Contrasting one's share of the shared life space

Symeonidou, Mariel; Mizokawa, Ai; Kabaya, Shinsuke; Doherty, Martin J.; Ross, Josephine

Published in:
Developmental Science

DOI:
10.1111/desc.13417

Publication date:
2023

Received: 23 July 2022 | Revised: 24 April 2023 | Accepted: 12 May 2023

DOI: 10.1111/desc.13417

Developmental Science WILEY

RESEARCH ARTICLE

Contrasting one's share of the shared life space: Comparing the roles of metacognition and inhibitory control in the development of theory of mind among Scottish and Japanese children

Mariel Symeonidou[1] | Ai Mizokawa[2] | Shinsuke Kabaya[3] | Martin J. Doherty[4] | Josephine Ross[1]

[1]Psychology, School of Humanities, Social Sciences and Law, University of Dundee, Dundee, Scotland

[2]Graduate School of Education and Human Development, Nagoya University, Nagoya, Japan

[3]Department of Psychology, Aichi Shukutoku University, Nagakute, Aichi, Japan

[4]School of Psychology, University of East Anglia, Norwich, UK

Correspondence
Josephine Ross, Psychology, School of Humanities, Social Sciences and Law, University of Dundee, Dundee, DD1 4HN, Scotland.
Email: juross@dundee.ac.uk

Funding information
Experimental Psychology Society, Grant/Award Number: 560808009

Abstract

Cultural comparisons suggest that an understanding of other minds may develop sooner in independent versus interdependent settings, and vice versa for inhibitory control. From a western lens, this pattern might be considered paradoxical, since there is a robust positive relationship between theory of mind (ToM) and inhibitory control in western samples. In independent cultures, an emphasis on one's own mind offers a clear route to 'simulate' other minds, and inhibitory control may be required to set aside one's own perspective to represent the perspective of others. However, in interdependent cultures, social norms are considered the key catalyst for behaviour, and metacognitive reflection and/or suppression of one's own perspective may not be necessary. The cross-cultural generalizability of the western developmental route to ToM is therefore questionable. The current study used an age-matched cross-sectional sample to contrast 56 Japanese and 56 Scottish 3–6-year-old's metacognition, ToM and inhibitory control skills. We replicated the expected cultural patterns for ToM (Scotland > Japan) and inhibitory control (Japan > Scotland). Supporting western developmental enrichment theories, we find that inhibitory control and metacognition predict theory of mind competence in Scotland. However, these variables cannot be used to predict Japanese ToM. This confirms that individualistic mechanisms do not capture the developmental mechanism underlying ToM in Japan, highlighting a bias in our understanding of ToM development.

KEYWORDS
cross-cultural, developmental mechanism, inhibitory control, metacognition, theory of mind

Research Highlights

- We replicate an independent cultural advantage for theory of mind (Scotland > Japan) and interdependent advantage for inhibitory control (Japan > Scotland).
- From a western lens, this pattern might be considered paradoxical, since there is a robust positive relationship between theory of mind and inhibitory control.
- Supporting western developmental enrichment theories, we find that the development of inhibitory control mediates the link between metacognition and theory of mind in Scotland.
- However, this model does not predict Japanese theory of mind, highlighting an individualistic bias in our mechanistic understanding of theory of mind development.

1 | INTRODUCTION

Theory of mind (ToM), defined as the ability to represent own and other's mental states, was first introduced by Premack and Woodruff (1978), who sought to determine if animals shared this special human capacity. In the decades following this paper, ToM became a dominant topic in developmental research, where research is focused on debating the onset of ToM (see Rakoczy, 2022, Wellman et al., 2001 for review). However, the vast majority of research on this topic, and in developmental psychology more generally (Nielsen et al., 2017), has focused on western children. This pervasive sampling bias, where the west is taken to represent the world, is a key challenge for 21st century developmental science to overcome (Nielsen et al., 2017). Whilst topics for developmental study are defined and operationalised by the west, we cannot claim to have developed a full understanding of human development. ToM offers a key case in point, since cross-cultural differences in this capacity have been framed as developmental delays, despite cultures varying in the emphasis placed on mental states as explanatory variables for human behaviour (Lillard, 1998).

Lillard (1998) argues that in independent cultures, an agent's mental states are perceived as the cause of action, whereas in interdependent cultures, external factors are viewed as driving behaviour. Perhaps as a result of this differing emphasis, a number of studies have reported cultural variation in the development of an understanding of false belief, where mental states diverge from reality (e.g. Hughes et al., 2014; Lecce & Hughes, 2010; Liu et al., 2008; Naito & Koyama, 2006; Sabbagh et al., 2006; Shahaeian et al., 2011; Wang et al., 2016; Wellman et al., 2006). These differences in the development of ToM have been directly linked to cultural differences in parents' tendency to reference mental states in conversation with their child (Fujita & Hughes, 2021; Hughes et al., 2018; Taumoepeau et al., 2019), and to children's tendency to use mental states as opposed to external situational factors to explain agent's actions in ToM tasks (Naito & Koyama, 2006). Interestingly, this cross-cultural pattern is reversed for executive function, which tends to develop earlier in interdependent cultures, specifically in relation to inhibitory control (Imada et al., 2013; Moriguchi et al., 2012; Oh & Lewis, 2008; Sabbagh et al., 2006; Senzaki et al., 2018; Wang et al., 2016). This developmental pattern might be considered paradoxical from a western lens, since there is a robust positive relationship between ToM and inhibitory control in western samples (see Devine & Hughes, 2014 for review), thought to arise from the need to suppress one's own mental perspective to represent others'. The current paper explores this paradox, providing direct evidence to suggest that established individualistic mechanisms—premised on the idea that we must overcome an egocentric focus on own mental states to represent other minds (e.g. Damen et al., 2020; Epley, 2008)—are not generalizable to other cultural contexts. Specifically, we find that although cognitive skills such as metacognition (the ability to reflect on own mental states) and inhibitory control (the ability to supress prepotent thoughts or behaviours) predict the development of ToM in Scotland, this model is of no explanatory value in Japan. Selecting our variables based on a western lens leaves only age as a predictor of Japanese theory of mind. This result provides a stark illustration of the inadequacy of our current understanding of the developmental mechanisms underpinning ToM.

1.1 | The development of theory of mind in independent cultures

Simulation theory and theory-theory have historically been offered as competing explanations for the development of ToM (Doherty, 2009). Simulation theorists (Gallese & Goldman, 1998; Goldman, 2006; Harris, 1992) argue that we come to understand other's mental states by reasoning by analogy from our own minds. This theory suggests that metacognitive self-reflection is prerequisite to ToM. Moreover, to the extent that our own perspective is salient, inhibitory control may also be prerequisite. We may need to suppress our own egocentric perspective to represent the perspective of others (Carlson et al., 2015). Theory-theory (Gopnik & Melzoff, 1997; Gopnik & Wellman, 1992) provides an alternative explanation, describing the development of ToM as the process of developing a theory concerning how people will behave given certain mental states or situational factors. It is now

widely accepted that both theories may be involved in the development of ToM, with 'developmental enrichment' accounts proposing that the development of inhibitory control and other executive functions allow implicit/learned forms of ToM to become explicit (Kloo et al., 2020). See Alcalá-López et al. (2019) for a recent review of the literature supporting theoretical mechanisms of theory of mind, including comment on the difficulty inherent in operationalising these theories for empirical testing.

In explanation of the strong relationship between inhibitory control and ToM in interdependent samples, *expression* accounts first contended that inhibitory control was a task requirement involved in false belief tasks, since the participant might require inhibitory control to avoid responding on the basis of current reality (where an object is, as compared to where someone thinks it is). However, decades of research have found little evidence to support the idea that attenuating inhibitory control demands allows younger children to pass explicit ToM tasks (Carlson et al., 2015; Perner & Lang, 1999; Wellman et al., 2001). Moreover, longitudinal evidence suggests that inhibitory control emerges before, and longitudinally predicts ToM (Flynn, 2007; Flynn et al., 2004; Hughes & Ensor, 2007; Marcovitch et al., 2015), suggesting that it is unlikely that inhibitory control demands are masking ToM. As a result, an *emergence* account of the relationship between ToM and inhibitory control is now widely accepted (Rakoczy, 2022). On this reading, the development of inhibitory control offers the cognitive flexibility necessary to reason about other minds, and so might be considered a 'developmental enricher' of ToM (Devine & Hughes, 2014). One reason for this may be the need to control one's own perspective in order to focus on others (in keeping with simulation theory), another possibility is that cognitive control is required to open explicit reflection on other minds (in keeping with theory-theory).

In keeping with the current consensus that aspects of both simulation and theory-theory may be involved in the development of ToM, we propose that basic forms of metacognition may provide the 'developmental enrichment' necessary to think about thinking. The capacity to reflect on current mental states might be the first step to understanding mental states separately from actions, on an explicit cognitive level. When inhibitory control is sufficiently developed to allow one to pause reflection on one's own mental states and turn this lens outwards, this thinking may be extended to more complex reasoning about other minds. This reasoning is in keeping with simulation theory, but on a developmental level. That is, children may not exclusively solve theory of mind tasks by simulating how they would think in another's shoes, but the capacity to reflect on their own mental states may be a necessary developmental stepping stone to explicitly entertaining the possibility of mental states in other people. Importantly, this stance does not preclude children also iteratively developing a theory of other minds by learning from their environment (as described in theory-theory). Rather, we propose that inhibitory control and metacognitive reflection may be the key developmental building blocks to make this theory explicit. However, although the link between inhibitory control and ToM has been robustly established, the role of metacognitive self-reflection in the 'developmental enrichment' of ToM has been little explored (Kim et al., 2020).

There is considerable theoretical overlap in the definitions of metacognition—the ability to reflect upon and report one's own mental states and processes—and ToM—our understanding of our own and others mental states (Flavell, 1979). Indeed, traditional ToM tasks sometimes involve representing one's own mental states. For example, in the unexpected contents task (Hogrefe et al., 1986), children are shown a familiar container (for example, an egg box) and asked what they believe is inside. The container is then opened to reveal unexpected contents (for example, pencils) and the child is asked to recall their original belief. Three-year-olds struggle to recall their own false belief, and instead report thinking that the container held its unusual contents from the outset. This is surprising given the temporal recency of the belief that the container held its usual contents. However, a clear distinction can be drawn between representing past mental states, as in the unexpected contents task, and reporting on current thoughts or perceptions. This form of metacognitive monitoring develops by at least three years, as reflected by children's ability to report on their own current thoughts and feelings (Flavell, 1979, 2002; Salmon & Lucas, 2011), their own ignorance about the contents of a container (Kim et al., 2020; Kim et al., 2016; Rowher et al., 2012; Kloo et al., 2021), or their level of certainty when asked to make a cognitive judgement (for example, whether the line on the left is longer or shorter than the line on the right) in a perceptual certainty monitoring task (Balcomb & Gerken, 2008; Beran et al., 2012; Lyons & Ghetti, 2011, 2013).

Some theorists suggest that we need to have awareness of our own mental states in order to open reflection on others (Goldman, 2006), whereas others argue we first develop a theory of others mental states, and use this understanding to take a 'meta-representational' stance on our own thinking (Carruthers, 2009). Several studies have reported that the ability to impute other's perspectives at 4 years predicts children's metacognitive competence in later childhood (Ebert, 2015; Lecce et al., 2014, 2015; Lockl & Schneider, 2006). However, these studies have focused on metamemory—children's ability to generate appropriate strategies to remember to perform specific tasks (e.g. leaving your lunch bag out that you remember to take it to school). Simpler measures of metacognition, as derived from perceptual certainty monitoring, have shown the reverse pattern, finding that early metacognition predicts later ToM (Feurer et al., 2015; Symeonidou et al., 2022; Cf. Kloo et al., 2021). This mixed pattern of results suggests that the developmental relationship between metacognition and ToM is complex, and likely to depend on the form of metacognition measured. Metacognition is not a unitary concept, and so it is unlikely that we can conclude that all forms of metacognition develop before all forms of ToM, and vice versa.

Moreover, it is important to note that each of the aforementioned studies and theories are derived from western, individualistic samples. Few studies have directly explored developmental mechanisms of theory of mind or explored whether established developmental mechanisms of ToM generalize cross-culturally. This gap in the literature is striking—not least since cultures vary in the explanatory value placed on mental states—but also, since broader cultural differences in the conception of self, and the permeability of boundaries between self and other, are well established (Kagitcibasi, 2005; Markus & Kitayama,

1991), and can be documented early in development (e.g. see Ross et al., 2017). A key aim of the current paper is therefore to determine if this individualistic model of ToM development generalizes to a different cultural context.

1.2 | The development of theory of mind in interdependent cultures

ToM tasks typically require predicting or explaining a behaviour that is difficult to anticipate or understand without recourse to mental states (see Doherty, 2009 for review). For example, in the traditional unexpected transfer task (Wimmer & Perner, 1983), one has to recognise that the protagonist will search for the object in the wrong location due to a mistaken belief that it is there. However, in interdependent cultures, children and parents have been found to focus less on mentalistic explanations or predictions when thinking about other's behaviour, and more on social norms and situational factors (Fujita & Hughes, 2021; Hughes et al., 2018; Mizokawa, 2020; Naito & Koyama, 2006; Taumoepeau et al., 2019). This different focus has been offered as an explanation for different developmental onsets of passing ToM tasks in interdependent relative to independent cultures (e.g. Hughes et al., 2014; Lecce & Hughes, 2010; Liu et al., 2008; Naito & Koyama, 2006; Sabbagh et al., 2006; Shahaeian et al., 2011; Wang et al., 2016; Wellman et al., 2006). Supporting the possibility that the 'developmental enrichment' of ToM varies cross-culturally, there is some evidence of a qualitatively different pathway to ToM in interdependent cultures. Western children are thought to develop an understanding of diverse belief (an understanding that others might believe something different to them) before knowledge access (an understanding that other's might be ignorant) (Kristen et al., 2006; Wellman & Liu, 2004). By contrast, Chinese (Wellman et al., 2006, 2011) and Iranian (Shahaeian et al., 2011) children have been shown to understand knowledge access before diverse belief. This difference has been explained as a process of socialisation placing different cultural emphasises on the role of beliefs versus external situational factors in shaping an individual's behaviour (Wellman et al., 2011).

Importantly, where external factors are the key catalyst for behaviour, one's own mental states may not be salient or relevant to ToM. Moreover, given an external focus, there may be a diminished need to supress one's own egocentric perspective. Metacognition and/or inhibitory control may therefore not provide relevant 'developmental enrichment' for ToM in an interdependent context. Indeed, as set out above, interdependent cultures show an advantage in the development of inhibitory control relative to independent cultures, but this is not accompanied by a boost in ToM reasoning (Oh & Lewis, 2008; Sabbagh et al., 2006; Wang et al., 2016). This implies that inhibitory control may not play a role in the emergence of ToM in interdependent cultures. In support, literature linking inhibitory control and ToM in interdependent samples is mixed. Positive relationships between ToM and inhibitory control are observed within some interdependent samples (e.g. Chasiotis et al., 2006; Duh et al., 2016; Sabbagh et al., 2006; Wang et al., 2016), but the relationship is absent or weaker in others (Kazama et al., 2013; Lewis et al., 2009; Ogawa & Koyasu, 2008; Oh & Lewis, 2008). In contrast, we know little about the cultural differences in the development of metacognition, and how they might relate to ToM. Empirical data documenting the extent to which adults and children in interdependent cultures tend to engage in reflection on their own mental states is lacking (Heyes et al., 2020).

To our knowledge, only one previous study has considered cultural differences in metacognition as related to the development of ToM. Kim et al. (2020) compared German and Japanese 4-year-olds performance on a knowledge access task, where children were asked twice to report on their own knowledge (knowing which toy was in a box), twice to report on their ignorance (not knowing what was in the box), and twice to report on their uncertainty (when they know that one of two toys was placed in a box, but not which toy). Children also chose whether or not to inform another person on the contents of the box in these conditions. Kim et al. (2020) found that metacognition as assessed by the knowledge access task was unrelated to 4-year-old's ToM as measured by a pair of false belief tasks, supporting the idea that these skills are dissociable. Moreover, whereas German 4-year-olds outperformed their Japanese counterparts in ToM, there were no cultural differences in metacognition. However, the metacognition task used by Kim et al. (2020) may lack developmental sensitivity. On the two occasions in which they had only partial knowledge of the contents of the box and were asked to explicitly report on their knowledge, fewer than 50% of trials involved an admission of ignorance. Previous work has shown that although children as young as 3 years can report on their own ignorance in the knowledge access task, children do not reliably volunteer ignorance under partial knowledge conditions (i.e. uncertainty) in this task until the age of 6 years (Kim et al., 2016; Rowher et al., 2012).

Thus, despite strong evidence of developmental connections between metacognition, inhibitory control and ToM in independent cultures, there is a lack of theoretical and empirical evidence to suggest that individualistic models of generalize to other settings. The available evidence suggests that the relationship between inhibitory control and ToM may be weaker in interdependent cultures, but few studies have considered the role of metacognition. Moreover, the majority of studies have focused on cultural comparisons of the developmental onset of skill sets, rather than directly testing models exploring the generalizability of developmental mechanisms of ToM.

1.3 | The current study

Given the evidence reviewed above, the main aim of the current paper is to explore whether individualistic mechanisms of ToM development generalize to a non-individualistic setting. Focusing on the period of emergence of ToM, metacognition and inhibitory control, our age matched sample of 3- to 6-year-olds is drawn from 'independent' Scotland and 'interdependent' Japan. Scotland follows an individualistic philosophy, where self and other are clearly delineated (Dai et al., 2016), parenting goals focus on children gaining independence and autonomy, and behaviour is driven by individual desires and

goals (Barrable, 2020). On the contrary, Japan follows a collectivist philosophy, in which the self is defined in terms of group harmony, and behaviour is driven by social rules (Markus & Kitayama, 1991). The concept of *amae*, or mutual dependence has been used to describe the Japanese mother-child relationship (Doi, 1973). Further emphasising the social nature of self in Japan, the Japanese word for self, *jibun*, translates as "one's share of the shared life space" (Hamaguchi, 1985, cited by Markus & Kitayama, 1991). In addition, developing *omoiyari* (a similar concept to empathy and sympathy), the ability and willingness to respond to the needs of others, is a key goal of child rearing and early childhood education in Japan (Hayashi et al., 2009). Japanese parents and teachers often emphasize to sympathize with other's feeling in their daily conversation with children. This emphasis on social conformity could positively influence Japanese children's development of inhibitory control, and explain a focus on the behaviour of the protagonists and the social rules of the situation in ToM tasks (Naito & Koyama, 2006). Although there are of course further nuanced cultural differences beyond independent/interdependent orientations, these contrasting perspectives offer a good test case for the idea that individualistic models of ToM may not be of explanatory value in collectivistic settings.

The developmental onset of passing of ToM tasks has been well documented in Japan (Fujita et al., 2022; Hughes et al., 2014; Lewis et al., 2009; Naito & Koyama, 2006; Wellman et al., 2001). Wellman et al.'s (2001) influential meta-analysis of ToM development concluded that on average Japanese children may pass ToM tasks up to 2 years after western children. Direct comparisons of Japanese and British children (Fujita et al., 2022; Hughes et al., 2014; Lewis et al., 2009; Naito & Koyama, 2006) have since replicated this finding in ToM performance to be robust. However, of the four previous studies to explore the relationship between inhibitory control and ToM within Japan, none found a correlation that survives correction for age (Fujita et al., 2022; Kazama et al., 2013; Lewis et al., 2009; Ogawa & Koyasu, 2008). In comparison, the relationship between inhibitory control and ToM in British samples is robust (e.g. Flynn, 2007; Flynn et al., 2004; Hughes & Ensor, 2007; cf. Fujita et al., 2022). This pattern of results is suggestive of mechanistic differences in the development of theory of mind in Japan.

As set out above, one possibility is that Japanese children develop ToM later than western children due to a fundamental culture difference in 'mind-mindedness' that is the tendency to invoke mental states (such as beliefs or desires) as explanations for behaviour. In a context where parents emphasise the external, social factors that influence behavioural choices (Hughes et al., 2018), children may find it more natural to appeal to social norms when predicting other's behaviour (Mizokawa, 2020; Naito & Koyama, 2006) than to consider their own or others' mental states. If this is the case, we would expect to replicate the developmental pattern of reduced ToM performance in Japan relative to the UK, and the diminished role for inhibitory control in Japanese ToM. We might also see a relatively weak connection between metacognition and ToM in Japan relative to Scotland, since a lack of emphasis on mental states as explanations for behaviour would reduce the theoretical links between reflecting on own mental states and solving theory of mind tasks. To test these hypotheses, we asked an age matched sample of Scottish and Japanese 3- to 6-year-old children to complete a battery of traditional ToM and inhibitory control tests designed to capture development change and cultural differences across this age range. The ToM battery included questions focused on other's mental states, allowing us to separate metacognitive reflection from theory of mind. Alongside these tests, children completed a perceptual certainty monitoring task. 3- and 4-year-old children have proven capable of reporting their own uncertainty in perceptual certainty monitoring paradigms; where the stimuli are patently ambiguous and the uncertainty response is modelled and encouraged (e.g. do you feel sure or unsure as indicated by a pictorial scale) (Balcomb & Gerken, 2008; Beran et al., 2012; Lyons & Ghetti, 2011, 2013). This paradigm was chosen because it offers a sensitive test of metacognition in the focal age range, which does not overlap in form with the traditional theory of mind paradigm. We can therefore be more confident that relations between theory of mind and this metacognition task, where present, might represent a common underlying mechanism, as opposed to common task demands.

The certainty monitoring task has rarely been explored from a cross-cultural lens. Kim et al. (2021) compared 3.5–5-year-old Japanese and German children's ability to accurately judge their memory for pictorial stimuli, and asked them to sort the stimuli in order of their confidence. They found that although both German and Japanese children performed similarly in the certainty monitoring task, Japanese children's sorting was more closely matched to their accuracy. Testing young adults in a more complex paradigm, van der Plas et al. (2022) found that when presented with post-decision clues to the correct answer in a psychophysical decision task, Chinese adults were better able to adjust confidence in their original decision than English adults. These studies suggest that metacognitive sensitivity may vary cross-culturally, in a direction that might be considered paradoxical to the western lens, that is despite an overlap in the definition of theory of mind and metacognition from a western perspective, and strong evidence of a developmental differences in Asian theory of mind, there is no evidence that metacognition develops differently for Asian children. Critically, we are focused not on cultural differences in skill sets, but on the generalisability of proposed developmental mechanisms. Accordingly, we test for the first time whether the inhibitory control and basic metacognitive reflection can be used to predict ToM in both Scottish and Japanese children, or if these proposed 'developmental enrichers' of ToM are culturally bound.

2 | METHODS

2.1 | Participants

Fifty-six Japanese 3–6-year olds took part, 25 male, $M = 59.70$ months, $SD = 8.59$. Japanese children were recruited in metropolitan Tokyo, Kyoto and Aichi, and tested in a traditional kindergarten setting in their native language by the second and third authors, all children were native to Japan and the majority were Asian. The Japanese children were age matched with 56 Scottish 3–6-year-olds, 28 male, $M = 59.55$

months, *SD* = 9.56. Scottish children were recruited in urban Dundee, and tested in a traditional nursery setting in their native language by the first author, all children were native to Scotland and the majority were white. Children were tested with parental consent and their own assent, and ethical review and approval was provided by University of Dundee non clinical research ethics committee in keeping UKRI ethics principles. Although individual demographic data was not available, both samples were drawn from urban cities from a selection of kindergartens, capturing a range of socio-economic conditions. Thirteen of the Scottish children included here went on to contribute follow up data to Symeonidou et al. (2022). The sample size was comparable to that of previous research (Hughes et al., 2014; Kim et al., 2020, 2021) and determined by convenience sampling. A univariate ANOVA with age (in months) as a dependent variable and culture (Japan vs. Scotland) as a between subjects variable confirmed that the different cultural groups were age-matched, F (1, 110) = 0.007, $p = 0.935$ $\eta^2 <$ 0.001. To allow comparison between age groups whilst preserving the age match across cultures, a median split ($M = 56$ months) was used to split the sample into 55 older, $M = 67.09$, $SD = 7.18$, 26 Scottish, and 57 younger children, $M = 52.42$, $SD = 3.54$, 30 Scottish. Sensitivity for moderate main effects and interactions was calculated before analysis using G Power predictions for a 2 × 2 design (power = 80%, $p = 0.05$, Cohen's $f = 0.30$) (Faul et al., 2007), which suggested a minimum sample size of 90 participants. Sensitivity for correlations was also calculated a priori, based on the robust ($r <= 0.35$) correlation between ToM, G power suggested a minimum sample size of 49 participants per cultural group to achieve 80% power. Finally, based on a moderate $f^2 = 0.25$, and three predictor variables, G power suggested a minimum sample size of 48 participants per cultural group to achieve 80% power in regression analyses.

2.2 | Materials and procedure

Assessments were administered over two sessions to ensure that children were not fatigued. In the first session children completed the inhibitory control and ToM tasks and in the second session (1 day to 1 week later depending on convenience) children completed the certainty monitoring task used to assess metacognition. Task instructions and narratives were translated from English to Japanese by the second and third authors, who are native to Japan, and fluent in English. Pictorial stimuli were the same across cultures, but physical objects used (for example, in the unexpected contents task) were sourced in the relevant country in order to be culturally appropriate.

2.3 | ToM scale

Children's understanding of other's mental states was assessed using a modified version of Wellman and Liu's (2004) five item ToM battery, where the other was represented by a puppet known as "Teddy" or the experimenter, dependent on the task. Like Wellman and Liu (2004), our scale contained an assessment of the child's understanding of diverse beliefs (where will Teddy look for an item if he thinks it is hidden somewhere different from you?), knowledge access (does Teddy know what is in the box?), false beliefs (what will Teddy think is inside a familiar package?) and hidden emotions (how will Teddy look when trying to hide disappointment?). To increase sensitivity to developmental change in the age range chosen, the diverse desires task in the original battery was substituted with an additional false belief task; Wimmer and Perner's (1983) original unexpected transfer task (where will Teddy look for an item that has been moved without his knowledge?). The diverse desires task can be passed by 18 months (Repacholi & Gopnik, 1997), whereas the unexpected transfer task is typically not passed until between 4 and 6 years, dependent on cultural setting (Wellman et al., 2001).

We included four additional questions in our ToM battery in order to extend the breadth and sensitivity of the scale. This included two additions to the original scale; asking the children to explain how Teddy could find out what was in the box in the knowledge access task, and to track changes in Teddy's feelings during the unexpected contents task, when the actual contents of an attractive container were revealed to be disappointing (adapted from Bender et al., 2011). We also included Russell et al. (2001) transparent intentions task. This task involved showing the child line drawings presented on transparent acetate slides. When the slides overlapped, they appeared to create one image (for example, a mouse), with a missing feature. The experimenter pointed out the missing feature (for example, the mouse had no tail) and proceeded to add it to the drawing. However, when the top acetate was lifted from the bottom, the picture was transformed (for example, the mouse's body, removed from context, became the shape of a balloon). This meant that the added feature changed interpretation (for example, the 'tail' looked like the string of the balloon). The child was asked "Did I" *think* I was drawing a ["tail on a mouse"] or ["a balloon string"]?; Did I *mean* to draw a ["tail on a mouse"] or a ["balloon string"]? (order counterbalanced). Children receive on point per correct answer. Previous research suggests that western children struggle with 'false intention' tasks, even when they have mastered false belief (Lang & Perner, 2002). Together with the scaled tasks, these additions resulted in a ToM score which summed to a possible total of nine points, the proportion of trials passed was then calculated to create the final score. As in Wellman and Liu (2004), the order of tasks was counterbalanced, and control questions were included where relevant (e.g. checking that children followed the narrative and knew the final location of the object in the unexpected transfer task). Appendix 1 describes the theory of mind task battery in detail.

Preliminary analyses (see Appendix 2) found no evidence of an 'interdependent' reversal in sensitivity to knowledge over desire previously observed in China (Wellman et al., 2006; Wellman et al., 2011) and Iran (Shahaeian et al., 2011). Instead, the tasks scaled in 'western' order (Kristen et al., 2006; Wellman & Liu, 2004) from an understanding of desire, to knowledge, to an understanding of false beliefs, to hidden emotions. This could point to differences between our sample and other interdependent samples, since broad distinctions/similarities between geographical location are of course inadequate to full characterise cross-cultural differences.

2.4 | Inhibitory control tasks

Inhibitory control was assessed using two established tests of inhibitory control, one before the ToM battery, and the other afterwards (order counterbalanced). We used Kochanska et al. (1997) instructions to implement an adaptation of Reed et al. (1984) *Bear-Dragon Task*. Children were introduced to two puppets, one of whom was "nice", and the other who was "mean". Both puppets gave the children simple action instructions (e.g. touch your head). Children were asked to follow the instructions of the nice puppet and ignore the instructions of the mean puppet. Following a set of practice trials, the children took part in 10 test trials, where the puppets gave instructions in pseudorandomised in order. Children were given 3 points for full imitation of the "nice" puppet's actions, 2 points for partial or delayed imitation, 1 point for a wrong movement, and 0 points for no movement. Conversely, children were given 3 points for ignoring "mean" puppet's action commands, 2 points for partially imitating the action, 1 point for a wrong movement, and 0 points for fully imitating the mean puppet. Children's score out of a possible total of 30 indicates their capacity for behavioural inhibitory control. Cognitive inhibitory control was measured using Gerstadt et al. (1994) *Day-Night Stroop Task*. During this game children were instructed to say "Night" when shown a picture of a sun and "Day" when shown a picture of a moon. Following a practice in which children established the rules of the game, children completed 16 test trials in a pseudorandomised order. A point was given for each correct trial, building to a possible total of 16. Since these tasks reliably correlate (and were correlated in our sample, $r = 0.477$, $p < 0.001$) the proportion of correct trials for each task was then averaged to create the final score.

2.5 | Metacognition task

An adaptation of Lyons and Ghetti's (2013) perceptual discrimination task was used to measure children's ability to monitor their own certainty. Children were required to identify a target from two line drawings, which had been digitally pixelated (to 80%) to create ambiguity. Taken from a set curated by Cycowicz et al. (1997), the line drawings were displayed on a laptop, presented using DMDX coding software (Forster & Forster, 2003). Children were asked to point to the target named by the experimenter (e.g. Can you find the "duck"?), and their choice (left or right) was recorded in DMDX by the experimenter. Following selection of the target, a confidence judgement screen appeared, where the child saw an image of a child of the same gender showing confident (positive) and unconfident (negative) facial expressions. The child was asked "are you sure" or "not so sure" that you have found the target (e.g. the duck)? The child's response (verbal, or pointing to the relevant prompt) was recorded in DMDX by the experimenter. Children completed 4 practice trials, and 10 test trials. Following Lyons and Ghetti (2011, 2013) and Feurer et al. (2015), we calculated the proportion of sure trials on which the child had accurately selected the target (correct trials where the child reported being sure/the total of trials where the child reported being sure), and subtracted from this to the proportion of sure trials when they were inaccurate (incorrect trials where the child reported being sure/the total number of trials where the child reported being sure). By calculating the accuracy of sure trials in this way, we were able to determine whether children were able to successfully match their certainty to their accuracy. Successful certainty-monitors should be more accurate when sure, and thus higher certainty difference scores indicate more metacognitive sensitivity. Univariate ANOVAs with culture (Japan vs. Scotland) as a between subjects variable confirmed there was no main effect of culture on accuracy, $F(1,108) = 0.001$, $p = 0.780$, $\eta p^2 = 0.001$, or certainty (tendency to use the unsure response), $F(1,108) = 0.001$, $p = 0.961$, $\eta p^2 = 000$ within the perceptual certainty monitoring task, ensuring that the children started from an equal baseline. Note: we do not include unsure responses in our metacognitive sensitivity calculations as where there is a binary choice, children are likely to be accurate on 50% of 'unsure' trials simply through guessing. As a result, we would not predict that certainty in unsure trials will correspond directly with ultimate accuracy. On the contrary, sensitive metacognitive monitors should be accurate for 100% of 'sure' trials, making this a clearer measure of metacognitive competence.

3 | RESULTS

Table 1 shows the average proportion of trials passed for theory of mind and inhibitory control tasks, split by age and culture. There was clear age-related improvement in both tasks. Multivariate ANOVA confirmed main effects of age for ToM performance, F (1, 108) = 35.594, $p < 0.001$, $\eta^2 = 0.248$. There was no main effect of culture, F (1, 108) = 1.586, $p = 0.211$, $\eta^2 = 0.014$. However, there was an interaction between culture and age predicting ToM performance, F (1, 108) = 4.078, $p = 0.046$, $\eta^2 = 0.036$. Simple comparisons showed that although there was a main effect of age on ToM scores in both Scotland, F (1, 54) = 32.807, $p < 0.001$, $\eta^2 = 0.378$, and Japan, F (1, 54) = 7.57, $p = 0.008$, $\eta^2 = 0.123$, the magnitude of the change was significantly larger in Scotland (.333) than Japan (.165), $t(110) = -2.02$, $p = 0.046$. Moreover, when compared directly, younger Japanese children performed similarly to their Scottish counterparts, F (1, 54) = 0.291, $p = 0.591$, $\eta^2 = 0.005$. However, older Scottish children had significantly higher ToM scores than older Japanese children, F (1, 54) = 5.327, $p = 0.025$, $\eta^2 = 0.09$.

For inhibitory control performance, there were main effects of age, F (1, 108) = 23.61, $p < 0.001$, $\eta^2 = 0.179$ and culture, F (1, 108) = 4.812, $p = 0.030$, $\eta^2 = 0.043$. There was no interaction between these main effects, F (1, 108) = 1.227, $p = 0.270$, $\eta^2 = 0.011$. As shown in Table 1, Japanese children of all ages had higher inhibitory control scores than Scottish children.

For Scottish children, both inhibitory control, $r = 0.464$, $p < 0.001$, and ToM, $r = 0.615$, $p < 0.001$, improved with age, and robustly correlated, $r = 0.667$, $p < 0.001$, $r^{partial} = 0.546$, $p < 0.001$. Inhibitory control also improved with age for Japanese children, $r = 0.431$, $p = 0.001$, confirming that although Japanese performance approached ceiling in inhibitory control, the sample retained sufficient variance to capture

TABLE 1 Average proportion (and standard deviation) of trials passed for theory of mind (ToM) and inhibitory control, split by age and culture.

	ToM	Younger	Older	Inhibitory control	Younger	Older
Japan	**0.587(0.237)**	0.502(0.239)	0.667(0.207)	**0.918(0.113)**	0.861(0.131)	0.970(0.057)
Scotland	**0.625(0.273)**	0.470(0.203)	0.803(0.232)	**0.846(0.210)**	0.765(0.205)	0.939(0.178)
Overall	0.606(0.255)	0.485(0.220)	0.731(0.228)	0.881(0.171)	0.811(0.179)	0.956(0.129)

Note: Younger Japanese children ($N = 27$, $M^{age} = 52.37$), Older Japanese children ($N = 29$, $M^{age} = 66.52$); Younger Scottish children ($N = 30$, $M^{age} = 52.37$), Older Scottish children ($N = 26$, $M^{age} = 66.52$).

TABLE 2 Average certainty difference (and standard deviation) score split by age and culture.

	Certainty difference score	Younger	Older
Japan	**0.525 (0.340)**	0.401(0.382)	0.642(0.251)
Scotland	**0.625(0.293)**	0.558(0.322)	0.705(0.234)
Overall	0.574(0.320)	0.481(0.357)	0.672(0.243)

Note: Younger Japanese children ($N = 27$, $M^{age} = 52.37$), Older Japanese children ($N = 29$, $M^{age} = 66.52$); Younger Scottish children ($N = 30$, $M^{age} = 52.37$), Older Scottish children ($N = 26$, $M^{age} = 66.52$).

developmental change. There was also age-related improvement in Japanese ToM, $r = 0.416$, $p = 0.001$. However, the correlation between ToM and inhibitory control in Japan failed to reach significance, $r = 0.260$, $p = 0.053$, $r^{partial} = 0.098$, $p = 0.478$. This pattern of results replicates and extends observations from previous literature, showing cultural variation in the association between ToM and inhibitory control performance.

Table 2 shows children's certainty difference scores. Metacognitive sensitivity increased with age, and Scottish children consistently outperformed Japanese children. Although multivariate analyses confirmed a moderate main effect of age, $F(1, 108) = 11.728$, $p < 0.001$, $\eta^2 = 0.098$, the small main effect of culture failed to reach significance, $F(1, 108) = 3.522$, $p = 0.063$, $\eta^2 = 0.032$, and there was no interaction between age and culture, $F(1, 108) = 0.599$, $p = 0.441$, $\eta^2 = 0.006$.

Pearson's correlation analyses suggested a positive link between metacognition and age for both Japan, $r = 0.356$, $p = 0.007$, and Scotland, $r = 0.357$, $p = 0.007$. However, although there was a robust link between metacognitive sensitivity and ToM for Scottish children, $r = 0.480$, $p < 0.001$, $r^{partial} = 0.353$, $p = 0.008$, this link was not robust to age correction for Japanese children, $r = 0.344$, $p = 0.009$, $r^{partial} = 0.231$, $p = 0.090$. Interestingly, there was a robust positive relationship between metacognitive sensitivity and inhibitory control in both Japan, $r = 0.433$, $p < 0.001$, $r^{partial} = 0.332$ $p = 0.013$ and Scotland, $r = 0.621$, $p < 0.001$, $r^{partial} = 0.550$, $p < 0.001$.

To further explore the interrelations between our variables, we modelled ToM performance in Scotland and Japan using metacognition, inhibitory control and age in months as predictors in linear regression analyses. For Scottish children, the model was significant, predicting 57% of the variance in ToM understanding, $F(3, 55) = 22.625$, $p < 0.001$. A larger proportion of this variance was attributable to age and inhibitory control than to metacognition ($\beta^{age} = 0.384$, $p < 0.001$; $\beta^{inhibitory\ control} = 0.449$, $p < 0.001$; $\beta^{metacognition} = 0.064$, $p = 0.588$). For Japanese children, the model was also significant, but predicted only 22% of the variance in ToM understanding, $F(3, 55) = 4.829$, $p = 0.005$, largely accounted for by age ($\beta^{age} = 0.329$, $p = 0.022$; $\beta^{inhibitory\ control} = 0.024$, $p = 0.869$; $\beta^{metacognition} = 0.217$, $p = 0.125$).

Since there was robust relationship between metacognition and inhibitory control in both Japan and Scotland, follow up analyses using Hayes' (2013) Macro Process Model 4 were used to consider whether metacognition had an indirect effect on theory of mind via inhibitory control in either Japan or Scotland, with age in months entered as a covariate. We used a bias corrected 95% CI around the indirect effect from 5000 bootstrap re-samples, accepting the indirect effect as statistically significant only if its bias corrected 95% CI excluded zero. As shown in Figure 1, there was a direct effect of inhibitory control, and a significant indirect effect of metacognition on ToM in Scotland. On the contrary, there was no evidence of a direct influence of inhibitory control on ToM in Japan, and no evidence of an indirect or direct effect of metacognition. Note: this analysis is aimed at thoroughly exploring the complex relations between inhibitory control and metacognition which may relate to ToM since regression analyses can only consider independent effects. This analysis should be interpreted with caution, and not be considered a causal mediation analysis given the data are cross-sectional.

4 | DISCUSSION

We observed cultural differences in development across Scotland and Japan, replicating increased ToM in an independent context (Hughes et al., 2014; Lewis et al., 2009; Naito & Koyama, 2006; Wellman et al., 2001) and increased self-control in an interdependent context (Imada et al., 2013; Senzaki et al., 2018). Despite similar levels of metacognitive sensitivity, metacognition and inhibitory control showed weaker relationships with ToM in Japan than Scotland. Once common age-related change was controlled for, no significant relationship was evident between metacognition and ToM, or inhibitory control and ToM in Japan. Measuring age, metacognition and inhibitory control predicted 57% of the variance in Scottish ToM, with metacognition relating to ToM indirectly through inhibitory control. However, only 22% of the variance in Japanese ToM could be explained by the combination of these factors, with age emerging as the only

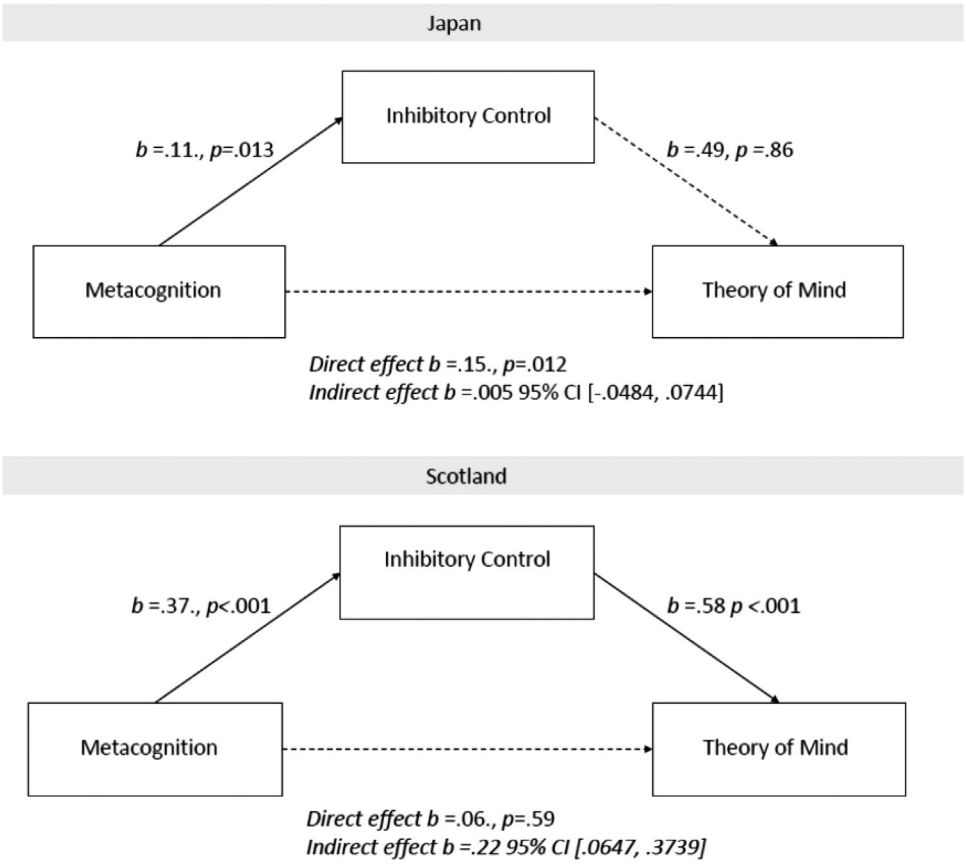

FIGURE 1 Developmental models linking metacognition, inhibitory control and theory of mind in Japan and Scotland.

independent predictor. There was no evidence of direct or indirect relations between inhibitory control mediating the link between and metacognition predicting ToM in Japan. Following the framework set out in the introduction, this pattern could be interpreted as consistent with cultural differences in the tendency to consider (and ultimately suppress) one's own perspective when computing the perspective of others', or the tendency to reason about other's mental perspective rather than their behaviour. Importantly, these results suggest confirm that there are likely to be mechanistic differences in the development of ToM across Scotland and Japan, which should be explored longitudinally.

Although younger Scottish and Japanese children performed similarly on the ToM scale, older Scottish children, passed 80% of tasks relative to their Japanese counterparts 67%. Alongside age related improvement in ToM for both cultures, this pattern is consistent with a small difference in the development of Scottish and Japanese children (Naito & Koyama, 2006; Wellman et al., 2001) which the evidence suggests will remedy with age (Bradford et al., 2018; Kobayashi et al., 2006; Wu & Keysar, 2007). Similarly, although Japanese children had increased inhibitory control relative to their western counterparts (replicating Imada et al., 2013; Senzaki et al., 2018), this advantage was unlikely to have real world relevance beyond 56 months, where both groups were close to ceiling in accuracy.

Metacognitive sensitivity improved with age, and although the raw scores showed a small Scottish advantage for certainty monitoring, this difference did not reach significance. This suggests that the *capacity* to reflect and report on one's mental states—as necessitated by the perceptual certainty monitoring task—did not differ with culture in our sample, replicating Kim et al. (2020). This is perhaps unsurprising, given the metacognitive monitoring task used here requires simply that children make public their conscious experience of certainty/uncertainty—an aspect of human experience that we can assume to be universal. We measure only one, simple, form of metacognition here, and it is clear that more work is needed to characterise cross-cultural differences in metacognitive capacity more broadly, including exploration of other forms of metacognition (Heyes et al., 2020; Kim et al., 2021; van der Plas et al., 2022).

It is interesting to note that metacognition and inhibitory control were robustly related in both Japan and Scotland, suggesting that some cognitive control may be required to take a metacognitive stance, as previously argued for western samples (Kälin & Roebers, 2022). However, Zhao et al. (2021) suggest that whereas children from the independent cultures may attribute their ability to control themselves to internal factors such as the ability to inhibit desires, children from interdependent cultures associate inhibitory control with complying to external rules, without reflection of their own willpower. This opens the possibility that despite similar capacities for metacognitive

reflection the *tendency* to reflect on one's own mental states may vary cross-culturally. Indeed, Naito and Koyama (2006) consider the possibility that Japanese children may find ToM tasks more challenging precisely because they are not used to being asked to reflect on their own or others' mental states, or hearing others speculate about these internal activities. Thus, in the context of ToM tasks, when asked to predict or explain behaviour, they may not anticipate that the researcher intends them to respond in terms of mental states, impacting on their performance.

It is possible that assessment of our mechanistic model was compromised by inhibitory control approaching ceiling for Japanese children. However, we did capture a relationship between age, inhibitory control, and metacognition in Japanese children, suggesting that there was sufficient variance in performance to support correlation. The Japanese disconnect between inhibitory control and ToM is in any case made clear by the developmental profile of Japanese children: high in inhibitory control, low in ToM. Alongside neurocognitive evidence from adults (Frank & Temple, 2009; Koelkebeck et al., 2011), these different cultural patterns suggest that the 'problem' of other minds is likely to be framed and solved differently dependent on cultural context. Specifically, we speculate that the Japanese lag in ToM development may be explained by the need for Japanese children to gradually build a theory of behaviour that incorporates misunderstandings.

However, it is also possible that the Japanese lag in ToM found here is due to bias in methodology. We rely on western framing of the ToM questions set, since established measures of ToM are needed to identify universality. However, previous research may suggest that when the western bias in ToM methodology is reduced, for example by using a less linguistically challenging narrative (Moriguchi et al., 2010), Japanese children perform better. Going forward, we suggest that a paradigm shift is needed, where researchers develop and adopt culturally appropriate tests to better understand the universality and mechanistic development of ToM. Although ToM has dominated the developmental literature for decades (Beaudoin et al., 2020), we need to recognise that the emphasis on divergent minds and mental states is a western preoccupation, dependent on a hard boundary between self and other which does not necessarily exist in other cultural settings. It is still an open question how ToM develops in interdependent contexts, and the term ToM may itself be misleading given the emphasis on external causes of action.

It is interesting to note that the traditional explanatory mechanisms for ToM, including simulation theory (Gallese & Goldman, 1998; Goldman, 2006; Harris, 1992) and theory-theory (Gopnik & Melzoff, 1997; Gopnik & Wellman, 1992), have rarely been considered through a cross-cultural lens, despite a sustained flurry of activity comparing performance in ToM cross-culturally (though see Bradford et al., 2018 for an exception). The lack of emphasis on theory is unfortunate, since paying close attention to the dominant mechanisms proposed to explain the development of ToM highlights culturally bound assumptions. As set out in the introduction, simulation theory proposes that we use 'ready' access to our own minds to become aware of mental states in general, and (with the help of inhibitory control) imagine ourselves in the position of other to simulate their internal state. Likewise, theory-theory proposes that we gradually build a picture of our own and other's minds, and the actions that may result from these states, through our experience of social discourse regarding mental states. Both theories depend on internal mental states to be foregrounded as the causes of action, and neither theory (nor the demands of the traditional ToM tasks themselves) allows for action to be explained on the basis of external factors. It is somewhat circular then, to consider cultural lags in ToM in interdependent relative to independent cultures as representative of a true developmental difference in our ability to understand others.

Instead, our emphasis might usefully be changed to a comparison of how ToM problems are routinely solved across cultures. When we take that position here, we find that although the development of metacognition and inhibitory control skills are of explanatory value in Scotland, they do not usefully relate to Japanese performance in ToM tasks. For Scottish children, metacognition and inhibitory control were robustly related to ToM. This is consistent with inhibitory control being necessary for the expression or emergence of ToM in Scotland, since one's own perspective must be actively supressed to represent other. However, for Japanese children, neither metacognition nor inhibitory control were robustly related to ToM, and age (acting as a proxy for unmeasured skills) was the only independent predictor.

Here we capture, but cannot explain, age related change in ToM in Japan. Other studies have considered and found evidence for cultural differences in linguistic competence (Moriguchi et al., 2010) and family circumstances (Naito & Koyama, 2006) which may influence the development of ToM universally, and it is unfortunate that we could not include these more nuanced variables here. It is also unfortunate that we rely on a cross-sectional comparison, since cross-sectional work is limited in its ability to explore developmental mechanisms, which are temporal in nature. Therefore, we would encourage the collection of longitudinal data in non western contexts. Our results make clear is that at least two developmental 'enrichers' of ToM in Scotland have limited relations to ToM in Japan. We would argue based on this data that characterisation of interdependent ToM from the perspective of what environmental variables matter in the west is insufficient. Future studies should address *theoretical* explanations for the development of ToM in Japan which are embedded in and emerge from the interdependent cultural context—based on qualitative and quantitative findings, and not requiring a western comparison to be judged relevant. More generally, 'developmental enrichment' accounts of ToM must be extended to recognise that the ways in which children's understanding of other's minds is enriched is likely to vary widely cross-culturally, parallel to different emphases on self.

5 | CONCLUSIONS

Our study demonstrates that there may be cultural differences in the developmental mechanisms underlying ToM. Replicating the expected cultural patterns for ToM (Scotland > Japan) and inhibitory control (Japan > Scotland), we find that inhibitory control, metacognitive sensitivity, and age offer a good model of ToM in Scotland (see also

Symeonidou et al., 2022). However, the same model was significantly less successful in predicting Japanese ToM, with age emerging as the only independent predictor. This pattern of results can be related to *theoretical* models of ToM. Specifically, we suggest that the independent construal of self as distinct from, but similar to, others offers the simulation of other minds as an efficient strategy to solve ToM problems—provided the child has developed sufficient inhibitory control. Children in independent contexts may also use metacognition to build a theory of other minds, involving flexibly reasoning about other's mental perspectives, requiring self-control to make this theory explicit. However, the interdependent construal of self as bound with others, and the focus on shared external perspective, may make simulation or reasoning about minds a less natural strategy for Japanese children. Based on a western lends, one might speculate that ToM problems may be solved more gradually in interdependent settings, through a process of building a culturally appropriate ToM based on external social cues. However, we call for further investigation of the definition and development of ToM in interdependent contexts, since the applicability of the term ToM may be culturally bound.

More generally, we join our colleagues (Amir & McAuliffe, 2020; Nielsen et al., 2017; Singh et al., 2023) in calling for broader sampling of developmental populations to allow for participative design of the questions, operationalisations and outcomes of developmental science. This is likely to necessitate combining qualitative and quantitative work (Amir & McAuliffe, 2020), and moving beyond the cross-sectional perspective. The current study neatly demonstrates that given a pervasive history of sampling bias, relying on established developmental research to propose and test theoretical predictions in other settings is unlikely to offer a fruitful starting point to understand experiences in other contexts. Indeed, attempting to focus on commonalities in development undermines the idea that the human mind is a product of our social context (Nielsen et al., 2017; Singh et al., 2023). Ironically, across the same time period as ToM has dominated developmental science, animal research has shown the human capacity for culture is a credible contender for unique human specialisation (O'Madagain & Tomasello, 2022). Thus, the 'special' nature of human cognition likely lies in our diversity, flexibility and capacity to learn from others. On this reading, inclusivity in our approach to research is critical for developmental science to thrive.

ACKNOWLEDGMENTS

This project was approved by the University of Dundee non-clinical research ethics committee, and made possible by funding received from the Experimental Psychology Society 560808009 and a University of Dundee PhD studentship held by the first author. We would like to thank the children and families who took part. There is no known conflict of interest. Data is freely available on the open science framework at https://osf.io/rhvmt/.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data is freely available on the open science framework at https://osf.io/rhvmt/.

REFERENCES

Alcalá-López, D., Vogeley, K., Binkofski, F., & Bzdok, D. (2019). Building blocks of social cognition: Mirror, mentalize, share? *Cortex*, *118*, 4–18. https://doi.org/10.1016/j.cortex.2018.05.006

Amir, D., & McAuliffe, K. (2020). Cross-cultural, developmental psychology: Integrating approaches and key insights. *Evolution and Human Behavior*, *41*(5), 430–444. https://doi.org/10.1016/j.evolhumbehav.2020.06.006

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*(5), 750–760. https://doi.org/10.1111/j.1467-7687.2008.00725.x

Barrable, A. (2020). Shaping space and practice to support autonomy: Lessons from natural settings in Scotland. *Learning Environments Research*, *23*, 291–305. https://doi.org/10.1007/s10984-019-09305-x

Beaudoin, C., Leblanc, É., Gagner, C., & Beauchamp, M. H. (2020). Systematic review and inventory of theory of mind measures for young children. *Frontiers in Psychology*, *10*, 2905. https://doi.org/10.3389/fpsyg.2019.02905

Bender, P. K., Pons, F., Harris, P. L., & de Rosnay, M. (2011). Do young children misunderstand their own emotions? *European Journal of Developmental Psychology*, *8*(3), 331–348. https://doi.org/10.1080/17405629.2010.495615

Beran, M. J., Decker, S., Schwartz, A., & Smith, J. D. (2012). Uncertainty monitoring by young children in a computerized task. *Scientifica*, Article ID 692890, 6 pages. https://doi.org/10.6064/2012/692890

Bradford, E. E. F., Jentzsch, I., Gomez, J-C., Chen, Y., Zhang, D., & Su, Y. (2018). Cross-cultural differences in adult ToM abilities: A comparison of native-English speakers and native-Chinese speakers on the self/other differentiation task. *The Quarterly Journal of Experimental Psychology*, *71*(12), 2665–2676. https://doi.org/10.1177/1747021818757170

Carlson, S. M., Claxton, L. J., & Moses, L. J. (2015). The relation between executive function and theory of mind is more than skin deep. *Journal of Cognition and Development*, *16*(1), 186–197. https://doi.org/10.1080/15248372.2013.824883

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *Behavioral and Brain Sciences*, *32*(2), 121–138. https://doi.org/10.1017/S0140525X09000545

Chasiotis, A., Kiessling, F., Hofer, J., & Campos, D. (2006). Theory of mind and inhibitory control in three cultures: Conflict inhibition predicts false belief understanding in Germany, Costa Rica and Cameroon. *International Journal of Behavioral Development*, *30*(3), 249–260. https://doi.org/10.1177/0165025406066759

Cycowicz, Y. M., Friedman, D., Rothstein, M., & Snodgrass, J. G. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology*, *65*(2), 171–237. https://doi.org/10.1006/jecp.1996.2356

Dai, Q., Williams, J., & McGregor, E. (2016). Who Am I? The Developing Self-Concept of Scottish-Born Chinese Children: A Comparison With White Scottish, Mainland Chinese, and Hong Kong Chinese Children. *Identity*, *16*(4), 239–249. https://doi.org/10.1080/15283488.2016.1229607

Damen, D., van der Wijst, P., van Amelsvoort, M., & Krahmer, E. (2020). Can the curse of knowing be lifted? The influence of explicit perspective-focus instructions on readers' perspective-taking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(8), 1407. https://psycnet.apa.org/doi/10.1037/xlm0000830

Devine, R. T., & Hughes, C. (2014). Relations between false belief understanding and executive function in early childhood: A meta-analysis. *Child Development*, *85*(5), 1777–1794. https://doi.org/10.1111/cdev.12237

Doherty, M. (2009). *Theory of mind: How children understand others' thoughts and feelings*. Psychology Press, Taylor & Francis.

Doi, L. T. (1973). The Japanese patterns of communication and the concept of amae. *Quarterly Journal of Speech*, *59*(2), 180–185. https://doi.org/10.1080/00335637309383166

Duh, S., Paik, J. H., Miller, P. H., Gluck, S. C., Li, H., & Himelfarb, I. (2016). Theory of mind and executive function in Chinese preschool children. *Developmental Psychology*, *52*(4), 582–591. https://doi.org/10.1037/a0040068

Ebert, S. (2015). Longitudinal relations between theory of mind and metacognition and the impact of language. *Journal of Cognition and Development*, *16*(4), 559–586. https://doi.org/10.1080/15248372.2014.926272

Epley, N. (2008). Solving the (real) other minds problem. *Social and Personality Psychology Compass*, *2*(3), 1455–1474. https://doi.org/10.1111/j.1751-9004.2008.00115.x

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Feurer, E., Sassu, R., Cimeli, P., & Roebers, C. (2015). Development of meta-representations: Procedural metacognition and the relationship to theory of mind. *Journal of Educational and Developmental Psychology*, *5*(1), 6–18. https://doi.org/10.5539/jedp.v5n1p6

Flavell, J. (1979). Metacognition and Cognitive Monitoring A New Area of Cognitive-Developmental Inquiry A Model of Cognitive Monitoring. *American Psychologist*, *34*(10), 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Flavell, J. (2002). Cognitive development: Children's Knowledge About the Mind. *Annual Review of Psychology*, *50*, 21–45. https://doi.org/10.1146/annurev.psych.50.1.21

Flynn, E. (2007). The role of inhibitory control in false belief understanding. *Infant and Child Development*, *16*, 53–69. https://doi.org/10.1002/icd.500

Flynn, E., O'Malley, C., & Wood, D. (2004). A longitudinal, microgenetic study of the emergence of false belief understanding and inhibition skills. *Developmental Science*, *7*(1), 103–115. https://doi.org/10.1111/j.1467-7687.2004.00326.x

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116–124. https://doi.org/10.3758/BF03195503

Frank, C. K., & Temple, E. (2009). Cultural effects on the neural basis of theory of mind. *Progress in Brain Research*, *178*, 213–223. https://doi.org/10.1016/S0079-6123(09)17815-9

Fujita, N., & Hughes, C. (2021). Mind-mindedness and self-other distinction: Contrasts between Japanese and British mothers' speech samples. *Social Development*, *30*, 57–72. https://doi.org/10.1111/sode.12454

Fujita, N., Devine, R. T., & Hughes, C. (2022). Theory of mind and executive function in early childhood: A cross-cultural investigation. *Cognitive Development*, *61*, 101150. https://doi.org/10.1016/j.cogdev.2021.101150

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in cognitive sciences*, *2*(12), 493–501. https://doi.org/10.1016/S1364-6613(98)01262-5

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition*, *53*(2), 129–153. https://doi.org/10.1016/0010-0277(94)90068-X

Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford University Press.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories*. Mit Press.

Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, *7*(1-2), 145–171. https://doi.org/10.1111/j.1468-0017.1992.tb00202.x

Harris, P. L. (1992). From simulation to folk psychology: The case for development. *Mind and Language*, *7*, 120–144. https://doi.org/10.1111/j.1468-0017.1992.tb00201.x

Hayashi, A., Karasawa, M., & Tobin, J. (2009). The Japanese preschool's pedagogy of feeling: Cultural strategies for supporting young children's emotional development. *Ethos*, *37*(1), 32–49. https://doi.org/10.1111/j.1548-1352.2009.01030.x

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*(5), 349–362. https://doi.org/10.1016/j.tics.2020.02.007

Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, *57*, 567–582. https://doi.org/10.2307/1130337

Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology*, *43*(6), 1447–1459. https://doi.org/10.1037/0012-1649.43.6.1447

Hughes, C., Devine, R. T., & Wang, Z. (2018). Does parental mind-mindedness account for cross-cultural differences in preschoolers' theory of mind? *Child development*, *89*(4), 1296–1310. https://doi.org/10.1111/cdev.12746

Hughes, C., Devine, R. T., Ensor, R., Koyasu, M., Mizokawa, A., & Lecce, S. (2014). Lost in translation? Comparing British, Japanese, and Italian children's theory-of-mind performance. *Child Development Research*, Article ID 893492, 10 pages. https://doi.org/10.1155/2014/893492

Imada, T., Carlson, S. M., & Itakura, S. (2013). East-West cultural differences in context-sensitivity are evident in early childhood. *Developmental Science*, *16*(2), 198–208. https://doi.org/10.1111/desc.12016

Kagitcibasi, C. (2005). Autonomy and relatedness in cultural context: Implications for self and family. *Journal of Cross-Cultural Psychology*, *36*(4), 403–422. https://doi.org/10.1177/0022022105275959

Kälin, S., & Roebers, C. M. (2022). Longitudinal associations between executive functions and metacognitive monitoring in 5–8-year-olds. *Metacognition and Learning*, *17*, 1079–1095. https://doi.org/10.1007/s11409-022-09306-x

Kazama, M., Hirabayashi, H., Karasawa, M., Tardif, T., & Olson, S. (2013). Ambiguous parenting and four-year-olds' understanding of others: A comparison between mothers in Japan and the U.S. *Japanese Journal of Developmental Psychology*, *24*(2), 138–149.

Kim, S., Paulus, M., Sodian, B., & Proust, J. (2016). Young children's sensitivity to their own ignorance in informing others. *PloS One*, *11*(3), e0152595. https://doi.org/10.1371/journal.pone.0152595

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*, 103017. https://doi.org/10.1016/j.concog.2020.103017

Kim, S., Senju, A., Sodian, B., Paulus, M., Itakura, S., Okuno, A., Ueno, M., & Proust, J. (2021). Memory monitoring and control in Japanese and German Preschoolers. *Memory & Cognition*, *51*, 708–717.

Kloo, D., Kristen-Antonow, S., & Sodian, B. (2020). Progressing from an implicit to an explicit false belief understanding: A matter of executive control? *International Journal of Behavioral Development*, *44*(2), 107–115. https://doi.org/10.1177/0165025419850901

Kloo, D., Sodian, B., Kristen-Antonow, S., Kim, S., & Paulus, M. (2021). Knowing minds: Linking early perspective taking and later metacognitive insight. *British Journal of Developmental Psychology*, *39*, 39–53. https://doi.org/10.1111/bjdp.12359

Kobayashi, C., Glover, G. H., & Temple, E. (2006). Cultural and linguistic influence on neural bases of 'Theory of Mind': An fMRI study with Japanese bilinguals. *Brain and Language*, *98*(2), 210–220. https://doi.org/10.1016/j.bandl.2006.04.013

Kochanska, G., Murray, K., & Coy, K. C. (1997). Inhibitory control as a contributor to conscience in childhood: From toddler to early school age. *Child Development*, *68*, 263–277. https://doi.org/10.2307/1131849

Koelkebeck, K., Hirao, K., Kawada, R., Miyata, J., Saze, T., Ubukata, S., Itakura, S., Kanakogi, Y., Ohrmann, P., Bauer, J., Pedersen, A., Sawamoto, N., Fukuyama, H., Takahashi, H., & Murai, T. (2011). Transcultural differences in brain activation patterns during theory of mind (ToM) task performance in Japanese and Caucasian participants. *Social Neuroscience*, *6*(5-6), 615–626. https://doi.org/10.1080/17470919.2011.620763

Kristen, S., Thoermer, C., Hofer, T., Aschersleben, G., & Sodian, B. (2006). Validation of the" Theory of Mind" Scale. *Zeitschrift Fur Entwicklungspsychologie Und Padagogische Psychologie*, *38*(4), 190–199. https://doi.org/10.1026/0049-8637.38.4.186

Lang, B., & Perner, J. (2002). Understanding of intention and false belief and the development of self-control. *British Journal of Developmental Psychology*, *20*(1), 67–76. https://doi.org/10.1348/026151002166325

Lecce, S., Caputi, M., & Pagnin, A. (2014). False-belief understanding at age 5 predicts beliefs about learning in year 3 of primary school. *European Journal of Developmental Psychology*, *12*(1), 40–53. https://doi.org/10.1080/17405629.2014.949665

Lecce, S., Demicheli, P., Zocchi, S., & Palladino, P. (2015). The origins of children's metamemory: The role of theory of mind. *Journal of Experimental Child Psychology*, *131*, 56–72. https://doi.org/10.1016/j.jecp.2014.11.005

Lecce, S., & Hughes, C. (2010). The Italian job?: Comparing theory of mind performance in British and Italian children. *British Journal of Developmental Psychology*, *28*(4), 747–766. https://doi.org/10.1348/026151009X479006

Lewis, C., Koyasu, M., Oh, S., Ogawa, A., Short, B., & Huang, Z. (2009). Culture, executive function, and social understanding. *New Directions for Child and Adolescent Development*, *2009*(123), 69–85. https://doi.org/10.1002/cd.236

Lillard, A. (1998). Ethnopsychologies: Cultural variations in theories of mind. *Psychological Bulletin*, *123*(1), 3–32. https://doi.org/10.1037/0033-2909.123.1.3

Liu, D., Wellman, H. M., Tardif, T., & Sabbagh, M. A. (2008). Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. *Developmental Psychology*, *44*(2), 523. https://doi.org/10.1037/0012-1649.44.2.523

Lockl, K., & Schneider, W. (2006). Precursors of metamemory in young children: The role of theory of mind and metacognitive vocabulary. *Metacognition and Learning*, *1*(1), 15–31. https://doi.org/10.1007/s11409-006-6585-9

Lyons, K. E., & Ghetti, S. (2011). The development of uncertainty monitoring in early childhood. *Child development*, *82*(6), 1778–1787. https://doi.org/10.1111/j.1467-8624.2011.01649.x

Lyons, K. E., & Ghetti, S. (2013). I don't want to pick! Introspection on uncertainty supports early strategic behavior. *Child development*, *84*(2), 726–736. https://doi.org/10.1111/cdev.12004

Marcovitch, S., O'Brien, M., Calkins, S. D., Leerkes, E. M., Weaver, J. M., & Levine, D. W. (2015). A longitudinal assessment of the relation between executive function and theory of mind at 3, 4, and 5 years. *Cognitive Development*, *33*, 40–55. https://doi.org/10.1016/j.cogdev.2014.07.001

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological review*, *98*(2), 224–253. https://doi.org/10.1037/0033-295X.98.2.224

Mizokawa, A. (2020). Japanese and British children's understanding of the social function of pretend crying. *Japanese Psychological Research*, *64*(1), 27–39. https://doi.org/10.1111/jpr.12313

Moriguchi, Y., Evans, A. D., Hiraki, K., Itakura, S., & Lee, K. (2012). Cultural differences in the development of cognitive shifting: East-West comparison. *Journal of Experimental Child Psychology*, *111*(2), 156–163. https://doi.org/10.1016/j.jecp.2011.09.001

Moriguchi, Y., Okumura, Y., Kanakogi, Y., & Itakura, S. (2010). Japanese children's difficulty with false belief understanding: Is it real or apparent? *Psychologia*, *53*(1), 36–43. https://doi.org/10.2117/psysoc.2010.36

Naito, M., & Koyama, K. (2006). The development of false-belief understanding in Japanese children: Delay and difference? *International Journal of Behavioral Development*, *30*(4), 290–304. https://doi.org/10.1177/0165025406063622

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

Ogawa, A., & Koyasu, M. (2008). The relation between components of executive function and theory of mind in young children. *The Japanese Journal of Developmental Psychology*, *19*, 171–182.

Oh, S., & Lewis, C. (2008). Korean re-schoolers' advanced inhibitory control and its relation to other executive skills and mental state understanding. *Child Development*, *79*(1), 80–99. https://doi.org/10.1111/j.1467-8624.2007.01112.x

O'Madagain, C., & Tomasello, M. (2022). Shared intentionality, reason-giving and the evolution of human culture. *Philosophical Transactions of the Royal Society B*, *377*(1843), 20200320. https://doi.org/10.1098/rstb.2020.0320

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, *3*(9), 337–344. https://doi.org/10.1016/S1364-6613(99)01362-5

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526. https://doi.org/10.1017/S0140525X00076512

Reed, M. A., Pien, D. L., & Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill-Palmer Quarterly*, *30*(2), 131–147.

Rakoczy, H. (2022). Foundations of theory of mind and its development in early childhood. *Nature Reviews Psychology*, *1*(4), 223–235. https://doi.org/10.1038/s44159-022-00037-z

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14-and 18-month-olds. *Developmental Psychology*, *33*(1), 12–21. https://doi.org/10.1037/0012-1649.33.1.12

Rohwer, M., Kloo, D., & Perner, J. (2012). Escape From Metaignorance: How Children Develop an Understanding of Their Own Lack of Knowledge. *Child Development*, *83*(6), 1869–1883. https://doi.org/10.1111/j.1467–8624.2012.01830.x

Ross, J., Yilmaz, M., Dale, R., Cassidy, R., Yildirim, I., & Suzanne Zeedyk, M. (2017). Cultural differences in self-recognition: The early development of autonomous and related selves? *Developmental Science*, *20*(3), e12387. https://doi.org/10.1111/desc.12387

Russell, J., Hill, E. L., & Franco, F. (2001). The role of belief veracity in understanding intentions-in-action: Preschool children's performance on the transparent intentions task. *Cognitive Development*, *16*(3), 775–792. https://doi.org/10.1016/S0885-2014(01)00057-0

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and US preschoolers. *Psychological science*, *17*(1), 74–81. https://doi.org/10.1111/j.1467-9280.2005.01667.x

Salmon, A. K., & Lucas, T. (2011). Exploring young children's conceptions about thinking. *Journal of Research in Childhood Education*, *25*(4), 364–375. https://doi.org/10.1080/02568543.2011.605206

Senzaki, S., Wiebe, S. A., Masuda, T., & Shimizu, Y. (2018). A cross-cultural examination of selective attention in Canada and Japan: The role of social context. *Cognitive Development*, *48*, 32–41. https://doi.org/10.1016/j.cogdev.2018.06.005

Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental Psychology*, *47*(5), 1239. https://doi.org/10.1037/a0023899

Singh, L., Cristia, A., Karasik, L. B., & Rajendra, S. J. (2023). Diversity and representation in infant research: Barriers and bridges toward a globalized science of infant development. *Infancy*, *28*(4), 708–737. https://doi.org/10.1111/infa.12545

Symeonidou, M., Doherty, M., & Ross, J. (2022). Thinking about thinking: A longitudinal investigation of the developmental link between metacognition, mind-reading and self-control. https://psyarxiv.com/prs2b

Taumoepeau, M., Sadeghi, S., & Nobilo, A. (2019). Cross-cultural differences in children's theory of mind in Iran and New Zealand: The role of caregiver mental state talk. *Cognitive Development*, *51*, 32–45. https://doi.org/10.1016/j.cogdev.2019.05.004

van der Plas, E., Zhang, S., Dong, K., Bang, D., Li, J., Wright, N. D., & Fleming, S. M. (2022). Identifying cultural differences in metacognition. *Journal of Experimental Psychology: General*, *151*(12), 3268–3280. https://doi.org/10.1037/xge0001209

Wang, Z., Devine, R. T., Wong, K. K., & Hughes, C. (2016). Theory of mind and executive function during middle childhood across cultures. *Journal of Experimental Child Psychology*, *149*, 6–22. https://doi.org/10.1016/j.jecp.2015.09.028

Wellman, H. M., & Liu, D. (2004). Scaling of theory of mind tasks. *Child Development*, *75*(2), 523–541. https://doi.org/10.1111/j.1467-8624.2004.00691.x

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta analysis of theory of mind development: The truth about false belief. *Child Development*, *72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304

Wellman, H. M., Fang, F., & Peterson, C. C. (2011). Sequential progressions in a theory of mind scale: Longitudinal perspectives. *Child Development*, *82*(3), 780–792. https://doi.org/10.1111/j.1467-8624.2011.01583.x

Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychological Science*, *17*(12), 1075–1081. https://doi.org/10.1111/j.1467-9280.2006.01830.x

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Wu, S., & Keysar, B. (2007). The effect of culture on perspective taking. *Psychological Science*, *18*(7), 600–606. https://doi.org/10.1111/j.1467-9280.2007.01946.x

Zhao, X., Wente, A., Flecha, M. F., Galvan, D. S., Gopnik, A., & Kushnir, T. (2021). Culture moderates the relationship between self-control ability and free will beliefs in hildhood. *Cognition*, *210*, 104609. https://doi.org/10.1016/j.cognition.2021.104609

How to cite this article: Symeonidou, M., Mizokawa, A., Kabaya, S., Doherty, M. J., & Ross, J. (2023). Contrasting one's share of the shared life space: Comparing the roles of metacognition and inhibitory control in the development of theory of mind among Scottish and Japanese children. *Developmental Science*, e13417. https://doi.org/10.1111/desc.13417

Appendix A: Theory of mind task battery

Diverse Belief task (1 point). In this task children were shown two pictures representing hiding places (for example, a bush and a tree) and were told that an animal for example, a cat) was hiding behind one of them (the hiding locations and animals were varied across phases to avoid practise effects). They were then asked to point to the picture they thought the cat was hiding behind. Once they did this, a puppet called Teddy joined the game. When the experimenter asked Teddy where he thought the cat was, Teddy would always respond with the opposite location to the one selected by the child. The experimenter then asked "Teddy thinks the cat is hiding in the tree, so where will Teddy look for the cat?". This task assesses the child's ability to predict how they and others will act in accordance with their beliefs.

Knowledge Access task (2 points). In this task children were presented with a small plastic box with a small toy figure hidden inside (the appearance of the box and its contents varied across phases to avoid practise effects). The experimenter asked the child if they knew what was in the box (target answer: no) and then, how they might find out (target answer: look inside). The box was then opened and the toy figure was revealed. The figure was then put inside the box and the lid placed back to hide the box's contents once again. At this point, Teddy returned to play the game. The child was then asked whether Teddy knew what was inside the box (target answer: No), and then, how he might find out (target answer: look inside/ask). This task assess the child's understanding of knowledge-ignorance, and acting on this basis.

Unexpected Transfer Task (1 point). This task involved participation by the experimenter, the child and Teddy. At the start of this game the child was shown two different sized and coloured boxes. One of these boxes belonged to Teddy and the other to the experimenter. The child was then shown a marble and told that it was Teddy's favourite toy, the child was told that Teddy was going to go away to play for a short moment, so he was going to put his marble in his box for safekeeping. Teddy was then placed out of the child's view, and while he was away the experimenter encouraged child to play a trick on Teddy. With the child's engagement the experimenter moved the marble from Teddy's box to their own box. At this stage the child was asked control questions about the current location of the toy ("where is the marble now?"), followed by a question relating to Teddy's initial action ("Where did Teddy put the marble in the beginning?") in order to establish that the child understood a change took place. Once these control questions were completed, Teddy returned to the game, the experimenter asked the child the test question "Where will Teddy look for his marble first?" [target: where he left it].

Unexpected contents task (2 points). To instil confidence in their guesses (especially given repeated testing across phases) children started the task with a control condition, in which they were shown canonical containers (for example, egg boxes, DVD cases, branded sweets) and asked what they thought might be inside. The contents were then revealed to be as expected. Where treats were revealed and parental permission was in place, the child was given a treat. Mirroring test trials, the contents were then replaced in the containers, and the child asked what they had thought was inside. However, in the test trials, the child was shown a branded packet of treats (treat varied across phases) and asked what they thought might be inside the box. The experimenter then excitedly opened the packet to find out if there was really a treat inside, only to express disappointment on revealing that the packet contained something unattractive ("Those aren't sweets! Those are pencil shavings. Boo!!"). The package was then closed up, and Teddy joined the game, indicating excitement upon seeing the treat container. The child was asked what Teddy thought was in the container, scoring a point if they thought he would guess in line with the packaging. Teddy then excitedly opened the box, only to express disappointment upon finding the unattractive contents inside. The child was then asked how Teddy felt before [target: happy] and after [target:

sad], scoring a point if they noted the change in emotion. Children were shown a drawing of a happy face and a sad face for reference.

Hidden Emotions Task (1 point) This task involved the child listening to a simple story accompanied by simple drawings about a boy called Matt receiving a disappointing present from his aunt. In the story Matt received a present from his aunt which he falsely thought would be a train set (for which he was excited for), only to then discover that it was actually a pair of "boring grey socks" making him feel sad. However, the child was told that Matt did not want to upset his aunt so he tried to hide how he really felt. Once the story was over the child was asked what present Matt received from his aunt (control), this was followed by test questions "How did Matt feel when he got a present he didn't like?" [target: sad] and "How did he try to look?" [target: happy], scoring a point if they noted the mismatch. Children were shown a drawing of a happy face and a sad face for reference.

Transparent Intentions task (2 points). This task involved a set of line drawings presented on transparent acetate slides. When the slides overlapped, they appeared to create one image (for example, a mouse or a robot), with a missing feature (drawings were varied across phases to avoid practise effects). The experimenter pointed out the missing feature (for example, the mouse had no tail, the robot had no ear). On self-focused trials, the child was invited to add the missing feature to the drawing ("Can you draw the mouse a tail? *or* Can you draw the robot an ear?); on other-focused trials, the experimenter stated this intention ("I will draw the mouse a tail" or "I will draw the robot an ear") and proceeded to add the feature to the drawing. However, when the top acetate was lifted from the bottom, the picture was transformed (for example, the mouse's body, removed from context, became the shape of a balloon, the robots head, removed from context became the shape of a mug). This meant that the added feature changed interpretation (for example, the 'tail' looked like the string of the balloon, the 'ear' looked like the handle of a mug). On self-focused trials, the child was asked "Did you think you were drawing a ["tail on a mouse"] or ["a balloon string"]?"; Did you mean to draw a ["tail on a mouse"] or a ["balloon string"]?" (order counterbalanced). On other focused trials the questions were phrased to reference the experimenter i.e. "Did I mean…" and "Did I think…"? This task allows children to demonstrate an understanding of false intentions (what they or the experimenter meant to draw) and false thoughts (what they or the experimenter thought they were drawing). In this task, self- and other-focused questions were closely matched in context, and could be counterbalanced for order and drawing content.

Appendix B: Preliminary Analyses on Sequencing of Theory of Mind

Table A shows the number and proportion of children who passed each ToM task, split by culture and age group. Scottish children showed a similar sequence of passing as observed for other western samples. More Scottish children passed diverse belief, followed by knowledge access, false belief (at least for the content task), and more complex tasks (false intentions, false thoughts, false emotions, hidden emotions). Japanese children showed the same sequence. Focusing on previous areas of cultural difference, pass rates for diverse desires where exactly equivalent, and a Pearson's chi square test confirmed that pass rates for the ignorance, χ^2 (1) = .169, p = .837, and seeing equals knowing component, χ^2 (1) = 3.187, p = .056, of the knowledge access task were statistically equivalent across cultures.

Extracting Wellman and Liu's (2004) original test questions, Table B shows individual sequences of performance, split by culture. 77% of Scottish children, and 75% of Japanese children showed a pattern which fit the established developmental sequence for western

TABLE A1 Proportion of Scottish and Japanese children passing each ToM task

Task		Japan	Scotland
Diverse belief	Overall ∝	.88	.88
	∝ Younger	.81	.80
	∝ Older	.93	.96
Knowledge Access –belief	Overall ∝	.71	.68
	∝ Younger	.59	.50
	∝ Older	.82	.88
Knowledge Access –practical	Overall ∝	.57	.73
	∝ Younger	.41	.60
	∝ Older	.72	.88
Unexpected contents – false belief	Overall ∝	.61	.66
	∝ Younger	.59	.46
	∝ Older	.62	.88
Unexpected contents – false emotion	Overall ∝	.36	.39
	∝ Younger	.37	.10
	∝ Older	.35	.73
Unexpected transfer – false belief	Overall ∝	.64	.75
	∝ Younger	.41	.63
	∝ Older	.86	.88
Transparent Intentions – mean	Overall ∝	.57	.61
	∝ Younger	.52	.50
	∝ Older	.59	.73
Transparent Intentions – think	Overall ∝	.55	.50
	∝ Younger	.56	.33
	∝ Older	.59	.69
Hidden emotions	Overall ∝	.39	.43
	∝ Younger	.26	.30
	∝ Older	.52	.57

TABLE B1 Pass sequences for Wellman & Liu's (2004) scale, split by culture

	Pass Sequence					
	1	2	3	4	5	Other
Diverse belief	–	+	+	+	+	
Knowledge Access –ignorance	–	–	+	+	+	
False belief – contents	–	–	–	+	+	
Hidden emotions	–	–	–	–	+	
Japan	2	6	8	15	11	14
Scotland	0	8	7	10	18	13

+ denotes pass, – denotes fail

children (diverse belief, knowledge access, false belief, hidden emotion). The number of children who fit the scale did not vary by culture, χ^2 (1) = .049, p = 1.000, and Green's (1956) coefficient of replicability suggested that the original scaling provided a good (defined as >.90) and equal fit for both Japan, Rep = .97, and Scotland, Rep = .97. Only 4 Japanese children showed evidence of the alternative sequence suggested for interdependent cultures, passing only the knowledge access task (1), or the knowledge access task plus some more complex tasks prior to the diverse desires task (3). The same number of Scottish children in the 'other' category passed only the knowledge access task (3), or the knowledge access task plus a more complex task prior to passing diverse desires (1). Thus, we found no evidence of cultural variation from the established ToM sequence in our sample.