors. In a study by Balcomb and Gerken (2008), for example, 3.5-year-old children skipped trials in which they answered incorrectly later, suggesting that they implicitly monitor their own memory uncertainty. In addition, preschool-aged children were less confident on incorrectly identified picture items than on correctly identified pictures; when uncertain, they withheld their responses (Lyons & Ghetti, 2013) and sought for another person's help (Coughlin et al., 2015). Interestingly, 18-month-old infants asked parental help when they did not remember where a toy had been hidden (Goupil et al., 2016) (see also Geurten & Bastin, 2019; Goupil, & Kouider, 2016).

Moreover, Hembacher and Ghetti (2014) further investigated both memory monitoring and control in 3- to 5- year-old children. They found metacognitive monitoring, measured by their confidence judgments about their memories of picture items, was present in children as young as 4 years old. The presence of metacognitive control, measured by children's ability to sort the remembered items for later evaluation for a prize, was also observed in 4- and 5-year-olds and even among some 3-year-olds.

The aforementioned studies, however, have explored the developmental pattern of metacognition by exclusively focusing on European and North American children. This is unfortunate because an answer to whether metacognition is culturally modulated will greatly improve our understanding of the nature and characteristics of human metacognition. More generally, studies of different ethnic and cultural populations will establish the generalizability of the findings to a wider human population (Henrich et al., 2010; Nielsen et al., 2017). Granting that memory monitoring and control are involved in everyday judgments of learning and associated decisions, which themselves are critical for survival, one might expect that memory monitoring and control develop similarly across different cultures from early childhood. In a related metacognitive domain, Kim et al. (2020) recently reported no significant difference in assessing their own knowledge states between Japanese and German 4-year-olds – both of which seem different from Yucatec Mayan children (Kim et al., in press). In these studies, children's knowledge states of the hidden contents of a container were manipulated across conditions and children were asked to verbally report whether they knew the hidden contents. The same children in a separate task were also asked to either agree or decline informing an ignorant person about the hidden contents. In a critical partial knowledge condition, children were shown two toys and were told that one of them would be hidden in the box; then one of the toys was hidden in the box unbeknownst to them. In this condition, children tended to decline informing whereas they tended to respond that they knew what was in the box – with no significant performance difference between Japanese and German children. The informing task did not facilitate Yucatec Mayan children's performance, however. Further studies should examine the extent to which the informing task is meaningful and congruent with participants' cultural ways in matters such as communication and child-adult interactions (Kim et al., in press). As far as we know, these are the only studies that examined metacognition in young children growing up in non-"WEIRD" cultures (Western, educated, industrialized, rich, and democratic) (Henrich et al., 2010), and thus our current understanding is limited.

On the other hand, we also have reason to expect that cultural diversity of metacognition might appear early in human development. Importantly, metacognition can be shaped through cultural learning: metacognitive judgments, discrimination and calibration of metacognitive feelings could be directly and indirectly influenced by social feedback and instruction (see Heyes et al., 2020; see also Loussouarn et al., 2011). First of all, learning and teaching are culturally varied (Kline, 2015). For example, Japanese classroom teaching draws attention to individual errors whereas teaching in the USA or Germany lacks – and even avoids –discussing the errors made by individual students (e.g., Tulis, 2013, Erikson et al., 2020). Furthermore, Japanese teachers prefer to anticipate children's needs, considering that children should learn to depend on their teacher. In contrast, U.S. teachers prefer to respond to students' explicit expressions of need, considering that children should learn to depend on themselves (Rothbaum et al., 2006). These culturally modulated instructions and teaching may already start in early childhood. Studies provide converging evidence that parental sensitivity in parent-child interaction is understood differently by culture (e.g., Bornstein et al., 1992; Keller et al., 2002). Particularly, Japanese mothers, whose cultural norms emphasize interdependence over autonomy, are more sensitive to their child's needs than German mothers (Friedlmeier & Trommsdorff, 1999; Rothbaum et al., 2000), which typically facilitates young children's emotional regulation (e.g., von Suchodoletz et al., 2011). Japanese mothers display child-oriented feelings (e.g., empathy toward the child) whereas German mothers display self-oriented feelings (e.g., anger) when facing conflicts with their children – but Japanese mothers are more likely to intervene in peer conflicts and guide their children to empathize with peers. As a result, Japanese children are more likely to readily accept mothers' demands (e.g., Trommsdorff & Kornadt, 2003; Trommsdorff & Friedlmeier, 1993). Such maternal sensitivity to emotional needs may also occur in the form of informational needs (e.g., offering relevant feedback). Thus, Japanese children as compared to German children are more likely to select teaching ignorant over knowledgeable persons, while German children as compared to Japanese counterparts are more likely to select knowledgeable over ignorant persons to learn from (Kim et al., 2018). These differences might reflect differences in social and cultural attunement to others' knowledge states depending on the role of a learner or a teacher.

One might further reason that, in general, accurate memory monitoring and control could be important to achieve group decision-making and cooperation (Shea et al., 2014). As an interdependent view of the self is more likely to enhance group cooperation than an independent view of the self (e.g., Utz, 2004), the Japanese culture (an interdependent culture) in comparison to their Western independent counterparts (such as Germany) may emphasize an earlier motivation in accurately reporting what one knows or remembers; Japanese parents might also offer more frequent and appropriate feedback related to error monitoring and control to their children, which should help them to calibrate and accurately report their uncertainty. Specifically, parental feedback may allow children not only to distinguish metacognitive from other non-metacognitive feelings but also to distinguish among different metacognitive feelings. It may also involve teaching children an appropriate translation of subjective feelings into conventional/linguistic (both verbal and nonverbal) expressions – and correspondence between different levels of metacognitive feelings (e.g., uncertainty) and linguistic expression.

In the present research, we investigated preschool-aged Japanese and German children's memory monitoring and control. Our reason for comparing these populations was that we already know– as discussed above – that parenting and teaching styles and early child-parent interactions are guided by different cultural norms and values in Japan and Germany, which might further modulate memory monitoring and control in young children. Notably, the two cultural groups are comparable in other respects: they belong to highly industrialized countries, where the nuclear families typically include few children, where formal education occurs early on, etc. We also did not expect differences between the two groups of children concerning our experimental approach. Therefore, differences in metacognitive monitoring and control – if obtained – are unlikely to be attributable to differences in the latter features. Because we intended to test both metacognitive monitoring and control, the design of Hembacher and Ghetti (2014) was suitable for our purposes as it examined both metacognitive monitoring and control during preschool years.

Method

Participants. A total of 100 children were tested: 3.5- to 5-year-old Japanese children (N = 54, 28 girls, 26 boys, Mean age = 4.81, range = 3.69 ~ 5.89) and German children (N = 46, 27 girls, 19 boys, Mean age = 4.80, range = 3.71 ~ 5.04) participated. The two groups did not differ in Age t (98) = .049, p = .961. The data from two additional children were excluded for the analyses: one German child due to a technical problem, and one additional German child due to a failure to complete the task (after 8 trials). The sample size was determined based on Hembacher and Ghetti (2014): The total sample size of 98 was considered sufficient to detect the effect size of .25 (Cohen's d) at the power of 80%. Children from both Germany and Japan came from middle or upper middle class families. Fifty-six precent of Japanese mothers and 67% of Japanese fathers had university or graduate degrees. Fifty-nine precent of German mothers and 56% of German fathers had university or graduate degrees. Sixty-seven percent of Japanese mothers and all Japanese fathers were working; all German mothers and fathers were working – a majority of them had white-collar jobs (74% of

Japanese parents and 68% of German parents). The present research was approved by the ethic committee at Ludwig Maximilian University - Munich in Germany and Kyoto University in Japan and was conducted according to the 1964 Declaration of Helsinki. Written informed consent was obtained from the parents.

Design and procedure. We closely followed Hembacher and Ghetti (2014) except for several changes as described below. The original research study included young 3-year-old American children. We tested older rather than younger 3-year-olds for the following reason. During our pilot test, an entire session took about 30 minutes; maintaining their attention throughout testing turned out to be extremely challenging for younger 3-year-old children. Individual children were tested by a female experimenter in a lab or in a separate room in their kindergarten. Children were presented with 20 picture stimuli one at a time for 2 seconds on a computer screen. During our pilot testing, we confirmed that both Japanese and German children in the age range of 3.5-5 were familiar with the picture items. The order of the presentation of the picture stimuli was fixed and maintained across participants. Unlike the original study in which children were tested in two separate sessions each involving 20 trials, and each session a week apart from one another, we tested children in one session of 20 trials. Importantly, as in the original study, half of the picture items were presented once, and the other, twice. Children were asked to touch the screen as soon as the picture item appeared on the screen to ensure that they paid attention ("I'm going to show you some pictures and please touch the picture as soon as it appears on the screen. Okay?"). After this encoding phase, children received successively a retrieval task, a confidence judgment task, and a sorting task. In the retrieval task, children were presented with the picture pairs, one of which was previously presented during the encoding phase and, the other, a novel item and were asked to choose the picture that they saw before ("Which one did you see before? Can you point at the picture you saw before?"). The novel-familiar item pairs were predetermined and their order was then randomized and fixed across trials . Upon their choice, the non-chosen picture item disappeared from the screen and the confidence scale appeared below the chosen item. Then the confidence judgment task followed. In this task, children were asked about their confidence concerning their picture choice in the retrieval task, using a 3-point scale of, “not so sure”, “kind of sure”, and “really sure” (“How sure are you? Can you point at one of the circles?”). Unlike Hembacher and Ghetti (2014) who used a confidence scale of corresponding face pictures, we used a 3-point scale of circles with increasing size and intensifying colors – because we reasoned that facial expressions of uncertainty might be culturally specific. Like in Hembacher and Ghetti (2014), we gave the same training trials with feedback concerning the usage of the scale. The experimenter explained the meaning of each circle corresponding to “not so sure”, “kind of sure”, or “really sure” and feedback was given for the correspondence between children’s behaviors and confidence judgments during the training (“Okay, there are three circles. If you are not really sure about your answer, whether you saw the picture, then you should select this circle (pale red); if you are sort of sure, you should select this circle (middle red); if you are really sure, then you should select this circle (bold red)?”). Children were then asked, “How sure are you? Can you point at one of the circles?” followed by “Okay, can you tell me, how sure you are?” If children’s verbal response matched pointing at the circle then, the experimenter said, “Right. This circle means that you are really sure/kind of sure/really not sure. So it is right that you pointed at this circle.” If not matched: “No, look, this circle means, that you are really sure/kind of sure/really not sure. You said that you are really sure/kind of sure/really not sure, so you should point at this circle.” Finally, after children made their confidence judgment, they received the sorting task in order to assess children’s behavioral control based on their subjective confidence. In this task, children were asked to choose either a box with a smiley face with open eyes or another box with a smiley face with closed eyes (“To which smiley face would you give your picture?"). They were explained that only those picture items they chose to put in the box with an open eye smiley face would be evaluated for a later prize ("Okay, during the game, you can give the pictures to either this smiley face with open eyes (pointing) or with this smiley face with closed eyes (pointing). If you think you did a good job and if you want me to look at your answer, then you can give the pictures to this smiley face (pointing) with open eyes. If you think you make a mistake and don't want me to look at them, then you can give the pictures to this smiley face with closed eyes (pointing). I will look at only those pictures that you give to this one (pointing at a smiley with open eyes) and you will get a prize if you have done a good job on only those pictures with this smiley face (pointing at the one with open eyes). Sounds good?"). Once they chose one of the boxes, they would see the lid of the box of their choice opened with the picture of their choice next to the box. Again, as in the original study, we administered the training trials with feedback ("Okay, do you want me to look at the picture later because you think you've done a good job or do you not want me to look at the picture because you think you might make a mistake?" If children answered correctly, then they were told: "Yes. The smiley with the open/closed eyes means that I will look/won't look at the picture later." If they answered incorrectly then: "No. The smiley with the open/closed eyes means that I will look/won't look at the picture later. If you want me to look/ not look at the picture later, you should give your picture to the smiley with the open/closed eyes."). See Figure 1 for the pictorial layout of the retrieval, confidence judgment, and sorting tasks.

Data analyses. We followed the analyses by Hembacher and Ghetti (2014) except for the following changes. First, we included Country as a factor. Unlike Hembacher and Ghetti (2014) we included all children – for example, those who did not have inaccurate responses; those with chance memory performance – which resulted in missing data in some cells. In order to handle the missing data as well as repeated measures, we used Linear Mixed models.

Unlike Hembacher and Ghetti (2014), we treated Age as a continuous variable rather than a categorical variable. Finally, we adopted a hierarchical Bayesian estimation of efficiency, called HMeta-d (Fleming, 2017) and compared metacognitive efficiency between our two cultural groups. Meta-d' (e.g., Maniscalco & Lau, 2012), as an index of metacognitive sensitivity, measures how well an individual's confidence judgment discriminates between accurate and inaccurate responses and controls for type 1 sensitivity (or d': how well an individual categorizes stimuli), type 1 response bias (tendency to select one stimulus classification more often than the other), or type 2 response bias (tendency to give high or low confidence estimates). In particular, by comparing directly to d' (as it is expressed in the same units), meta-d' provides a measure of an individual's metacognitive efficiency: it is ideal if meta-d'/d' (Mratio or metacognitive efficiency) = 1. Estimated log (Mratio), based on signal detection theory, provides an index of metacognitive efficiency and HMeta-d provides group-level estimates of metacognitive efficiency (Fleming, 2017). Comparing estimates of group level metacognitive efficiency , the 95% highest-density interval (HDI) of difference of the group parameters (computed from the posterior samples) allowed us to test whether one group had a higher metacognitive efficiency than the other. HMeta-d controls for confounds such as memory performance or confidence bias and incorporates subject level uncertainty into the model; moreover, it performs well with small data sets (Fleming, 2017) as was the case in our study. We measured HMeta-d for the confidence-memory accuracy and for the sorting-memory accuracy.

Results

All data are included as a supplementary information file. The preliminary analyses found that there was no main effect of item order or gender, nor did they interact with any other main factors. Thus, we did not consider these factors for the rest of the analyses. In all reported analyses, we compared the models with and without Age – Age was dropped if the model including Age did not significantly improve the overall model fit.

Memory performance. We first asked whether our manipulation of memory strength during memory encoding was effective. Figure 2 presents the mean proportion of trials in which children accurately remembered items as a function of Encoding type and Country. Children remembered the repeated items (M = .90, SD = .16) more accurately than the non-repeated items (M = .82, SD = .17), F (1, 98) = 33.650, p < .001 but after controlling for Age (F (1, 96) = 12.994, p < .001) it was no longer significant F (1, 96) = 2.439, p = .122. The model including Age significantly improved the model fit (-2LL = -181.351 vs. -195.936). No other effects were significant including Country F (1, 96) = .330, p = .567 or Country X Encoding F (1, 96) = .283, p = .596.

Monitoring memory accuracy and strength. Importantly, we asked whether children's confidence judgments corresponded to their memory accuracy and strength. Figure 3 presents children's mean confidence scores (score 0- low confidence; score 1- medium confidence; score 2- high confidence) as a function of Country and Item types (Repeated-accurately remembered items vs. Non-repeated accurately remembered vs. Inaccurately remembered items). Because we reasoned that memory strength would not be differentiated between repeated and non-repeated conditions for the inaccurately remembered items we combined these conditions for the inaccurately remembered items. The model including Age (F (1, 92.535) = 4.728, p = .032) did not significantly improve the model fit (-2LL = 257.274 vs. 269.702). Item type was significant, F (2, 84.568) = 19.010, p < .001. Children rated the inaccurately remembered items (M = 1.19, SD = .74) as less confident than either the non-repeated-accurately remembered (M = 1.66, SD = .39) (t (75.697) = 5.987, p < .001) or the repeated-accurately remembered items (M = 1.70, SD = .37) (t (77.324) = 6.309, p < .001). Country was not significant F (1, 96.762) = .005 p = .945, nor was Country X Item type F (2, 84.568) = .277, p = .759.

Sorting decisions. An important question was whether children's confidence judgments corresponded to their sorting decisions. Figure 4 presents children's mean confidence scores as a function of Country and Sorting decision. The model including Age (F (1, 73.55) = 4.638, p = .035) did not significantly improve the model fit (-2LL = 246.056 vs. 246.743). Children's confidence was higher for the items they chose for later evaluation (open-eyes box) (M = 1.71, SD = .39) than for the items they did not choose for later evaluation (closed-eyes box) (M = .94, SD = .71), F (1, 74.601) = 63.072, p < .001. Sorting X Country was not significant F (1, 74.601) = .027, p = .870 as well as Country F (1, 73.316) = .359, p < .551.

Next, we asked whether children's sorting decision based on their confidence level as demonstrated by the previous analysis corresponded to their memory accuracy. Figure 5 presents children's mean memory scores as a function of Country and Sorting. The model including Age (F (1, 49.385) = 5.310, p = .025) did not significantly improve the model fit (-2LL = -18.622 vs. -20.397). There was a significant interaction of Sorting X Country F (1, 96.108) = 8.260, p = .005 as well as Country F (1, 103.280) = 9.491, p = .003 and Sorting F (1, 96.108) = 30.523, p < .001. Japanese children sorted accurately remembered items more frequently into the open-eye box (M = .88, SD = .16) than into the closed-eye box (M = .60, SD = .33) F (1, 96.967) = 37.803, p < .001 whereas German children sorted accurately remembered items equally into the two boxes (open-eye box: M = .89, SD = .15; closed-eye box: M = .81, SD = .24) F (1, 95.362) = 3.293, p = .073. German children sorted accurately remembered items more frequently into the closed-eye box than Japanese children F (1, 161.946) = 14.785, p < .001 whereas there was no group difference with respect to the open-eye box F (1, 159.822) = .198, p = .657.

HMeta- d. We found no group difference for children's memory accuracy-confidence as HDI of difference of posterior group estimates overlapped with 0: [-1.866 ~ .569]. By contrast, there was a group difference with respect to children's memory accuracy-sorting: [.000123 ~ 1.4264]. See Figure 6. Note that there was no group difference in terms of memory performance $t(98) = -1.63$, $p = .106$, confidence level $t(98) = -.283$, $p = .778$ or sorting level $t(98) = .228$, $p = .820$. [1]

Discussion

In the present research, we investigated 3.5 ~ 5-year-old Japanese and German children's memory monitoring and control. As noted in the Introduction, the empirical data are scarce concerning metacognitive abilities in young children growing up in non-WEIRD cultures. To our knowledge, the present research provides initial evidence both for cross-cultural similarity and diversity in young children's memory monitoring and control. Specifically, we found, first, that children rated accurate memory responses more confidently than inaccurate responses. Importantly, this was similar for both cultural groups. Second, both cultural groups similarly sorted out for prize attribution the items they felt more confident about. However, Japanese children's sorting decisions corresponded more closely to their memory accuracy than German children's decisions. Below we discuss these findings in more detail.

Memory monitoring. Both Japanese and German children rated their accurate memory responses more confidently than their inaccurate responses. This finding is consistent with prior research showing that young children are able to monitor their own uncertainty (e.g., Coughlin et al., 2014; Goupil & Kouider, 2016, Lyons & Ghetti, 2013). *Memory control.* Both Japanese and German children tended to exclude their less confidently rated responses for later prize evaluation. However, Japanese children's sorting corresponded to their memory accuracy, whereas German children' sorting did not.

---

[1] Although our interest was a direct group level comparison of Mratio, we also provide values of Meta-d': For confidence-accuracy: Meta-d' = 1.67 for Japanese children, Meta-d' = 1.75 for German children; for sorting-accuracy: Meta-d' = 2 for Japanese children, Meta-d' = 1.47 for German children.

In summary, we found similarities in Japanese and German children's memory monitoring and control. Accurately remembered items were rated as more confident than inaccurately remembered items. Additionally, more confidently rated items were appropriately selected for later evaluation. Kim et al. (2020) also reported a similarity between Japanese and German children in their metacognitive ability of assessing their own knowledge. The findings of cultural similarity in the present study together with those of Kim et al. (2020) may indicate the universal importance of metacognitive monitoring and control in adaptive learning and behaviors, and thus in human and nonhuman survival. Indeed, basic memory monitoring and control are present even in preverbal infants (e.g., Goupil & Kouider, 2016).

In the Introduction, we discussed several related reasons for cultural differences of metacognitive monitoring and control appearing early in childhood. If different cultural practices surrounding teaching and instructions as well as parental sensitivity in parent-child interactions – and in relation to different notions of self (interdependent vs. independent) – contributed to metacognitive monitoring and control even in early childhood, then we would have observed more pronounced and consistent differences between the two cultural groups targeted in our study. On the contrary, the cultural difference emerged only with respect to the sorting decision: Japanese children's sorting for evaluation corresponded more closely to the accuracy of their responses than German children's. These findings were further confirmed by HMeta-d controlling for potential confounds (also note that both groups of children had a similar level of memory performance and of response biases (confidence or sorting level)).

How can we explain our finding of a cultural difference in the sorting decision? Cross-cultural studies demonstrate that internal motivation as well as behaviors and even individual well-being are modulated by a culturally different construal of the self – whether the self is understood independently of others (as in Germany) or interrelated to others (as in Japan). For example, when primed by public eyes (thus eliciting social evaluation) during an IQ test, Japanese American adults scored higher than Caucasian American adults, whereas a reversed pattern was found in the absence of priming conditions (Na & Kitayama, 2012). This suggests that Japanese American participants' internal motivation to perform and behave is more likely to be influenced by others' evaluation than participants from a European descent because the relation to others is a more salient dimension in selfhood in an interdependent than in an independent conception of selfhood. In this research, the sorting task required children to sort the items that might qualify for a reward. Thus, it is possible that even at this young age, children's performance is modulated by a cultural difference in self-construal related to internal motivation and public evaluation (see also Iyengar & Lepper, 1999). Japanese children might have a stronger motivation to perform well in general – accompanied by higher social expectation –compared to German children. Alternatively, German children may be more prone to risk aversion (over-estimating the risk of error when evaluating their chances of getting a prize) – and this might be because German children are under higher pressure to perform rather than Japanese children.

The cultural difference of a tighter correspondence between sorting and accuracy may also be explained by a difference in executive function (defined as a range of cognitive abilities whose function is to control behaviors and other cognitive processes in order to attain a certain goal). Roebers (2017) proposed that metacognition and executive function are guided by the same underlying cognitive self-regulatory processes: executive function might be causally involved in the development of metacognitive control especially in early years. In the sorting task, children were asked to sort the items for later prize winning consideration, which is likely to involve executive function. Studies report a higher executive function among children growing up in East Asian countries than among those in Western countries (e.g., Imada et al., 2013). Thus, to the extent that metacognitive control involves executive function (e.g., Koren et al., 2006) – whose development varies across cultures – we might explain our data by a cultural difference of metacognitive control. Future studies should replicate the present findings and address whether the cultural differences observed in the present study become increasingly heightened or diminished with increasing age. Nevertheless, the present study is the first step toward answering cross-cultural variations and similarities in young children's metacognitive development.

The above explanations are consistent with a dual-process theory of metacognition (e.g., Koriat & Levy-Sadot, 1999; Koriat & Ackerman, 2010; Proust, 2013). According to the theory, two kinds of inputs are used in metacognitive evaluations: 1) Experience-based evaluations are guided by automatic inferences based on mnemonic and heuristic cues gathered from immediate feedback from task performance. 2) Information-based evaluations are guided by analytic and deliberate inferences and theories. In the present study, the tasks seem to depend primarily on experience-based evaluations (i.e., subjective feelings of confidence). While it is not excluded that children also rely upon information-based evaluations (e.g., beliefs about the task), this form of metacognition has been shown to typically develop later in development – around school age) (e.g., Schneider & Lockl, 2008). Given the observed cultural difference specific to the sorting task, therefore, it is plausible that metacognitive *control* is modulated by cultural factors such as social values and overall learning goals. Moreover, the difference between a prospective and a retrospective form of evaluation (respectively used in sorting one's responses and expressing a confidence judgement) might be itself sensitive to cultural influences. Indeed, different brain regions are involved in prospective judgments (e.g., ease of learning or feeling of knowing) and in retrospective monitoring judgments (confidence monitoring) (Fleming & Dolan, 2012). Although the sorting task may not exclusively exploit a predictive source of information different from confidence (because decision was likely to be made partly on the basis of retrospective judgments of memory and/or confidence), further research still might scrutinize the existence of a potential contrast between future-oriented and past-oriented metacognitive evaluations in young children – whether they are subject to different external influences. Finally, the absence of cultural difference in confidence monitoring could be due to small effects in the present study.

What cannot be denied is that accuracy in metacognitive monitoring and control is beneficial not only at the individual level (e.g., Dunlosky & Rawson, 2012; Dunlosky et al., 2013; Sodian & Frith, 2008) but also at the group level (Heyes et al. 2020; Shea et al., 2014). Societies and cultures where individuals reliably report to others their own confidence and subsequently use it in their own decisions and behaviors are likely to be more successful and more efficient in their cooperative endeavors (e.g., Bahrami et al., 2010). While metacognitive monitoring and control underpin the emergence and evolution of every culture, and are involved in all kinds of learning and decision-making, the ways in which individual metacognitive monitoring and control are scaffolded by social structures and practices early on in development promotes sensitivity to specific metacognitive cues and to associated epistemic norms (Heyes et al., 2020; Proust & Fortier, 2018).

Although we found culturally variable as well as similar aspects of metacognition in a perceptual memory task, future studies should address whether this finding generalizes to other types of tasks. There is no reason to expect that the development of metacognitive abilities should be aligned across cognitive tasks in a given culture. In fact, even among the same cultural population, developmental studies document different ages for the emergence of the various metacognitive abilities. For example, in a memory task, verbal assessment of confidence in one's own memory is observed in children as young as 4 years old (Hembacher & Ghetti, 2014), whereas in a perceptual discrimination task, 3.5 -year-olds are able to opt out from providing responses when they are uncertain (Bernard et al., 2015; Lyons & Ghetti, 2013).

Likewise, there is no a priori reason to expect that these cultural differences and similarities would generalize to other populations. In adult studies, for example, reasoning and attention styles of Central and East Europeans are more similar to those of East Asians than West Europeans (Varnum et al., 2008); similarly, South Italians are more similar to East Asians than North Italians (Knight & Nisbett, 2007). Developmentally, we may also find a different pattern of metacognitive abilities in other ethnic populations (e.g., Kim et al., in press). Depending on the social and cultural emphasis on specific epistemic practices and attentional guidance, metacognitive abilities might sensibly differ across cultures. Engaging in investigations of diverse human populations, therefore, will promote a better understanding of the social dimensions of human cognition as well as of human metacognition.

Acknowledgments. We thank all the families and children who participated in the present research. The research was supported by an ERC advanced grant (#269616) and ANR-17-EURE-0017 FrontCog to JP and Japanese Ministry of Education, Culture, Sports, Science and Technology grant (#16H06301) to SI. Part of the data were presented in Budapest CEU Conference on Cognitive Development (2017), Budapest, Hungary.

The data are available as supplementary material, and the present research was not preregistered.

References

Baars, M., Wijnia, L., de Bruin, A., & Paas, F. (2020). The relation between students' effort and monitoring judgments during learning: A meta-analysis. *Educational Psychology Review, 32,* 979–1002. https://doi.org/10.1007/s10648-020-09569-3

Bahrami, B., Olsen, K., Latham, P. E., Roepstorff, A., Rees, G., & Frith, C. D. (2010) Optimally interacting minds. *Science, 329,* 1081–1085. doi: 10.1126/science.1185718

Balcomb, F. K., & Gerken, L. (2008). Three-year-old children can access their own memory to guide responses on a visual matching task. *Developmental Science, 11*, 750–760. doi: 10.1111/j.1467-7687.2008.00725.x

Bernard, S., Proust, J., & Clément, F. (2015). Procedural metacognition and false belief understanding in 3-to 5-year-old children. *PloS one*, *10*(10), e0141321.

Bornstein, M.H., Tamis-LeMonda, C.S., Tal, J., Ludemann P, Toda, S., Rahn, C.W., Pêcheux, M.G., Azuma, H., & Vardi, D. (1992). Maternal responsiveness to infants in three societies: the United States, France, and Japan. *Child Development, 63*, 808-821.

Coughlin, C., Hembacher, E., Lyons, K. E., & Ghetti, S. (2015). Introspection on uncertainty and judicious help-seeking during the preschool years. *Developmental Science, 18*, 957-971. doi: 10.1111/desc.12271

Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and Instruction, 22*, 271–280. doi: 10.1016/j.learninstruc.2011.08.003

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4-58. doi : 10.1177/1529100612453266

Eriksson, K., Lindvall, J., Helenius, O., & Ryve, A. (2020). Cultural Variation in the Effectiveness of Feedback on Students' Mistakes. *Frontiers in Psychology*, *10*, 3053- https://doi.org/10.3389/fpsyg.2019.03053

Fleming, S. M. (2017). HMeta-d: hierarchical Bayesian estimation of metacognitive efficiency from confidence ratings, *Neuroscience of Consciousness*, 1, nix007, https://doi.org/10.1093/nc/nix007

Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B Biological Science*, 367, 1338–49.

Friedlmeier, W., Trommsdorff, G. (1999). Emotion regulation in early childhood: A cross-cultural comparison between German and Japanese toddlers. *Journal of Cross-Cultural Psychology*. *30,* 684-711. doi:10.1177/0022022199030006002

Geurten, M., & Bastin, C. (2019). Behaviors speak louder than explicit reports: Implicit metacognition in 2.5-year-old children. *Developmental Science, 22*:e12742. doi: 10.1111/desc.12742

Gönül, G., Tsalas, N., & Paulus, M. (2021). The effect of time pressure on metacognitive control: developmental changes in self-regulation and efficiency during learning. *Metacognition and Learning,* https://doi.org/10.1007/s11409-021-09262-y

Goupil, L., & Kouider, S. (2016). Behavioral and neural indices of metacognitive sensitivity in preverbal infants, *Current Biology, 26*, 1-8. doi: 10.1016/j.cub.2016.09.004

Hembacher, E., & Ghetti, S. (2014). Don't look at my answer: Subjective uncertainty underlies preschoolers' exclusion of their least accurate memories. *Psychological Science, 25*, 1768-1776. doi: 10.1177/0956797614542273

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33(2-3*), 61–83. doi:10.1017/s0140525x0999152x

Heyes, C., & Bang, D., Shea, N., Frith, C. D. & Fleming, S. M. (2020). Knowing ourselves together: The cultural origins of metacognition. *Trends in Cognitive Sciences, 24*, 349-362. doi:10.1016/j.tics.2020.02.007

Imada, T., Carlson, S. M., & Itakura, S. (2013). East-West cultural differences in context-sensitivity are evident in early childhood. *Developmental Science,16*,198-208. doi: 10.1111/desc.12016

Iyengar, S. S., & Lepper, M. R. (1999). Rethinking the value of choice: A cultural perspective on intrinsic motivation. *Journal of Personality and Social Psychology, 76*(3), 349–366. doi; 10.1037/0022-3514.76.3.349

Keller, H., Yovsi, R. D., & Voelker, S. (2002). The role of motor stimulation in parental ethnotheories: the case of Cameroonian nso and German women. *Journal of Cross Cultural Psychology. 33*, 398–414. 10.1177/00222102033004003

Kim, S., LeGuen**,** O., Sodian, B., & Proust, J. (in press). Are children sensitive to what they know?: An insight from Yucatec Mayan children. *Journal of Cognition and Culture.*

Kim, S., Paulus, M., Sodian, B., Itakura, S., Ueno, M., Senju, A., & Prout, J. (2018). Selective learning and teaching among Japanese and German children. *Developmental Psychology. 54*, 536–542. doi: 10.1037/dev0000441

Kim, S., Sodian, B., Paulus, M., Senju, A., Okuno, A., Ueno, M., Itakura, S., & Proust, J. (2020). Metacognition and mindreading in young children: A cross-cultural study. *Consciousness and Cognition, 85*. doi: 10.1016/j.concog.2020.103017

Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences, 38*, E31. doi: 10.1017/S0140525X14000090

Knight, N., & Nisbett, R. (2007). Culture, class and cognition: Evidence from Italy. *Journal of Culture and Cognition,* 7, 283–291. doi: 10.1163/156853707X208512

Koren, D., Seidman, L. J., Goldsmith, M., & Harvey, P. D. (2006). Real-world cognitive—and metacognitive—dysfunction in schizophrenia: a new approach for measuring (and remediating) more "right stuff". *Schizophrenia Bulletin*, *32*, 310-326.

Koriat, A. (1997). Monitoring one's own knowledge during study: a cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, 349–370. doi: 10.1037/0096-3445.126.4.349

Koriat, A., & Ackerman, R. (2010). Choice latency as a cue for children's subjective confidence in the correctness of their answers. *Developmental Science. 13*, 441–453. doi: 10.1111/j.1467-7687.2009.00907.x

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), Dual process theories in social psychology (pp. 483–502). New York: Guilford Press.

Lyons, K. E., & Ghetti, S. (2013). I don't want to pick! Introspection on uncertainty supports early strategic behavior. *Child Development, 84*, 726-736. doi: 10.1111/cdev.12004

Loussouarn, A., Gabriel, D., & Proust, J. (2011). Exploring the informational sources of metaperception: The case of change blindness blindness. *Consciousness and Cognition*, *20*(4), 1489-1501.

Maniscalco, B. & Lau, H. (2012). A signal detection theoretic approach for estimating metacognitive sensitivity from confidence ratings. *Consciousness & Cognition. 21,* 422–430.

Na, J., & Kitayama, S. (2012). Will people work hard on a task they choose? Social-eyes priming in different cultural contexts. *Journal of Experimental Social Psychology, 48*(1), 284–290. doi:10.1016/j.jesp.2011.09.003

Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal Experimental Child Psychology. 162*, 31-38. doi:10.1016/j.jecp.2017.04.017.

Paulus, M., Tsalas, N., Proust, J., & Sodian, B. (2014). Metacognitive monitoring of oneself and others: Developmental changes in childhood and adolescence. *Journal of Experimental Child Psychology, 122,* 153-165. doi: 10.1016/j.jecp.2013.12.011

Proust, J. (2013).

Proust, J. & Fortier, M. (2018). Metacognitive Diversity: An Interdisciplinary Approach. Oxford University Press.

Roebers, C. M. (2017). Executive function and metacognition: Towards a unifying framework of cognitive self-regulation. *Developmental Review, 45*, 31–51. https://doi.org/10.1016/j.dr.2017.04.001

Rothbaum F., Nagaoka R., & Ponte I. C. (2006). Caregiver sensitivity in cultural context: Japanese and U.S. teachers' beliefs about anticipating and responding to children's needs. J. Res. Child. Educ. 21, 23–40 10.1080/02568540609594576

Rothbaum, F., Weisz, J., Pott, M, Miyake, K., & Morelli, G. (2000). Attachment and culture: Security in the United States and Japan. *American Psychologist, 55*, 1093-1104. doi:10.1037/0003-066X.55.10.1093

Schneider, W. (2008). The development of metacognitive knowledge in children and adolescents: Major trends and implications for education. *Mind, Brain, and Education, 2,* 114-121. doi: 10.1111/j.1751-228X.2008.00041.x

Schneider, W., & Lockl, K. (2008). Procedural metacognition in children: Evidence for developmental trends. In J. Dunlosky & R. A. Bjork (Eds.), A handbook of metamemory and memory. (pp. 391-410). New York: Psychology Press.

Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., & Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends in Cognitive Science, 18,*186-93. doi: 10.1016/j.tics.2014.01.006.

Sodian, B., & Frith, U. (2008). Metacognition, theory of mind, and self-control: The relevance of high-level cognitive processes in development, neuroscience, and education. *Mind, Brain, and Education, 2*, 111-113. doi:10.1111/j.1751-228X.2008.00040.x

Tulis, M., (2013). Error management behavior in classrooms: Teachers' responses to student mistakes. *Teaching and Teacher Education, 33*, 56-68. doi:10.1016/J.TATE.2013.02.003

Trommsdorff, G., & Friedlmeier, W. (1993). Control and responsiveness in Japanese and German mother-child interactions. *Early Development and Parenting, 2,* 65-78.

Trommsdorff, G. & Kornadt, H-J. (2003). Parent-child relations in cross-cultural perspective. *Handbook of dynamics in parent-child relations.* L. Kuczynski (ed.). London: Sage. 271-306.

Utz, S. (2004). Self-construal and cooperation: Is the interdependent self more cooperative than the independent self? *Self and Identity, 3*, 177–190. doi: 10.1080/13576500444000001

Varnum, M., Grossmann, I., Katunar, D., Nisbett, R., & Kitayama, S. (2008). Holism in a European cultural context: Differences in cognitive style between central and East Europeans and Westerners. *Journal of Cognition and Culture, 8,* 321-333. doi: 10.1163/156853708X358209

von Suchodoletz, A., Trommsdorff, G., & Heikamp, T. (2011). Linking maternal warmth and responsiveness to children's self-regulation. *Social Development, 20*, 486-503. doi: 10.1111/j.1467-9507.2010.00588.x

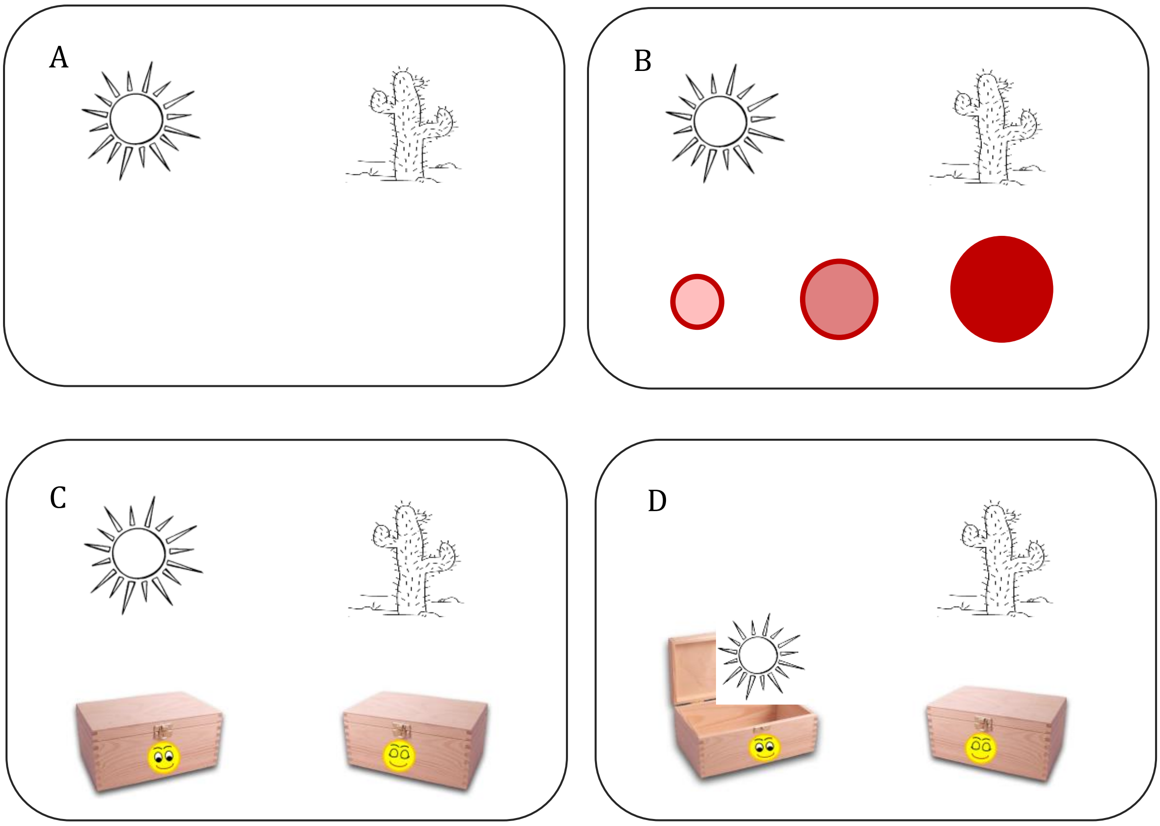

Figure 1. After the encoding phase, children received a retrieval task (A), a confidence judgment task (B), and a sorting task (C & D).

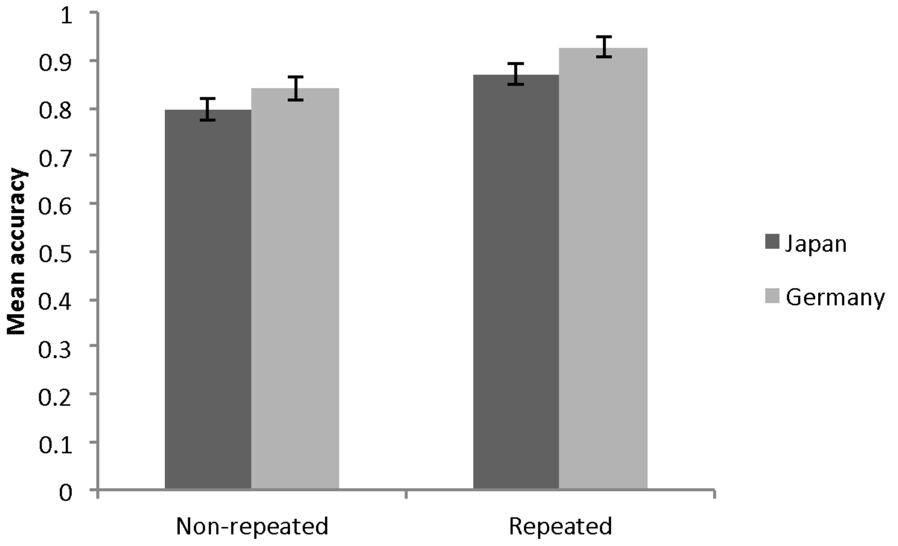

Figure 2. Mean proportion of trials in which children accurately remembered items as a function of Encoding type and Country. Error bars indicate standard errors.

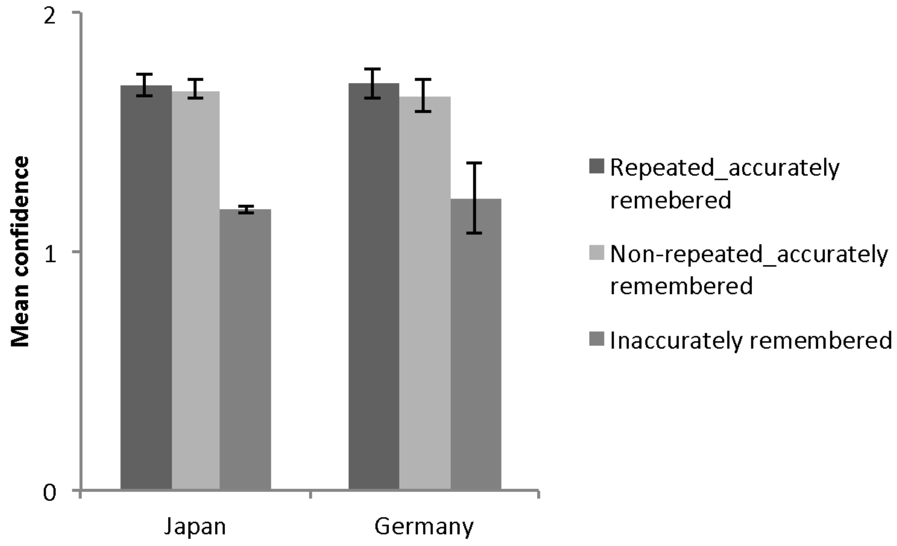

Figure 3. Mean confidence as a function of Item type and Country. Error bars indicate standard errors.

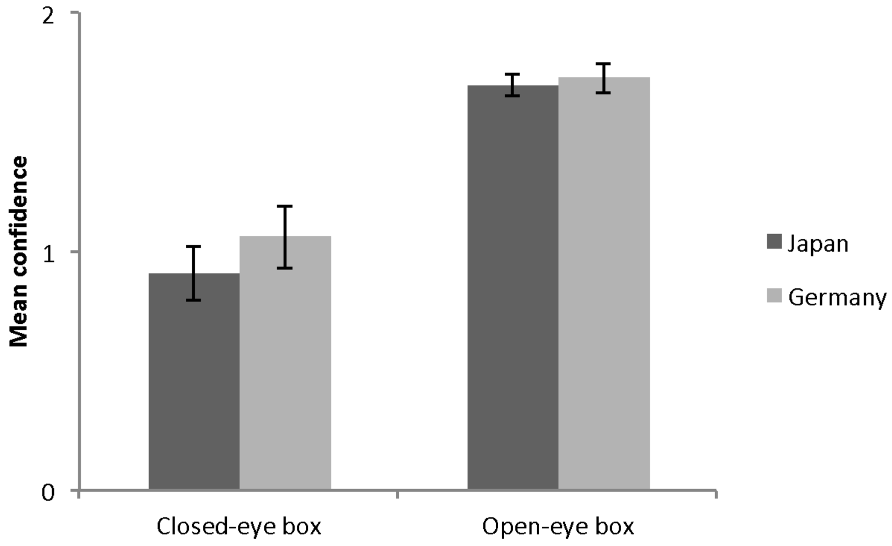

Figure 4. Mean confidence as a function of Sorting decision and Country. Error bars indicate standard errors.

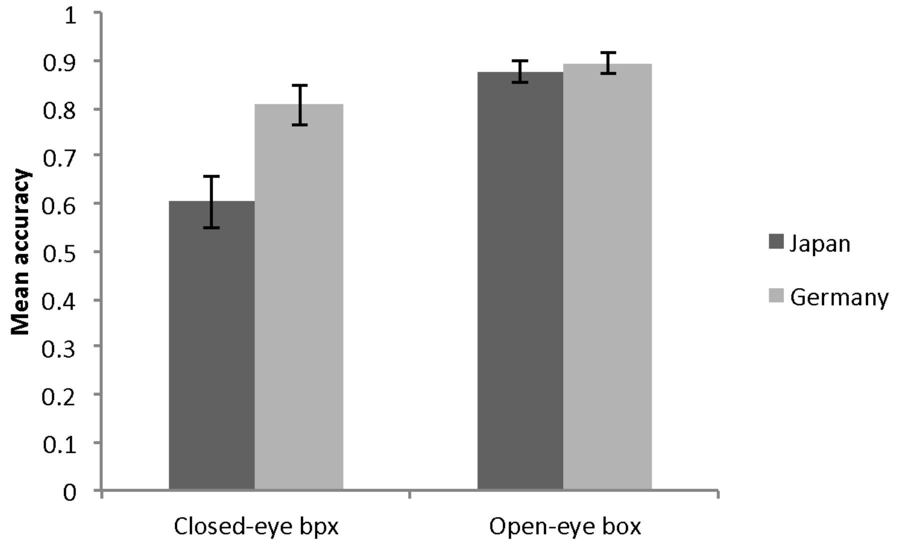

Figure 5. Mean accuracy as a function of Sorting decision and Country. Error bars indicate standard errors.

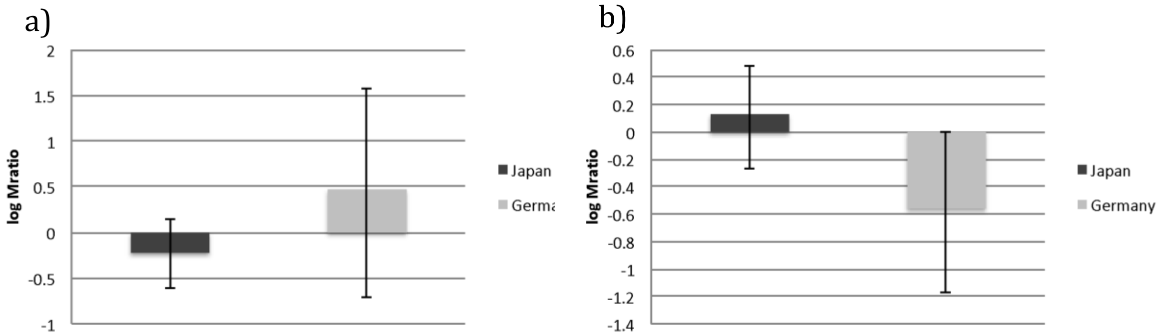

Figure 6. Metacognitive efficiency measured by log (Mratio). Mratio = 0 indicates that meta-d’ equals d’. a) confidence-memory accuracy and b) sorting-memory accuracy. Error bars indicate 95% HDI from posterior samples.