COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 45 (2021) e12966

ISSN: 1551-6709 online
DOI: 10.1111/cogs.12966

The Development of Cognitive Reflection in China

Tianwei Gong,[a] Andrew G. Young,[b] Andrew Shtulman[c]

[a]*Faculty of Psychology, Beijing Normal University*
[b]*Department of Psychology, Northeastern Illinois University*
[c]*Department of Psychology, Occidental College*

Received 31 March 2020; received in revised form 3 February 2021; accepted 22 February 2021

Abstract

Cognitive reflection is the tendency to override an intuitive response so as to engage in the reflection necessary to derive a correct response. Here, we examine the emergence of cognitive reflection in a culture that values nonanalytic thinking styles, Chinese culture. We administered a child-friendly version of the cognitive reflection test, the CRT-D, to 130 adults and 111 school-age children in China and compared performance on the CRT-D to several measures of rational thinking (belief bias syllogisms, base rate sensitivity, denominator neglect, and other-side thinking) and normative thinking dispositions (actively open-minded thinking and need for cognition). The CRT-D was a significant predictor of rational thinking and normative thinking dispositions in both children and adults, as previously found in American samples. Adults' performance on the CRT-D correlated with their performance on the original CRT, and children's performance on the CRT-D predicted rational thinking and normative thinking dispositions even after adjusting for age. These results demonstrate that cognitive reflection, rational thinking, and normative thinking dispositions converge even in a culture that emphasizes holistic, nonanalytic reasoning.

Keywords: Cognitive development; Cognitive reflection; Cross-cultural cognition; Rational thought; Thinking dispositions

Individual differences in rationality are pervasive and robust (Kahneman, 2011; Sloman, 1996; Stanovich & West, 2000). While some people routinely engage in reflective, analytic thought, others are more likely to rely on immediate intuition. The most widely used measure of adults' reliance on analytic versus intuitive thought is the cognitive reflection test (CRT; Frederick, 2005). This three-item measure probes adults' tendency to reflect on, and override, an inaccurate gut response. Consider the well-known bat-and-ball problem: "A bat and a ball

Tianwei Gong is now at the Department of Psychology, University of Edinburgh.
Correspondence should be sent to Andrew Shtulman, Department of Psychology, Occidental College, 1600 Campus Road, Los Angeles, CA 91104. Email: shtulman@oxy.edu

Highlights

- How does cognitive reflection, or the tendency to privilege analytic responses over intuitive ones, develop in a culture that values nonanalytic thinking styles?
- A child-friendly version of the cognitive reflection test (the CRT-D) was administered to Chinese children and adults, along with several heuristics-and-biases tasks.
- The CRT-D was a significant predictor of rational thinking and normative thinking dispositions in both children and adults, as previously found in American samples.
- Cognitive reflection, rational thinking, and normative thinking dispositions converge even in a culture that emphasizes holistic reasoning.

cost $1.10 in total. The bat costs $1 more than ball. How much does the ball cost?" If we subtract the two numbers given, we arrive at an answer of 10 cents, but the correct answer is 5 cents (because a 5-cent ball costs a dollar less than a $1.05 bat, and together they cost $1.10, as stipulated). Adults who provide the correct answer avoid the error of simple subtraction and generate a correct response through reflection.

Adults who perform well on the CRT demonstrate superior performance on many other tasks as well, including logical reasoning, probabilistic reasoning, argumentation, and temporal discounting (Frederick, 2005; Toplak, West, & Stanovich, 2011). Toplak et al. (2011) administered 15 different heuristics-and-biases tasks to college-educated adults and found that the CRT was a stronger predictor of performance than either general intelligence or executive functioning. The CRT also correlates with normative thinking dispositions, including the need for cognition, open-minded thinking, and lack of superstitious thinking (Toplak, West, & Stanovich, 2014a), as well as conceptual reasoning (Pennycook, Fugelsang, & Koehler, 2015a), including causal reasoning (Don, Goldwater, Otto, & Livesey, 2016), moral reasoning (Royzman, Landy, & Leeman, 2015), and science understanding (Gervais, 2015; Shtulman & McCallum, 2014).

Despite its centrality to rational thought, cognitive reflection has been studied mainly as an individual difference variable among Western adults. No research has explored its developmental origins or its stability across cultures. Here, we seek to address these gaps by studying cognitive reflection (a) in children and (b) in a culture that often emphasizes holistic reasoning over analytic reasoning, Chinese culture (Nisbett, Peng, Choi, & Norenzayan, 2001; Norenzayan, Smith, Kim, & Nisbett, 2002). Our approach was to assess whether cognitive reflection predicts two forms of higher-order cognition: rational thinking and normative thinking dispositions.

Rational thinking is conforming to normative models of decision making or problem solving, yielding outcomes that maximize accuracy or utility (Stanovich & West, 2000). Rational thinking is often identified with "Type 2" or "System 2" reasoning, characterized by slow and effortful processes (Evans, 2008), and it is traditionally measured with tasks that pit normatively valid solution strategies against information-processing shortcuts, that is, heuristics and

biases. Consider the task of deciding whether the conclusion of a syllogism follows from its premises. The normative approach is to focus on the syllogism's structure independent of its content, but the heuristic approach is to focus on content, namely, whether the conclusion is empirically true. Given the premises (a) all birds can fly and (b) penguins are birds, it logically follows that penguins can fly, but many people maintain that this conclusion does not follow, privileging content over logic (Ball & Thompson, 2018; Evans, Barston, & Pollard, 1983).

Normative thinking dispositions are cognitive strategies or styles associated with accurate or well-justified outcomes. They include actively open-minded thinking, or the willingness to engage with new evidence and alternative viewpoints (Stanovich & West, 1997) and need for cognition, or the tendency to engage in and enjoy cognitive activities (Cacioppo & Petty, 1982). These dispositions are assessed with self-report scales, in which participants rate their agreement with statements, such as "People should take into consideration evidence that goes against their beliefs" (characteristic of actively open-minded thinking) or "I prefer complex to simple problems" (characteristic of need for cognition).

Cognitive reflection is a distinct yet related disposition. It is the tendency to privilege deliberation over intuition, and it is measured with brainteasers that elicit an erroneous "gut" response that can be corrected upon further reflection. To succeed on these brainteasers, a person must (a) detect the conflict between their intuitive response and the question it is meant to answer, (b) inhibit the intuitive response, (c) reanalyze the question, and (d) derive a more accurate response (Pennycook, Fugelsang, & Koehler, 2015b; Travers, Rolison, & Feeney, 2016). While it is possible to provide a correct response without first considering and rejecting an intuitive response (Bago & De Neys, 2019), intuitive responses are the default response in most populations (Frederick, 2005), rendering the cognitive reflection task a better measure of reflective thinking than intuitive thinking (Pennycook, Cheyne, Koehler, & Fugelsang, 2016).

In short, cognitive reflection is the disposition to suppress intuitive reasoning in favor of analytic reasoning, and it predicts performance on a wide range of tasks that diagnose rational thinking (Frederick, 2005; Pennycook et al., 2015a; Toplak et al., 2011) and reliance on normative thinking dispositions (Pennycook et al., 2015a; Toplak et al., 2014a, 2014b). Cognitive reflection also actively promotes rational thought; individuals who are primed to think reflectively, by taking the CRT, provide more analytic responses on subsequent reasoning tasks (Paxton, Unger, & Greene, 2012; Pinillos, Smith, Nair, Marchetto, & Mun, 2011).

Our motivation for investigating the development of cognitive reflection in China are both methodological and theoretical. Methodologically, we seek to validate the use of a child-friendly version of the cognitive reflection test—the CRT-D—beyond Western culture. This test, developed by Young and colleagues (Young, Powers, Pilgrim, & Shtulman, 2018; Young & Shtulman, 2020a), consists of brain teasers designed to elicit wrong answers that even children are able to correct upon further reflection. A sample item is "What do cows drink?" The intuitive response is milk, but the correct response is water. While several measures of rational thinking have been shown to operate similarly across cultures, including the original CRT (Gervais et al., 2018; Peng et al., 2019), these measures rely heavily on mathematics (Byrd & Conway, 2019; Cokely & Kelley, 2009). The CRT-D draws more on verbal

knowledge than mathematical knowledge, so cross-cultural samples are required to examine its validity as a general test of analytic reasoning, preferably samples that vary in age and thus exposure to non-Western thinking styles.

Theoretically, we seek to determine whether cognitive reflection emerges in tandem with the capacities and dispositions it predicts in adulthood. Cognitively reflective adults are less likely to rely on heuristics and biases and more likely to endorse normative thinking dispositions, as noted above (Stanovich, West, & Toplak, 2016; Toplak et al., 2014b). These relations have not been studied in non-Western samples, let alone from a developmental perspective. China, as an Eastern culture, embraces a holistic style of thinking and knowing. While Westerners tend to decompose objects and events into discrete attributes, organizing the world by categories and rules, Easterners pay more attention to the surrounding context, organizing the world by relationships and similarities (Nisbett et al., 2001). The apparent conflict between holistic and analytic reasoning implies that individuals who value holistic reasoning may not perform as well, or as consistently, on the kinds of analytic tasks traditionally associated with rational thought. Likewise, individuals who excel at holistic reasoning may draw upon different competencies than those that tapped by the CRT—competencies like conflict detection, response inhibition, and systematic deliberation. On the other hand, holistic reasoning may require as much cognitive reflection as analytic reasoning. Tracking relationships, comparing and contrasting objects, and attending to the global context are resource-demanding tasks (Gentner, 2010; Richland & Simms, 2015), which may require the same reflective capacities as required by heuristics-and-biases tasks and measured by thinking disposition scales.

Consistent with the idea that Chinese culture emphasizes holistic reasoning over analytic reasoning, Chinese teachers have been observed to emphasize concrete, knowledge-based forms of reasoning over abstract, logical reasoning (Tweed & Lehman, 2002), and Chinese students report an inclination toward intuition rather than analysis (Buchtel & Norenzayan, 2008). Chinese adults are also particularly susceptible to several cognitive biases, including belief bias (Norenzayan et al., 2002), my-side thinking (Choi, Choi, & Norenzayan, 2004), hindsight (Pohl, Bender, & Lachmann, 2002), and overconfidence (Yates, Lee, Sieck, Choi, & Price, 2002). On the other hand, Chinese children outperform American children on tests of mathematical reasoning (Siegler & Mu, 2008; Wang & Lin, 2009), scientific reasoning (Bao et al., 2009), and analogical reasoning (Carstensen et al., 2019), which implies that their early analytic abilities are comparable to those of American children, if not superior. This mixed set of results leaves open the possibility that cognitive reflection develops similarly in Chinese and American cultures and exerts the same predictive power, even if cultural norms support some forms of reasoning more than others.

In the present study, we assessed the coherence of cognitive reflection, rational thinking, and normative thinking dispositions across cultures. We modeled our study on previous research by Kokis, Macpherson, Toplak, West, and Stanovich (2002) and Toplak et al. (2014b), who took tasks and scales from the adult literature on rational thought and adapted them for use with children. The rational thinking tasks include denominator neglect, which measures the ability to compare desired outcomes to total possible outcomes; belief bias syllogisms, which measures the prioritization of logic over prior beliefs; base rate sensitivity,

which measures a preference for statistical trends over anecdotes; and other-side thinking, which measures the ability to reason from another person's perspective. The normative thinking dispositions included child-friendly versions of the two scales mentioned previously, Need for Cognition and Actively Open-minded Thinking. Details on all tasks and scales are provided below.

These measures span a variety of reasoning contexts, from logical reasoning to probabilistic reasoning to argumentation, and cognitive reflection has been shown to predict success on all such measures in U.S. adults (Pennycook et al., 2016; Toplak et al., 2011). Here, we test whether the predictive power of cognitive reflection extends beyond the cultural context in which these measure were originally developed. Precedent for this approach comes from the cross-cultural study of executive function. Researchers have found that Chinese children surpass American children on tests of inhibition, working memory, and attentional control, but these skills predict social reasoning (Sabbagh, Xu, Carlson, Moses, & Lee, 2006) and academic achievement (Lan, Legare, Ponitz, Li, & Morrison, 2011) to similar extents in both cultures, suggesting that executive function skills facilitate conceptual development regardless of overall levels of executive functioning.

In sum, we investigated the development of cognitive reflection in China, from children to adults, by assessing whether a child-friendly test of cognitive reflection (the CRT-D) predicts rational thinking and normative thinking dispositions as it does in U.S. samples. Our cross-sectional design provides a particularly strong test of the relation between cognitive reflection and rational thought because it allows us to determine whether this relation emerges early in development or is forged later, through culturally prescribed habits of mind.

1. Method

1.1. Participants

One-hundred and thirty adults were recruited from an open online participant pool established by researchers from the Chinese Academy of Sciences (M age = 20.78 years, SD = 3.49 years, 89 females). Nineteen additional adults were excluded because they responded incorrectly to at least one of two attention checks.

One-hundred and eleven 5- to 12-year-old children were recruited from public playgrounds and completed the study onsite (M age = 9 years, 5 months, SD = 1 year, 11 months, 57 females). The playgrounds were located in several regions across China: North (Hebei, Beijing, and Jilin, $n = 35$), North Central (Shanxi, $n = 29$), Southeastern (Fujian, $n = 26$), and West Central (Sichuan, $n = 21$). Nine additional children were tested but removed from the dataset due to inattention (four), experimenter error (one), or parent coaching (four). The age distribution for the final sample of children is shown in Fig. 1. Regional differences in task performance are presented in the Supplementary Materials. While overall performance did vary, there were no interactions between region and age for any task (all $F < 2.45$, $p > .05$), indicating that the developmental patterns observed for the sample as a whole were also observed in each region.

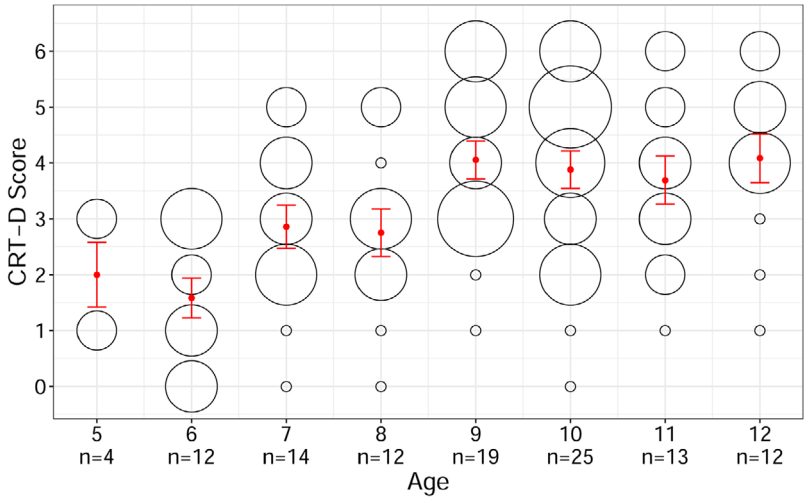

Fig. 1. Distribution of CRT-D scores by age (in years). Bubble size is proportional to the number of participants. Red dots represent means. Error bars indicate standard deviations.

The focus of the current paper is on the developmental relations between cognitive reflection, rational thinking, and normative thinking dispositions in a Chinese sample. These relations have been observed in a U.S. sample (Young et al., 2018), and we use the data from that study to determine whether the findings from the present study are comparable. The U.S. sample ($n = 96$) was recruited from public parks, similar to the Chinese sample, and was administered the same tasks using the same protocol. They ranged in age from 5 to 12, with a mean age of 8 years, 1 month ($SD = 2$ years, 2 months; 49% female). Task performance for the U.S. sample is summarized in the Supplementary Materials.

1.2. Materials

The original CRT (Frederick, 2005) has been translated into Chinese by Gervais et al. (2018), and we borrowed their translation. All other materials were translated into Chinese and then back-translated into English by two Chinese natives proficient in both Chinese and English. To ensure translation accuracy, the final Chinese version was improved by comparing the original materials to the back-translated materials.

Children and adults completed the same measures of rational thinking but slightly different measures of cognitive reflection and normative thinking dispositions. Adults completed three measures of cognitive reflection (CRT, CRT-ALT, and CRT-D), whereas children completed only the CRT-D, and adults completed the original versions of the Need for Cognition Scale and the Actively Open-Minded Thinking Scale, whereas children completed child-friendly versions. The exact tasks are described below.

For our scales, we report Cronbach's alpha and McDonald's ω_{Total} separately for each age group. McDonald's ω_{Total} is a less biased estimate of reliability than Cronbach's alpha in most circumstances, and equivalent to Cronbach's alpha when the latter's often unrealistic assump-

tions about unidimensionality are met (Zinbarg, Revelle, Yovel, & Li, 2005). Some scales exhibited a low Cronbach's alpha but a satisfactory McDonald's ω_{Total}. We retained these scales in their original form, without dropping items, to facilitate comparison with Western samples, though the lower-than-ideal reliabilities suggest that these scales may need adjustment in future research with Chinese samples.

1.3. Cognitive reflection measures

1.3.1. Cognitive reflection test

The original CRT consists of three questions: the bat-and-ball question mentioned above, the lily pad question ("In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake?," intuitive answer: 24, correct answer: 47), and the widget question ("If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?," intuitive answer: 100, correct answer: 5). For all three items, we used the number of correct responses as the final score, with higher scores indicating greater cognitive reflection. This measure was administered only to adults. McDonald's ω_{Total} for the measure was .58 and Cronbach's alpha was .57.

1.3.2. Cognitive reflection test—alternative version (CRT-ALT)

Another six items were taken from extended versions of the CRT developed by Primi, Morsanyi, Chiesi, Donati, and Hamilton (2015) and Stanovich et al. (2016). A sample item is "Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class?" (the intuitive answer is 30 but the correct answer is 29). This measure was administered only to adults. McDonald's ω_{Total} for the measure was .55 and Cronbach's alpha was .36.

1.3.3. Cognitive reflection test—developmental version (CRT-D)

Children's cognitive reflection was measured using the items developed by Young et al. (2018). The original, English-language test contained eight items, which can be found in Table 1, along with their Chinese translations. These items are brain teasers, taken from alternative versions of the CRT (Thomson & Oppenheimer, 2016) or found on the Internet. One item was removed in the present study because it related to Christmas, and Chinese children are not familiar with that holiday. McDonald's ω_{Total} for the measure was .48 in adults and .68 in children. Cronbach's alpha was .31 in adults and .56 in children.

1.4. Rational thinking measures

1.4.1. Denominator neglect

This bias entails focusing on favored outcomes (the numerator in a probability estimate) and neglecting all possible outcomes (the denominator). A child-friendly version was adopted from Kokis et al. (2002). Participants were shown trays of black and white marbles and told that the goal was to blindly pick a black marble after the marbles were scrambled. They were then asked to choose between a smaller tray with fewer black marbles but a higher

Table 1
CRT-D questions and the percentage of Chinese adults and Chinese children who selected each type of response (C: correct, I: intuitive, O: other)

			Adults			Children		
Question	Correct answer	Intuitive answer	C	I	O	C	I	O
1. If you're running a race and you pass the person in second place, what place are you in? (如果在跑步比赛中, 你超过了第二名, 那么现在你是第几名?)	Second (第二名)	First (第一名)	62.3	36.2	1.5	17.1	77.5	5.4
2. Emily's father has three daughters. The first two are named Monday and Tuesday. What is the third daughter's name? (莉莉的爸爸有三个女儿。前两个分别叫星期一和星期二。第三个女儿叫什么?)	Emily (莉莉)	Wednesday (星期三)	75.4	22.3	2.3	19.8	75.7	4.5
3. A farmer has 5 sheep, all but 3 run away. How many are left? (一个农民有 5 只羊, 除了 3 只以外都跑了。还剩几只?)	Three (三只)	Two (两只)	92.3	6.2	1.5	42.3	45.0	12.6
4. If there are 3 apples and you take away 2, how many do you have? (我有 3 个苹果, 你从我这里拿走了两个, 你有几个?)	Two (两个)	One (一个)	89.2	3.8	6.9	61.3	30.6	8.1
5. What do cows drink? (奶牛喝什么?)	Water (水)	Milk (奶)	88.5	9.2	2.3	73.9	22.5	3.6
6. What weighs more, a pound of rocks or a pound of feathers? (一斤的石头和一斤的羽毛, 哪个更重?)	Same Weight (一样重)	Rocks (石头)	97.7	1.5	0.8	52.3	43.3	4.5
7. What hatches from a butterfly egg? (从蝴蝶的卵里孵出来的是什么?)	Caterpillar (毛毛虫)	Baby Butterfly (小蝴蝶)	67.7	29.2	3.1	67.6	28.8	3.6
[Not Included] Who makes Christmas presents at the North Pole? (是谁在北极制作圣诞礼物?)	Elves (精灵)	Santa (圣诞老人)	—	—	—	—	—	—

probability of winning, or a larger tray with more black marbles but a lower probability of winning (e.g., 1:10 vs. 9:100). Participants responded to nine items of varying ratios. For three items, the larger tray also had a higher probability of winning. We used the number of highest probability selections as the final score, with higher scores indicating more resistance to denominator neglect.

1.4.2. Belief bias syllogisms

A child-friendly assessment of belief bias was adopted from Toplak et al. (2014b). Belief bias is endorsing arguments with true conclusions or rejecting arguments with false conclusions regardless of whether those conclusions follow from the arguments' premises. Participants evaluated the logical validity of eight syllogisms consisting of either an invalid argument paired with a believable conclusion (e.g., "All vegetables have vitamins; Carrots have vitamins; If we pretend the sentences are true, is it certain that carrots are vegetables?") or a valid argument paired with an unbelievable conclusion (e.g., "All birds can fly; Penguins are birds; If we pretend the sentences are true, is it certain that penguins can fly?"). We used the total number of correct responses as the final score.

1.4.3. Base rate sensitivity

A child-friendly version of base rate sensitivity was adopted from Kokis et al. (2002) and Toplak et al. (2014b). Participants evaluated five scenarios in which probabilistic base rate information conflicted with concrete, personal information. For example, "Anna wants to ride the scariest roller coaster at the amusement park. A survey of hundreds of kids in magazines said that the LizardLoop is scarier than the TurboFlip. While Anna is having lunch, she heard someone say that the TurboFlip is much scarier than the LizardLoop. Which roller coaster is most likely the scariest ride?" The names of the people and places were adapted to match Chinese customs. We used the total number of correct responses (probabilistic base rate choices) as the final score, following Toplak et al. (2014b).

1.4.4. Other-side thinking

Participants completed a standard other-side thinking task with a child-appropriate topic (Toplak et al., 2014b). Participants were asked to state their position on the issue of whether kids should have cell phones. Participants were then asked to give reasons for and against their position. Participants were encouraged to give as many reasons as possible. The key measure was the number of conceptually unique reasons a participant provided against their endorsed position.

1.4.5. Composite score

A rational thinking composite was created by averaging z scores from the four tasks above: denominator neglect, belief bias syllogism, base rate sensitivity, and other-side thinking (Toplak et al., 2014b; Young et al., 2018). Correlations among our measures of rational thinking were lower than expected (Tables 2 and 3). The low correlations among adults reflect the fact that most adults performed near ceiling. The low correlations among children, on the other hand, may be due to complications in adapting tasks designed for

Table 2
Descriptive statistics and correlations among variables for Chinese adults ($n = 130$)

	Mean	SD	1	2	3	4	5	6	7	8	9	10
1. CRT-D	5.73	1.13	—									
2. CRT	2.38	0.89	.35***	—								
3. CRT-Alt	4.72	1.18	.34***	.51***	—							
4. Denominator neglect	6.92	2.29	.17	.15	.09	—						
5. Belief bias syllogisms	6.09	2.15	.32***	.53***	.39***	.16	—					
6. Base rate sensitivity	3.29	1.22	.28**	.13	.16	.29***	.16	—				
7. Other-side thinking	3.42	2.13	.06	.08	.11	.01	.17	-.07	—			
8. Rational thinking composite	0.00	0.58	.35***	.38***	.32***	.63***	.64***	.59***	.47***	—		
9. Need for cognition	2.78	0.40	.09	.01	.06	.15	0.06	.10	.12	.19*	—	
10. Actively open-minded thinking	3.05	0.34	.23**	.24**	.13	.24**	.35***	.11	.20*	.39***	.30***	—
11. Normative thinking composite	0.00	0.81	.20*	.15	.12	.24**	.26**	.13	.20*	.36***	.81***	.81***

Note. Variables 1–3 measure cognitive reflection; variables 4–7 measure rational thinking (variable 8 is their composite); and variables 9–10 measure normative thinking (variable 11 is their composite).

*$p < .05$.

**$p < .01$.

***$p < .001$.

Table 3
Descriptive statistics and correlations among variables for Chinese children ($n = 111$)

	Mean	SD	1	2	3	4	5	6	7	8	9
1. CRT-D	3.34	1.68	—								
2. Denominator neglect	5.95	1.91	.21*	—							
3. Belief bias syllogisms	3.47	1.88	.40***	.10	—						
4. Base rate sensitivity	2.76	1.21	.20*	.11	.22*	—					
5. Other-side thinking	2.06	1.39	.09	.11	.29**	.10	—				
6. Rational thinking composite	0.00	0.61	.37***	.55***	.67***	.59***	.62***	—			
7. Need for cognition	3.09	0.42	.14	.14	.35***	.13	.19*	.33**	—		
8. Actively open-minded thinking	2.88	0.36	.20*	.21*	.07	.02	.06	.15	.21*	—	
9. Normative thinking composite	0.00	0.78	.23*	.22*	.28**	.11	.16	.32**	.78***	.78***	—
10. Age (Month)	113.18	23.69	.45***	.25**	.21*	.23*	.27**	.40***	-.13	.20*	.05

Note. Variable 1 measures cognitive reflection; variables 2–5 measure rational thinking (variable 6 is their composite); and variables 7–8 measure normative thinking (variable 9 is their composite).

*$p < .05$.
**$p < .01$.
***$p < .001$.

English-speaking adults for use with Chinese-speaking children. Still, correlations among the four tasks were in the predicted direction, as were correlations between each task and the CRT-D. Because no task was an outlier, we followed the conventions of prior research and pooled their variance, allowing us to capture general trends in rational thinking not well captured by a single task.

1.5. Normative thinking disposition measures

1.5.1. Need for cognition

Adults completed a nine-item need for cognition scale (NFC) of Kokis et al. (2002). They rated their agreement with statements like "I like hard problems instead of easy ones" and "It's really cool to figure out a new way to do something." Children completed a 14-item child-friendly version of NFC developed by Keller et al. (2016). They rated their agreement with statements like "Thinking is fun for me" and "I like learning new things." Children and adults responded on a four-point agreement scale, with higher scores indicating greater motivation to engage in effortful cognitive activities. McDonald's ω_{Total} for the measure was .85 in adults and .80 in children. Cronbach's alpha was .81 in adults and .73 in children.

1.5.2. Actively open-minded thinking

Adults completed an eight-item actively open-minded thinking scale (AOT), adapted from Haran, Ritov, and Mellers (2013). They rated their agreement with statements like "Allowing oneself to be convinced by an opposing argument is a sign of good character" and "People should take into consideration evidence that goes against their beliefs." Children completed a seven-item version, with several items modified to be more child-friendly (Young et al., 2018), such as "It is good to listen to the other side of an argument" and "Changing your mind is a bad thing." Participants responded on a four-point agreement scale, with higher scores indicating a greater tendency toward open-minded thinking. McDonald's ω_{Total} for the measure was .76 in adults and .41 in children. Cronbach's alpha was .66 in adults and .19 in children.

1.5.3. Composite score

A normative thinking disposition composite was created by averaging z scores from the NFC and AOT scales. Correlations among these measures can be seen in Tables 2 and 3.

1.6. Procedure

Participants completed the battery of tasks in the following order: CRT-D, denominator neglect, belief bias syllogisms, base rate sensitivity, other-side thinking, need for cognition, and actively open-minded thinking (as done by Young et al., 2018). Adults answered the CRT-D questions intermixed with questions from the CRT-ALT and the original CRT following the other-side thinking task. Adults completed the study online via Qualtrics. Children completed the study one-on-one with trained research assistants using paper-and-pencil

tests. Research assistants read the assessment items aloud and then recorded children's verbal answers on answer sheets.

2. Results

We first investigate whether the CRT-D, when given to Chinese adults, corresponds to measures of rational thinking and normative thinking dispositions, as the original CRT does. We then examine whether the same relations hold for Chinese children. Finally, we examined the CRT-D's unique contribution to children's rational thinking and normative thinking dispositions after adjusting for age, and we compared our results to those from the U.S. sample (Young et al., 2018). We focus on higher-order relations between cognitive reflection, rational thinking, and normative thinking dispositions rather than task-by-task comparisons of Chinese and U.S. samples because such comparisons are potentially confounded by socioeconomic status, educational background, and other such variables. We include a preliminary exploration of China–U.S. comparisons in the Supplementary Materials. Chinese children outperformed U.S. children on three of our nine tasks (CRT-D, belief bias syllogism, and actively open-minded thinking), but changes in performance with age were largely the same across countries. Only one task revealed an age-by-country interaction (actively open-minded thinking), and the effect was small. Data and R scripts to reproduce all analyses are available at https://osf.io/yhn7b/.

2.1. Adults' cognitive reflection

Table 2 presents descriptive statistics and bivariate correlations among the adult measures. For rational thinking, all three measures of cognitive reflection were positively correlated with evaluation of belief bias syllogisms and the overall rational thinking composite. Only the CRT-D yielded a significant correlation with base rate sensitivity. For normative thinking dispositions, only the CRT-D and CRT yielded significant correlations with actively open-minded thinking. And only the CRT-D yielded a significant correlation with the normative thinking composite as a whole.

Hoerger's (2013) corrected version of Stieger's z-test revealed that the correlation between the CRT-D and belief bias was not as strong as the correlation between the original CRT and belief bias ($r = .32$ vs. $r = .53$, $Z_H = -2.40$, $p = .016$). Apart from that, there were no differences between the CRT-D and original CRT in the strength of their correlations with the other variables, and no differences between the CRT-D and CRT-ALT in the strength of their correlations with the other variables. Finally, the correlation between the CRT-D and original CRT was about as strong as the correlation between the CRT-ALT and original CRT ($r = .35$ vs. $r = .51$, $Z_H = -1.81$, $p = .069$).

These results suggest the CRT-D functioned similarly to established measures of cognitive reflection (the CRT and CRT-ALT) for Chinese adults, with the exception of predicting belief bias, which the original CRT predicts better. Moreover, the CRT-D predicted both rational thinking and normative thinking dispositions, while the original CRT and CRT-ALT predicted only rational thinking.

Table 4
Regression analyses using age and CRT-D to predict rational thinking and normative disposition composites in Chinese children and U.S. children (from Young et al., 2018)

	China			United States		
	R^2	Standardized β	95% CI	R^2	Standardized β	95% CI
Rational thinking composite						
1. Age (months)	.157	.396**	[.222, .570]	.274	.301**	[.199, .404]
2. Age (months)	.203	.289**	[.099, .479]	.311	.225**	[.132, .357]
CRT-D		.241*	[.051, .431]		.123*	[.012, .235]
Normative disposition composite						
1. Age (months)	.002	.047	[–.143, .237]	.073	.203*	[.047, .359]
2. Age (months)	.056	-.068	[–.275, .140]	.125	.110	[–.064, .283]
CRT-D		.259*	[.051, .467]		.188*	[.021, .355]

*$p < .05$.
**$p < .001$.

2.2. Children's cognitive reflection

Table 3 presents descriptive statistics and bivariate correlations among the measures administered to children. The sample size for two measures (NFC and AOT) was 110, rather than 111, due to missing data. For rational thinking, the CRT-D was significantly correlated with children's denominator neglect, base rate sensitivity, and evaluation of belief bias syllogisms, as well as the overall composite. For normative thinking dispositions, the CRT-D yielded significant correlations with actively open-mind thinking and the overall composite. The CRT-D thus predicted several distinct measures of rationality associated with cognitive reflection in U.S. adults (Thomson & Oppenheimer, 2016; Toplak et al., 2014a) and U.S. children (Young et al., 2018), suggesting that it is a valid tool for measuring children's cognitive reflection in China.

Children's age was also correlated with the CRT-D, denominator neglect, belief bias, otherside thinking, the rational thinking composite, and need for cognition. To assess whether the CRT-D remained a significant predictor independent of age, we performed hierarchical regressions for both composite scores (Table 4). After adjusting for age, the CRT-D explained an additional 4.6% of the variance in the rational thinking composite, $F(1,108) = 6.30$, $p = .014$, and explained an additional 5.4% of the variance in the normative thinking composite, $F(1,106) = 6.08$, $p = .015$. Thus, the CRT-D's contribution to predicting rational thinking and normative thinking dispositions is not fully explained by the variance shared with age.

The CRT-D's predictive power, adjusting for age, appeared highly consistent across Chinese and U.S. samples (4.6% vs. 3.7% for rational thinking and 5.4% vs. 5.2% for normative thinking dispositions). We tested this consistency by regressing children's rational thinking and normative thinking dispositions against their age, CRT-D score, and country (China vs. United States; see Table 5). Country did not predict additional variance in rational thinking,

Table 5
Regression analyses using age, CRT-D, and country to predict rational thinking and normative disposition composites in a combined sample of Chinese and U.S. children

	Rational Thinking Composite		Normative Disposition Composite	
	Standardized β	95% CI	Standardized β	95% CI
Age (months)	.311**	[.123, .498]	–.070	[–.282, .142]
CRT-D	.236**	[.059, .412]	.226**	[.026, .427]
Country	–.139	[–.395, .117]	–.152	[–.442, .139]
Age * Country	.129	[–.157, .415]	.251	[–.074, .576]
CRT-D * Country	–.018	[–.311, .275]	.047	[–.283, .378]

*$p < .05$.
**$p < .01$.

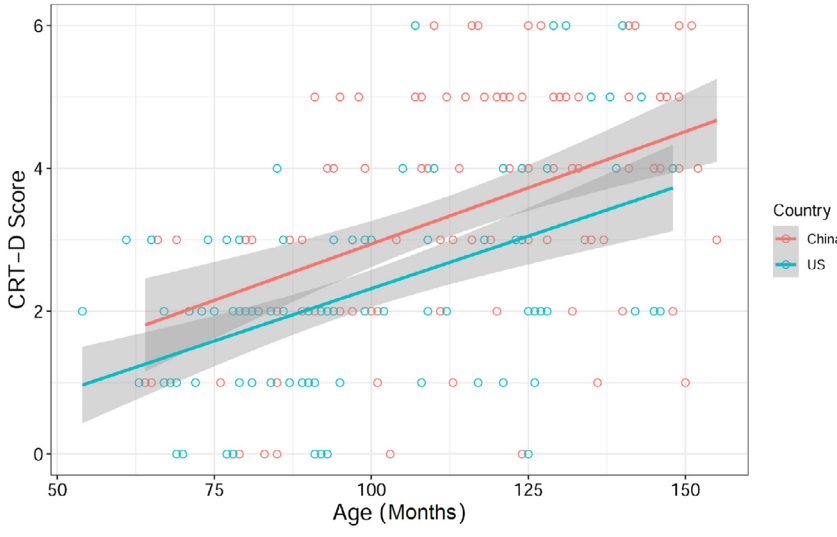

Fig. 2. Predicted CRT-D scores by age and culture (including their interaction). Dots represent individual children and ribbons indicate 95% CIs.

$F(2,197) = 0.73$, $p = .538$, or normative thinking dispositions, $F(2,193) = 1.61$, $p = .188$, nor did it interact with age or CRT-D scores. CRT-D scores were thus similarly predictive of the target measures in both countries.

Finally, we examined whether the developmental trajectories of cognitive reflection were similar across Chinese and U.S. samples. To do so, we fit a linear regression model on children's CRT-D scores with age, country, and their interaction as predictors. Note that we only considered the seven CRT-D items administered to both samples. As can be seen in Fig. 2, there was no age-by-country interaction, $F(1,203) = 0.07$, $p = .792$, indicating that Chinese and U.S. children demonstrate similar increases in CRT-D scores with age.

3. Discussion

This study validated a child-friendly measure of cognitive reflection, the CRT-D, in a non-Western culture. Consistent with findings from Western adults (Thomson & Oppenheimer, 2016; Toplak et al., 2014a), Chinese adults' cognitive reflection, as measured by the original CRT, was related to rational thinking and certain thinking dispositions (actively open-minded thinking). The CRT-D resembled the original CRT in these relations, correlating with most measures of rationality to a similar degree (i.e., denominator neglect, base rate sensitivity, otherside thinking, need for cognition, actively open-minded thinking, and alternative CRT items). Correspondences between cognitive reflection and rational thought are robust in Western contexts (Baron, Scott, Fincher, & Metz, 2015; Thomson & Oppenheimer, 2016; Toplak et al., 2014a; Young et al., 2018), but it was unknown whether such correlations depend on Western-endorsed thinking styles. Our findings indicate that they hold even in an Eastern context, where holistic styles of reasoning are endorsed and often preferred.

For Chinese children, the CRT-D correlated with largely the same measures of rationality: belief bias, denominator neglect, base rate sensitivity, and actively open-minded thinking. The CRT-D predicted children's overall rational thinking and normative thinking dispositions above and beyond children's age, and the predictive powers were similar to the corresponding results from U.S. samples (Young et al., 2018). Indeed, age-related increases in cognitive reflection were largely the same in China and the United States. These results support the CRT-D as a valid measure of cognitive reflection in Chinese children.

Our results suggest that cognitive reflection tracks rational thinking across the lifespan and across cultures. The present research provides the first evidence for the coherence of cognitive reflection, rational thinking, and normative thinking dispositions in China, both in terms of individual differences among adults and developmental differences among children. Although Westerners are observed to favor analytic thinking styles more than Easterners (Nisbett et al., 2001; Norenzayan et al., 2002), the structure of rational thought appears similar in Chinese and American cultures. While the two cultures value different decision-making strategies on the whole, individuals within those cultures may adopt stable preferences for one strategy over another. Some Chinese individuals may adopt a preference for analytic strategies, and some American individuals may adopt a preference for nonanalytic strategies, and these preferences are well predicted by cognitive reflection within both cultures. Group-level differences in thinking dispositions (Buchtel & Norenzayan, 2008; Choi et al., 2004; Tweed & Lehman, 2002) or susceptibility to inductive errors (Norenzayan et al., 2002; Pohl et al., 2002; Yates et al., 2002) mask individual differences within each group, and these individual differences appear to track the more general tendency to reflect on one's cognition in both China and the United States and among both children and adults.

While we found that cognitive reflection develops similarly across cultures (Fig. 2), our analyses do not control for related abilities, such as inhibitory control, task shifting, and working memory. We did not assess these executive function skills because their assessment is time-consuming and attention-demanding, and we preferred to devote our testing sessions to surveying rational thinking and normative thinking as broadly as possible. Research with American children has found that cognitive reflection predicts rational thinking even when

controlling for executive function (Young & Shtulman, 2020a), indicating that the two competencies are not redundant. Cognitive reflection requires inhibition and working memory—two facets of executive function—but they must be applied and coordinated in ways that traditional measures of executive function do not tap. Nevertheless, future research is needed to assess the cross-cultural relation between cognitive reflection and related competencies, in light of the possibility that the emergence of cognitive reflection depends on executive function skills that develop differently in different cultures.

Another limitation is that some of our measurement scales had lower than ideal reliability. We approached these scales conservatively, retaining their original items to facilitate comparison with prior research, but we acknowledge that higher-reliability measures of cognitive reflection and normative thinking dispositions are needed for future research with Chinese samples. Scales developed for use with English-speaking adults in Western countries may not travel well, even if the competencies they are meant to measure remain consistent across cultures. Our measure of cognitive reflection, the CRT-D, may have addressed developmental differences in the kinds of brainteasers children could be expected to answer (relative to adults), but we invite further revision or expansion of this battery, to include items that elicit an erroneous-but-correctable response in children of all cultures.

Finally, although we were able to demonstrate that Chinese children's cognitive reflection tracks their analytic reasoning, we do not know whether it influences their learning. Cognitive reflection has been shown to facilitate science learning in American children (Young & Shtulman, 2020b) and American adults (Shtulman & McCallum, 2014), but Chinese individuals may not learn science the same way given different instructional strategies (Tweed & Lehman, 2002) and learning goals (Cheng & Guan, 2015). In this environment, cognitive reflection may not facilitate science learning above and beyond general intelligence (Zhang, Hu, Ren, & Fan, 2017). American and Chinese children also deviate in their moral education, with regard to what constitutes a social convention and what constitutes a moral prescription (Zhao & Kushnir, 2019). Future research could investigate whether cognitive reflection predicts moral judgment in China as it does in the United States (Paxton et al., 2012; Royzman et al., 2015) and, if so, whether cognitive reflection moderates cultural differences in moral development.

In conclusion, the present study provides evidence that cognitive reflection develops similarly in China and the United States, despite cultural differences in preferred thinking styles. It also demonstrates that cognitive reflection, rational thinking, and normative thinking dispositions cohere across development. We anticipate the CRT-D will be a valuable tool for studying the emergence of rational thought across ages and cultures.

Acknowledgments

The authors would like to thank Yulin Wu, Xinchao Ma, Ling Zhang, Yusi Hao, Xuchi Luo, Hanqi Yun, and Yanyan Chen for assistance with data collection and data coding, and Wenkai Li for assistance with translation. We would also like to thank the James S. McDonnell Foundation for supporting this research with an Understanding Human Cognition Scholar Award

to Andrew Shtulman. The data that support the findings of this study are openly available in the Open Science Framework at https://osf.io/yhn7b/.

Open Research Badges

This article has earned Open Data and Open Materials badges. Data and materials are available at https://tinyurl.com/yjo2q6ez.

REFERENCES

Bago, B., & De Neys, W. (2019). The smart system 1: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning*, *25*, 257–299. https://doi.org/10.1080/13546783.2018.1507949

Ball, L. J., & Thompson, V. A. (2018). Belief bias and reasoning. In L. J. Ball & V. A. Thompson (Eds.), *The Routledge international handbook of thinking and reasoning* (pp. 16–36). New York: Routledge.

Bao, L., Cai, T., Koenig, K., Fang, K., Han, J., Wang, J., & Wang, Y. (2009). Learning and scientific reasoning. *Science*, *323*, 586–587. https://doi.org/10.1126/science.1167740

Baron, J., Scott, S., Fincher, K., & Metz, S. E. (2015). Why does the cognitive reflection test (sometimes) predict utilitarian moral judgment (and other things)? *Journal of Applied Research in Memory and Cognition*, *4*, 265–284. https://doi.org/10.1016/j.jarmac.2014.09.003

Buchtel, E. E., & Norenzayan, A. (2008). Which should you use, intuition or logic? Cultural differences in injunctive norms about reasoning. *Asian Journal of Social Psychology*, *11*, 264–273. https://doi.org/10.1111/j.1467-839X.2008.00266.x

Byrd, N., & Conway, P. (2019). Not all who ponder count costs: Arithmetic reflection predicts utilitarian tendencies, but logical reflection predicts both deontological and utilitarian tendencies. *Cognition*, *192*, 103995. https://doi.org/10.1016/j.cognition.2019.06.007

Cacioppo, J., & Petty, R. (1982). The need for cognition. *Journal of Personality and Social Psychology*, *42*, 116–131. https://doi.org/10.1037/0022-3514.42.1.116

Carstensen, A., Zhang, J., Heyman, G. D., Fu, G., Lee, K., & Walker, C. M. (2019). Context shapes early diversity in abstract thought. *Proceedings of the National Academy of Sciences*, *116*, 13891–13896. https://doi.org/10.1073/pnas.1818365116

Cheng, H. Y., & Guan, S. Y. (2015). Unravelling the influence of cognitive style on Chinese students' classroom behaviours: The mediating effects of the structure-oriented/depth-oriented learning approach. *Educational Psychology*, *35*, 497–512. https://doi.org/10.1080/01443410.2013.817539

Choi, I., Choi, J. A., & Norenzayan, A. (2004). Culture and decisions. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgement and decision making* (pp. 505–524).Malden, MA: Blackwell.

Cokely, E. T., & Kelley, C. M. (2009). Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation. *Judgment and Decision Making*, *4*, 20–33. http://www.sjdm.org/journal/81125/jdm81125.pdf

Don, H. J., Goldwater, M. B., Otto, A. R., & Livesey, E. J. (2016). Rule abstraction, model-based choice, and cognitive reflection. *Psychonomic Bulletin & Review*, *23*, 1615–1623. https://doi.org/10.3758/s13423-016-1012-y

Evans, J. S. B. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255–278. https://doi.org/10.1146/annurev.psych.59.103006.093629

Evans, J. S. B., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, *11*, 295–306. https://doi.org/10.3758/BF03196976

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives*, *19*, 25–42. https://doi.org/10.1257/089533005775196732

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, *34*, 752–775. https://doi.org/10.1111/j.1551-6709.2010.01114.x

Gervais, W. M. (2015). Override the controversy: Analytic thinking predicts endorsement of evolution. *Cognition*, *142*, 312–321. https://doi.org/10.1016/j.cognition.2015.05.011

Gervais, W. M., van Elk, M., Xygalatas, D., McKay, R. T., Aveyard, M., Buchtel, E. E., & Svedholm-Häkkinen, A. M. (2018). Analytic atheism: A cross-culturally weak and fickle phenomenon. *Judgment & Decision Making*, *13*, 268–274. http://www.sjdm.org/journal/18/18228/jdm18228.pdf

Haran, U., Ritov, I., & Mellers, B. A. (2013). The role of actively open-minded thinking in information acquisition, accuracy, and calibration. *Judgment and Decision Making*, *8*, 188–201. http://www.sjdm.org/journal/13/13124a/jdm13124a.pdf

Hoerger, M. (2013). ZH: An updated version of Steiger's Z and web-based calculator for testing the statistical significance of the difference between dependent correlations. Retrieved from http://www.psychmike.com/dependent_correlations.php.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Straus & Giroux.

Keller, U., Strobel, A., Wollschläger, R., Greiff, S., Martin, R., Vainikainen, M. P., & Preckel, F. (2016). A Need for Cognition Scale for children and adolescents. *European Journal of Psychological Assessment*, *35*, 137–149. https://doi.org/10.1027/1015-5759/a000370

Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology*, *83*, 26–52. https://doi.org/10.1016/S0022-0965(02)00121-2

Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, *108*, 677–692. https://doi.org/10.1016/j.jecp.2010.11.001

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*, 291–310. https://doi.org/10.1037/0033-295X.108.2.291

Norenzayan, A., Smith, E. E., Kim, B. J., & Nisbett, R. E. (2002). Cultural preferences for formal versus intuitive reasoning. *Cognitive Science*, *26*, 653–684. https://doi.org/10.1207/s15516709cog2605_4

Paxton, J. M., Unger, L., & Greene, J. D. (2012). Reflection and reasoning in moral judgment. *Cognitive Science*, *36*, 163–177. https://doi.org/10.1111/j.1551-6709.2011.01210.x

Peng, J., Feng, T., Zhang, J., Zhao, L., Zhang, Y., Chang, Y., & Xiao, W. (2019). Measuring decision making competence in Chinese adults. *Journal of Behavioral Decision Making*, *32*, 266–279. https://doi.org/10.1002/bdm.2114

Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2016). Is the cognitive reflection test a measure of both reflection and intuition? *Behavior Research Methods*, *48*, 341–348. https://doi.org/10.3758/s13428-015-0576-1

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015a). Everyday consequences of analytic thinking. *Current Directions in Psychological Science*, *24*, 425–432. https://doi.org/10.1177/0963721415604610

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015b). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology*, *80*, 34–72. https://doi.org/10.1016/j.cogpsych.2015.05.001

Pinillos, N. Á., Smith, N., Nair, G. S., Marchetto, P., & Mun, C. (2011). Philosophy's new challenge: Experiments and intentional action. *Mind & Language*, *26*, 115–139. https://doi.org/10.1111/j.1468-0017.2010.01412.x

Pohl, R., Bender, M., & Lachmann, G. (2002). Hindsight bias around the world. *Experimental Psychology*, *49*, 270–282. https://doi.org/10.1026/1618-3169.49.4.270

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2015). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making*, *29*, 453–469. https://doi.org/10.1002/bdm.1883

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*, 177–192. https://doi.org/10.1002/wcs.1336

Royzman, E. B., Landy, J. F., & Leeman, R. F. (2015). Are thoughtful people more utilitarian? CRT as a unique predictor of moral minimalism in the dilemmatic context. *Cognitive Science*, *39*, 325–352. https://doi.org/10.1111/cogs.12136

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*, 74–81. https://doi.org/10.1111/j.1467-9280.2005.01667.x

Shtulman, A., & McCallum, K. (2014). Cognitive reflection predicts science understanding. *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, 2937–2942.Retrieved from https://cogsci.mindmodeling.org/2014/papers/507/paper507.pdf

Siegler, R. S., & Mu, Y. (2008). Chinese children excel on novel mathematics problems even before elementary school. *Psychological Science*, *19*, 759–763. https://doi.org/10.1111/j.1467-9280.2008.02153.x

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22. https://doi.org/10.1037/0033-2909.119.1.3

Stanovich, K. E., & West, R. F. (1997). Reasoning independently of prior belief and individual differences in actively open-minded thinking. *Journal of Educational Psychology*, *89*, 342–357. https://doi.org/10.1037/0022-0663.89.2.342

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, *23*, 645–665. https://doi.org/10.1017/S0140525X0028343X

Stanovich, K., West, R., & Toplak, M. (2016). *The rationality quotient: Toward a test of rational thinking*. Cambridge, MA: MIT Press.

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, *11*, 99–113. http://journal.sjdm.org/15/151029/jdm151029.pdf

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition*, *39*, 1275–1289. https://doi.org/10.3758/s13421-011-0104-1

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014a). Assessing miserly information processing: An expansion of the cognitive reflection test. *Thinking & Reasoning*, *20*, 147–168. https://doi.org/10.1080/13546783.2013.844729

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014b). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology*, *50*, 1037–1048. https://doi.org/10.1037/a0034910

Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the cognitive reflection test. *Cognition*, *150*, 109–118. https://doi.org/10.1016/j.cognition.2016.01.015

Tweed, R. G., & Lehman, D. R. (2002). Learning considered within a cultural context: Confucian and Socratic approaches. *American Psychologist*, *57*, 89–99. https://doi.org/10.1037/0003-066X.57.2.89

Wang, J., & Lin, E. (2009). A meta-analysis of comparative studies on Chinese and US students' mathematics performance: Implications for mathematics education reform and research. *Educational Research Review*, *4*, 177–195. https://doi.org/10.1016/j.edurev.2009.06.003

Yates, J. F., Lee, J. W., Sieck, W. R., Choi, I., & Price, P. C. (2002). Probability judgment across cultures. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 271–291). Cambridge: Cambridge University Press.

Young, A. G., Powers, A., Pilgrim, L., & Shtulman, A. (2018). Developing a cognitive reflection test for school-age children. *Proceedings of the 40th Annual Conference of the Cognitive Science Society*, 1238–1243.Retrieved from https://cogsci.mindmodeling.org/2018/papers/0241/0241.pdf

Young, A. G., & Shtulman, A. (2020a). Children's cognitive reflection predicts conceptual understanding in science and mathematics. *Psychological Science*, *31*, 1396–1408. https://doi.org/10.1177/0956797620954449

Young, A. G., & Shtulman, A. (2020b). How children's cognitive reflection shapes their science understanding. *Frontiers in Psychology*, *11*, 1247. https://doi.org/10.3389/fpsyg.2020.01247

Zhang, X., Hu, B. Y., Ren, L., & Fan, X. (2017). Pathways to reading, mathematics, and science: Examining domain-general correlates in young Chinese children. *Contemporary Educational Psychology*, *51*, 366–377. https://doi.org/10.1016/j.cedpsych.2017.09.004

Zhao, X., & Kushnir, T. (2019). How U.S. and Chinese children talk about personal, moral and conventional choices. *Cognitive Development*, *52*, article 100804. https://doi.org/10.1016/j.cogdev.2019.100804

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's alpha, Revelle's beta, McDonald's omega: Their relations with each and two alternative conceptualizations of reliability. *Psychometrika*, *70*, 123–133. https://doi.org/10.1007/s11336-003-0974-7

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Regression analyses using age and region (within China) to predict task performance.

Table S2. Regression analyses using age and country (China vs. US) to predict task performance.

Figure S1. Task performance by age and region (within China).

Figure S2. Task performance by age and country (China vs. US).