

ARTICLE

The development of cognitive reflection

Andrew Shtulman¹  | Andrew G. Young²¹Occidental College, Los Angeles, California, USA²Northeastern Illinois University, Chicago, Illinois, USA

Correspondence

Andrew Shtulman, Department of Psychology, Occidental College, Los Angeles, CA 90041, USA.
Email: shtulman@oxy.edu

Funding information

James S. McDonnell Foundation

Abstract

What do cows drink? The correct answer is water, but many are tempted to say milk. The disposition to override an intuitive response (milk) with a more analytic response (water) is known as *cognitive reflection*. Tests of cognitive reflection predict a wide range of skills and abilities in adults. In this article, we discuss the construction of a developmental version of the cognitive reflection test and explore how it predicts rational thinking and normative thinking dispositions in elementary school-aged children, independent of age, executive function, and cultural context. We also explore how the test predicts children's mastery of counterintuitive concepts in science and mathematics. Findings suggest that cognitive reflection may be a prerequisite for developing, and improving, analytic thought, thus highlighting the value of studying cognitive reflection from a developmental perspective.

KEYWORDS

cognitive development, conceptual understanding, rational thought

Suppose a patch of lily pads has begun to grow at a nearby lake. You have noticed that the patch doubles in size each day. After 48 days, the entire lake is covered in lily pads. How many days did it take for the patch to cover half the lake? Your first thought may have been 24, as the prompt would seem to imply that you need to divide 48 by 2. But if the patch doubled each day, then it would have covered half the lake 1 day prior to covering the entire lake, so the correct answer is 47. This question is one of three devised by Frederick (2005) to assess cognitive reflection, or the propensity to reflect on one's own cognitive activity. The questions on Frederick's cognitive reflection test (CRT) are designed to elicit an erroneous intuitive response that can be recognized as incorrect on reflection and then corrected.

Cognitive reflection likely entails a variety of skills, including error monitoring, response inhibition, cognitive flexibility, and analytic reasoning. To succeed on the lily pad question, a person whose first thought is 24 must recognize that this answer does not satisfy the question's parameters and then search for an alternative. The

process of searching for alternatives requires not only inhibiting the incorrect answer but also reanalyzing the problem to find an answer that satisfies its parameters.

While it is debatable whether all such steps are necessary to succeed on the CRT (Bago & De Neys, 2019; Raelison et al., 2020), the test has proven a powerful predictor of cognitive activities that pit intuition against analysis. Adults with higher CRT scores have higher levels of rational thought (Toplak et al., 2011) and are more likely to endorse normative thinking dispositions, like the need for cognition and actively open-minded thinking (Stanovich et al., 2016), than adults with lower scores. They have a more comprehensive understanding of science (Shtulman & McCallum, 2014) and are more likely to accept science as true (Gervais, 2015). They use superior causal-inference strategies (Don et al., 2016) and social-coordination strategies (Corgnet et al., 2015), and they are better at rejecting unsubstantiated claims, including fake news (Pennycook & Rand, 2019), paranormal beliefs (Pennycook et al., 2012), and generic stereotypes (Hammond & Cimpian, 2017).

Abbreviation: CRT, cognitive reflection test.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Child Development Perspectives* published by Wiley Periodicals LLC on behalf of Society for Research in Child Development.

Research using the CRT has tended to treat cognitive reflection as a stable individual difference, concluding that some adults are just more reflective than others, but where do these differences come from? How does the disposition to reflect on one's cognition develop? Cognitive reflection can and should be studied from a developmental perspective for several reasons. This type of reflection is a domain-general capacity that predicts many domain-specific competencies, providing a window onto the processing demands involved in the development of higher-order cognition. Methodologically, cognitive reflection simply yet powerfully predicts many different cognitive activities, providing an efficient tool for assessing domain-general contributions to children's task-specific performance. Pedagogically, cognitive reflection is associated with rational thought and behavior, providing a pathway for improving children's analytic reasoning and reducing their susceptibility to cognitive biases.

We highlight insights from the developmental study of cognitive reflection using a child-friendly version of the CRT dubbed the CRT-D, where D stands for developmental (Young & Shtulman, 2020a). The nine-item CRT-D measures children's ability to privilege analysis over intuition. We describe the CRT-D's construction and validation, as well as its relation to other domain-general measures of cognitive ability. The CRT-D's success at predicting rational thought and conceptual understanding across ages and cultures suggests that cognitive reflection may be an ideal vantage point for studying—and improving—children's higher-order cognition.

MEASURING CHILDREN'S COGNITIVE REFLECTION

Frederick's (2005) original test of cognitive reflection is unsuitable for children because the questions involve mathematical operations that many children do not know. Alternative versions of the CRT have been constructed, such as the CRT-ALT (Primi et al., 2016), but

these tests also tend to involve math. In fact, a common criticism of cognitive-reflection tests is that they test numeracy as much as reflection (Otero et al., 2022; Thomson & Oppenheimer, 2016).

A suitable measure for children should elicit the same response structure as the CRT but without involving math, drawing instead on early-developing semantic knowledge, both in terms of the intuitions elicited and the analytic considerations needed to override those intuitions. With these criteria in mind, we identified nine brainteasers that could be posed to children as young as age 5 (Young & Shtulman, 2020a). Some were taken from a verbal version of the CRT (Thomson & Oppenheimer, 2016), but most were culled from the Internet. One such brainteaser is “What do cows drink?” This question elicits the intuitive response “milk,” given the strong semantic association between cows and milk, but a moment's reflection reveals that, while cows produce milk, they actually drink water.

The nine questions on the CRT-D (Young & Shtulman, 2020a) are listed in Table 1. When administered to 5- to 12-year-olds, they tend to elicit the associated intuitive response. The next most common response is the correct, analytic response. Rarely do the questions elicit random responses, indicating that they tap the desired conflict between intuition and analysis and are not just confusing or obscure. We have observed this response pattern across the age range for which the test was developed, though the proportion of correct, analytic responses to incorrect, intuitive responses increases with age (Young & Shtulman, 2020a). This finding, on its own, contributes to the study of cognitive reflection by showing that cognitive reflection improves with age: 5-year-olds answer an average of one to three questions correctly; 9-year-olds, two to four; 12-year-olds, three to five; and adults, seven to nine.

As a construct, cognitive reflection is traditionally measured using items of the same structure—brainteasers—but these items can vary widely in content. The content covered by the CRT-D ranges from trick associations (questions 5 and 8) to trick patterns (questions

TABLE 1 The nine questions on the CRT-D, along with each question's correct (analytic) response and incorrect (intuitive) response

Question	Correct	Incorrect
1. If you are running a race and you pass the person in second place, what place are you in?	Second	First
2. Emily's father has three daughters. The first two are named Monday and Tuesday. What is the third daughter's name?	Emily	Wednesday
3. A farmer has five sheep, all but three run away. How many are left?	Three	Two
4. If there are three apples and you take away two, how many do you have?	Two	One
5. What do cows drink?	Water	Milk
6. What weighs more, a pound of rocks or a pound of feathers?	Same weight	Rocks
7. What hatches from a butterfly egg?	Caterpillar	Baby butterfly
8. Who makes Christmas presents at the North Pole?	Elves	Santa
9. Anna is playing four square with her three friends: Eeny, Meeny, and Miny. Who is the fourth player?	Anna	Mo

2 and 9) to trick operations (questions 3 and 4), and spans topics as varied as cows, apples, and races. Yet despite this variation, the items on the CRT-D exhibit high internal reliability—as high as that of adult measures, if not higher (Gong et al., 2021; Young & Shtulman, 2020a). The reasoning processes tapped by the CRT-D appear to cohere across distinct facets of semantic knowledge.

COGNITIVE REFLECTION AND RATIONAL THOUGHT

The CRT was initially validated in the context of heuristics and biases tasks. In several studies, adults' CRT scores strongly predicted their adherence to normative principles of decision making (Frederick, 2005; Otero et al., 2022; Stanovich et al., 2016; Toplak et al., 2011). To determine whether the CRT-D functions similarly for children, we and our colleagues recruited elementary school-aged children from playgrounds in Pasadena, CA ($M_{\text{age}} = 8$ years, 1 month; 49% female) and assessed whether their performance on the CRT-D predicted their performance on child-friendly versions of heuristics and biases tasks (Gong et al., 2021; Young et al., 2018). The children in this study, as well as the other studies described later (Young & Shtulman, 2020a, 2020b), came from a community that is approximately 35% White (excluding Hispanic/Latino), 35% Hispanic or Latino, 18% Asian, and 8% Black; the community is largely middle class, with 14% of the population living below the poverty line (U.S. Census Bureau, 2021).

The tasks in our assessment examined four facets of rational thought (or lack thereof): belief bias, denominator neglect, base-rate sensitivity, and other-side thinking. Belief bias was measured by asking children to decide whether unbelievable conclusions could follow from valid arguments and whether believable conclusions could follow from invalid ones (Toplak et al., 2014). Denominator neglect was measured by asking children to decide between bets that pitted frequency against probability (9 chances to win out of 100 versus 1 chance to win out of 10; Kokis et al., 2002). Base-rate sensitivity was measured by asking children to decide between claims supported by statistics and claims supported by anecdotes (Kokis et al., 2002). Other-side thinking was measured by seeing whether children could generate reasons against a position they supported (Toplak et al., 2014).

Children were also queried on whether they endorsed two normative thinking dispositions: the need for cognition and actively open-minded thinking. Children were asked whether they agreed with statements like “Thinking is fun for me” and “I like learning new things” as measures of the need for cognition (Keller et al., 2016), and whether they agreed with statements like “It is good to listen to the other side of an argument” and “Changing your mind is a bad thing” as measures of actively open-minded thinking (Haran et al., 2013).

As predicted, children's CRT-D scores correlated with composite measures of rational thinking and normative thinking dispositions, as well as most individual tasks. Older children scored higher on these tasks than younger children, but the correlations between task performance and CRT-D scores held even when controlling for children's age (in months). The same correlations were found for adults. While adults scored substantially higher on the CRT-D than did children, they still exhibited variability, and this variability tracked individual differences in rational thought and normative thinking dispositions, similar to scores on the original CRT.

In a follow-up study (Gong et al., 2021), the same tasks were administered to Chinese participants to determine whether the results from the US study reflected a general feature of cognitive reflection or were a byproduct of Western culture. Western culture's emphasis on analytic reasoning may drive the relation between cognitive reflection and rational thought. This relation could emerge later, or in a different form, in cultures that emphasize holistic reasoning, such as Chinese culture.

The children in the Chinese study were recruited from public playgrounds in several regions, including Northern China (Hebei, Beijing, and Jilin), North Central China (Shanxi), West Central China (Sichuan), and Southeastern China (Fujian); they ranged in age from 5 to 12 years, similar to the US sample ($M_{\text{age}} = 9$ years, 5 months). Children from all regions exhibited the same developmental patterns observed in children in the US study: Their CRT-D scores correlated with composite measures of rational thinking and normative thinking dispositions, as well as most individual tasks, regardless of age. Chinese adults exhibited the same pattern of correlations as US adults, indicating that the relation between cognitive reflection and rational thought is robust across ages and cultures (Gong et al., 2021).

Responses in the Chinese study were so similar to those in the US study that culture did not significantly predict performance on the heuristics and biases tasks in combined analyses. The only significant predictor (aside from age) was participants' CRT-D scores. Thus, cognitive reflection appears to predict rational thought early in development and across diverse cultural contexts.

COGNITIVE REFLECTION AND CONCEPTUAL UNDERSTANDING

Decision making often pits intuition against analysis, which is likely why the ability to make rational decisions correlates with cognitive reflection. Another form of cognition that pits intuition against analysis is scientific cognition. Prior to learning scientific theories, we form intuitive theories of natural phenomena, which allow us to explain and predict everyday experiences like heat, motion, illness, and growth (Carey, 2009; Shtulman, 2017). Intuitive theories are useful in daily life, but they carve

the world into categories that are largely incompatible with science. Science learning requires grappling with these intuitions, as does scientific reasoning (Chi, 1992; Shtulman, 2022; Vosniadou, 1994).

The connection between cognitive reflection and scientific cognition was first demonstrated in adults, whose understanding and acceptance of science track their CRT scores. That is, adults with higher CRT scores understand astronomy, evolution, geology, mechanics, perception, and thermodynamics more accurately (Shtulman & McCallum, 2014) and are more likely to accept controversial scientific ideas, like evolution (Gervais, 2015) and climate change (da Rosa, 2021), than are adults with lower CRT scores.

To explore how the connection between cognitive reflection and scientific cognition develops, we administered the CRT-D to elementary school-aged children recruited from playgrounds in Pasadena ($M_{\text{age}} = 8$ years, 2 months; 56% female); we also tested them on whether they understood biology from a vitalist perspective (Young & Shtulman, 2020a). Vitalism is the idea that observable biological activities, like eating and breathing, are linked to unobservable organs, like stomachs and lungs, which extract energy from the environment and use that energy to support and maintain life (Inagaki & Hatano, 2004). Vitalism is counterintuitive because children initially understand life as the capacity for self-directed motion, leading to the misconception that animate phenomena, like the sun and the wind, are alive, and that seemingly inanimate organisms, like flowers and trees, are not.

To measure vitalism, we asked children to explain the functions of vital organs (the body parts task) and classify various entities as alive or not alive (the living things task; Bascandziev et al., 2018). We also administered a test of mathematical equivalence to older children (ages 8 to 12). This test consisted of addition problems with operations on both sides of the equation, like $1 + 5 = _ + 2$, which assesses whether children understand the equal sign as a mathematical relation or as a prompt to “put the answer here” (McNeil & Alibali, 2005). Equivalence, like vitalism, defies intuition, because it requires children to suppress an intuitive conception of arithmetic and use a more analytic one.

All three measures of conceptual understanding—the body parts task, the living things task, and the mathematical equivalence task—strongly correlated with children's cognitive reflection. In fact, children's CRT-D scores were generally a stronger predictor of conceptual understanding than their age or their performance on standard measures of executive function, which we discuss in the next section. Children's CRT-D scores thus predict their understanding of counterintuitive science and math concepts, but do they also predict their propensity to learn such concepts? Might children with higher CRT-D scores benefit more from instruction than children with lower scores?

We explored this possibility by teaching elementary school-aged children about the counterintuitive aspects of two domains of science—life and matter—and measuring their learning gains in relation to their cognitive reflection (Young & Shtulman, 2020b). The participants were again recruited from playgrounds in Pasadena ($M_{\text{age}} = 8$ years, 5 months; 58% female). In this study, we used a pre-post design in which children made true or false judgments for scientific statements before and after a tutorial on the relevant science. Some statements were intuitive and others were counterintuitive. For example, the statements “tulips are alive” and “tigers are alive” are both scientifically true, but only the second statement is intuitively true because only tigers appear to move on their own. Likewise, the statements “rivers are alive” and “rocks are alive” are both scientifically false, but only the second statement is intuitively false because only rocks lack motion. This task required children to prioritize their scientific understanding of everyday phenomena over their intuitive understanding.

The tutorials were designed to emphasize the scientific properties of life or matter, as well as to refute common misconceptions. For example, the tutorial about life emphasized that all living things need energy and nutrients, grow and develop, react to their environment, and reproduce. It addressed the misconception that life is synonymous with self-directed motion with examples of entities that do not move but are alive, like moss, and entities that move but are not alive, like comets.

The tutorials were effective at improving children's scientific reasoning. While children were consistently accurate at verifying intuitive statements, their accuracy at verifying counterintuitive statements increased substantially from pretest to posttest. In addition, children's CRT-D scores predicted their accuracy. At pretest, children with higher CRT-D scores verified counterintuitive statements more accurately than those with lower scores, replicating the earlier finding that CRT-D scores predict science understanding (independent of age). Extending these findings, children with higher CRT-D scores also showed greater improvements from pretest to posttest, indicating that they profited more from instruction (Young & Shtulman, 2020b).

Thus, cognitive reflection may be a prerequisite for achieving conceptual change. Cognitively reflective individuals may have more success at identifying gaps in their understanding or filling those gaps with new information. They may be more receptive to instruction or they may be better at monitoring and resolving conflicts in online reasoning. That said, the primary benefit of cognitive reflection may be fostering metaconceptual awareness. While children routinely reason *with* their concepts, they do not necessarily reason *about* their concepts, and the latter may be required for changing these concepts. Children who are disposed to reason about their concepts may have more opportunities to discover inconsistencies between their intuitive theories and the

expert theories modeled by parents and teachers. They may also be better positioned to resolve those inconsistencies by actively comparing the two theories and weighing their inferential value.

COGNITIVE REFLECTION AND EXECUTIVE FUNCTION

The finding that cognitive reflection predicts science learning parallels findings from studies showing that executive function predicts science learning (Bascandziev et al., 2018; Tardiff et al., 2020; Vosniadou et al., 2018). How might cognitive reflection and executive function be related?

Executive function has three main components: working memory, inhibitory control, and set shifting. Working memory is the capacity to maintain and manipulate task-relevant information; a common measure is backward digit recall, where children hear multidigit numbers of increasing length and repeat them back in reverse order (Alloway et al., 2009). Inhibitory control is the ability to suppress an undesired response in favor of a desired one; a common measure is the flanker task, where children indicate which direction an arrow is pointing when arrows on either side are pointing in the opposite direction (Rueda et al., 2004). Set shifting is moving adaptively between distinct sets of information; a common measure is verbal fluency, where children are asked to name as many examples of a category as they can. This task requires children to move adaptively between subcategories, such as “zoo animals” and “farm animals” when naming animals (Munakata et al., 2012).

All three executive function skills predict children's understanding of counterintuitive concepts, especially inhibitory control and set shifting (Tardiff et al., 2020; Vosniadou et al., 2018), but cognitive reflection is a stronger predictor. CRT-D scores predicted children's understanding of vitalism and mathematical equivalence more strongly than—and independently of—their performance on the three executive-function tasks noted earlier (Young & Shtulman, 2020a). Similarly, adults' CRT scores predicted their performance on heuristics and biases tasks independent of their executive function skills (Toplak et al., 2011).

Still, cognitive reflection and executive function are not entirely distinct. CRT-D scores correlate with performance on executive function tasks (Young & Shtulman, 2020a), possibly because cognitive reflection draws on executive function skills. To answer a brainteaser like “What do cows drink?” children must *inhibit* the gut response “milk.” They must also hold this gut response in *working memory* as they reanalyze the question—a process that requires *shifting* from an intuitive approach to an analytic one. Yet cognitive reflection is not wholly redundant with executive functions because it requires something more: the ability to recruit and

coordinate those functions. Inhibition alone will not suffice to answer a brainteaser correctly, nor will working memory or set shifting. All three skills must be used and their implementation must be initiated by the reasoner. Metacognitive awareness may mediate this process; executive function skills likely support and are supported by an awareness of one's thinking and decision making.

LOOKING AHEAD

Researchers need to determine which aspects of cognitive reflection are most critical to facilitating rational thought and conceptual understanding, in children as well as adults. Increasingly, evidence suggests that traditional math-based CRTs largely measure numeracy skills and general intelligence rather than reflection, calling into question the validity of math-based CRTs (Erceg et al., 2020; Otero et al., 2022). However, verbal-based CRTs more clearly tap reflective components beyond cognitive ability and numeracy (Sobkow et al., 2022). Cognitive reflection may also tap other distinct resources, such as the disposition to stop and think before responding (Wilkinson et al., 2020), identify alternative responses (Walker & Nyhout, 2020), or be vigilant against tricks (Bialek & Pennycook, 2018). Psychometrically focused studies might incorporate a range of cognitive abilities (e.g., IQ, executive function), thinking dispositions, and personality traits to understand more fully how the CRT-D relates to established individual differences.

Progress in understanding the development of cognitive reflection and its consequences will also require longitudinal studies. Only one study has used longitudinal data, finding that performance on a traditional math-based CRT increased from early to middle adolescence, but not from middle adolescence to early adulthood (Toplak, 2021). A few cross-sectional studies have found similar developmental differences from adolescence to adulthood using traditional math-based CRTs (Carriedo et al., 2020; Primi et al., 2016; Suzuki et al., 2021). While current research using the verbal-based CRT-D clearly suggests an overall improvement in cognitive reflection during the elementary school years, we know little about developmental trajectories beginning in early childhood.

More broadly, we suggest three directions for future research on children's cognitive reflection. First, a considerable portion of research on adult cognitive reflection is focused on outcomes of societal importance, such as rejecting stereotypes (Blanchar & Sparkman, 2020), rejecting conspiratorial beliefs (Stanley et al., 2021), and improving information literacy (Pennycook & Rand, 2019). Researchers might use the CRT-D to gain a deeper understanding of these critical behaviors across development.

Second, training studies could help establish the malleability of cognitive reflection and further reveal

which aspects of cognitive reflection are most critical. Training domain-general capacities are admittedly difficult, but researchers have had success with executive function (Diamond & Lee, 2011). Moreover, interventions with adults, such as debiasing training and decision justification, have improved CRT performance for up to 2 months (Boissin et al., 2021; Isler & Yilmaz, 2022). If children's cognitive reflection improves with age, it might also improve with targeted instruction.

Finally, research on children's cognitive reflection should interface with recent results and proposals regarding dual process theory (De Neys, 2022). In particular, recent evidence suggests that many individuals who provide correct answers on CRT problems do so without need for reflection; their intuitive response is the correct response (Bago & De Neys, 2019; Raelison et al., 2020). Accurate performance on the CRT may arise from a history of reflective thinking, during which certain counterintuitive responses are increasingly activated and automatized. Researchers might examine whether developmental differences on the CRT-D correspond to individual differences in mature dual-process reasoning. They should also examine whether the developmental correspondences between cognitive reflection and rational thought documented thus far hold in a broader range of populations, including countries beyond the United States and China, particularly non-Western ones, and in communities with different racial, ethnic, and socioeconomic characteristics. Cognitive reflection could develop at different rates in different populations or predict different facets of higher-order cognition.

CONCLUSIONS

The studies we have reviewed demonstrate that elementary school-aged children vary in their propensity to reflect on their cognition and that this variation, as measured by the CRT-D, predicts rational judgment, normative thinking dispositions, conceptual understanding of math and science, and the ability to learn counterintuitive concepts. The relation between cognitive reflection and these facets of higher-order cognition emerges early in development and remains consistent across the lifespan. Cognitive reflection tests, like the CRT-D, provide an efficient way to assess the domain-general prerequisites for many aspects of domain-specific cognition. Studies of this nature promise to improve our understanding of how we achieve rational thought and conceptual change, as well as our ability to facilitate those achievements.

FUNDING INFORMATION

The research reported in this article was supported by a James S. McDonnell Understanding Human Cognition Scholar Award awarded to Andrew Shtulman.

ORCID

Andrew Shtulman  <https://orcid.org/0000-0002-4687-3099>

REFERENCES

- Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliott, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development, 80*(2), 606–621. <https://doi.org/10.1111/j.1467-8624.2009.01282.x>
- Bago, B., & De Neys, W. (2019). The smart system I: Evidence for the intuitive nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning, 25*(3), 257–299. <https://doi.org/10.1080/13546783.2018.1507949>
- Bascandziev, I., Tardiff, N., Zaitchik, D., & Carey, S. (2018). The role of domain-general cognitive resources in children's construction of a vitalist theory of biology. *Cognitive Psychology, 104*(August), 1–28. <https://doi.org/10.1016/j.cogpsych.2018.03.002>
- Bialek, M., & Pennycook, G. (2018). The cognitive reflection test is robust to multiple exposures. *Behavior Research Methods, 50*, 1953–1959. <https://doi.org/10.3758/s13428-017-0963-x>
- Blanchar, J. C., & Sparkman, D. J. (2020). Individual differences in miserly thinking predict endorsement of racial/ethnic stereotypes. *Social Cognition, 38*, 405–421. <https://doi.org/10.1521/soco.2020.38.5.405>
- Boissin, E., Caparos, S., Raelison, M., & De Neys, W. (2021). From bias to sound intuiting: Boosting correct intuitive reasoning. *Cognition, 211*(5), 104645. <https://doi.org/10.1016/j.cognition.2021.104645>
- Carey, S. (2009). *The origin of concepts*. Oxford University Press.
- Carriedo, N., Corral, A., Montoro, P. R., & Herrero, L. (2020). A developmental study of the bat/ball problem of CRT: How to override the bias and its relation to executive functioning. *British Journal of Psychology, 111*(2), 335–356. <https://doi.org/10.1111/bjop.12400>
- Chi, M. (1992). Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.), *Cognitive models of science* (pp. 129–186). University of Minnesota Press.
- Corgnet, B., Espín, A. M., & Hernán-González, R. (2015). The cognitive basis of social behavior: Cognitive reflection overrides antisocial but not always prosocial motives. *Frontiers in Behavioral Neuroscience, 9*, 287. <https://doi.org/10.3389/fnbeh.2015.00287>
- da Rosa, J. A. (2021). Comparing climate science misconceptions with worldview and cognitive reflection suggests poor understanding and motivated reasoning among undergraduates. *Journal of Geoscience Education, 70*(4), 501–516. <https://doi.org/10.1080/10899995.2021.2006549>
- De Neys, W. (2022). Advancing theorizing about fast-and-slow thinking. *Behavioral and Brain Sciences, 1*–68. <https://doi.org/10.1017/S0140525X2200142X>
- Diamond, A., & Lee, K. (2011). Interventions shown to aid executive function development in children 4 to 12 years old. *Science, 333*(6045), 959–964. <https://doi.org/10.1126/science.1204529>
- 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>
- Erceg, N., Galić, Z., & Ružojčić, M. (2020). A reflection on cognitive reflection-testing convergent/divergent validity of two measures of cognitive reflection. *Judgment & Decision Making, 15*(5), 741–755. <https://journal.sjdm.org/20/200109a/jdm200109a.pdf>
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42. <https://doi.org/10.1257/089533005775196732>

- Gervais, W. M. (2015). Override the controversy: Analytic thinking predicts endorsement of evolution. *Cognition*, *142*(September), 312–321. <https://doi.org/10.1016/j.cognition.2015.05.011>
- Gong, T., Young, A. G., & Shtulman, A. (2021). The development of cognitive reflection in China. *Cognitive Science*, *45*(4), 12966. <https://doi.org/10.1111/cogs.12966>
- Hammond, M. D., & Cimpian, A. (2017). Investigating the cognitive structure of stereotypes: Generic beliefs about groups predict social judgments better than statistical beliefs. *Journal of Experimental Psychology: General*, *146*(5), 607–614. <https://doi.org/10.1037/xge0000297>
- 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*(3), 188–201. <https://sajdm.org/~baron/journal/13/13124a/jdm13124a.html>
- Inagaki, K., & Hatano, G. (2004). Vitalistic causality in young children's naive biology. *Trends in Cognitive Sciences*, *8*(1–3), 356–362. [https://doi.org/10.1016/0010-0277\(94\)90027-2](https://doi.org/10.1016/0010-0277(94)90027-2)
- Isler, O., & Yilmaz, O. (2022). How to activate intuitive and reflective thinking in behavior research? A comprehensive examination of experimental techniques. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-022-01984-4>
- Keller, U., Strobel, A., Wollschläger, R., Greiff, S., Martin, R., Vainikainen, M., & Preckel, F. (2016). A need for cognition scale for children and adolescents. Structural analysis and measurement invariance. *European Journal of Psychological Assessment*, *35*(1), 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*(1), 26–52. [https://doi.org/10.1016/S0022-0965\(02\)00121-2](https://doi.org/10.1016/S0022-0965(02)00121-2)
- McNeil, N. M., & Alibali, M. W. (2005). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*, *76*(4), 883–899. <https://doi.org/10.1111/j.1467-8624.2005.00884.x>
- Munakata, Y., Snyder, H. R., & Chatham, C. H. (2012). Developing cognitive control: Three key transitions. *Current Directions in Psychological Science*, *21*(2), 71–77. <https://doi.org/10.1177/0963721412436807>
- Otero, I., Salgado, J. F., & Moscoso, S. (2022). Cognitive reflection, cognitive intelligence, and cognitive abilities: A meta-analysis. *Intelligence*, *90*(January–February), 101614. <https://doi.org/10.1016/j.intell.2021.101614>
- Pennycook, G., Cheyne, J. A., Seli, P., Koehler, D. J., & Fugelsang, J. A. (2012). Analytic cognitive style predicts religious and paranormal belief. *Cognition*, *123*(3), 335–346. <https://doi.org/10.1016/j.cognition.2012.03.003>
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*(July), 39–50. <https://doi.org/10.1016/j.cognition.2018.06.011>
- Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (IRT). *Journal of Behavioral Decision Making*, *29*(5), 453–469. <https://doi.org/10.1002/bdm.1883>
- Raoelison, M., Thompson, V. A., & De Neys, W. (2020). The smart intuitor: Cognitive capacity predicts intuitive rather than deliberate thinking. *Cognition*, *204*(November), 104381. <https://doi.org/10.1016/j.cognition.2020.104381>
- Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, *42*(8), 1029–1040. <https://doi.org/10.1016/j.neuropsychologia.2003.12.012>
- Shtulman, A. (2017). *Scienceblind: Why our intuitive theories about the world are so often wrong*. Basic Books.
- Shtulman, A. (2022). Navigating the conflict between science and intuition. In M. Belanger, P. Potvin, S. Horst, A. Shtulman, & E. Mortimer (Eds.), *Multidisciplinary perspectives on representational pluralism in human cognition* (pp. 122–140). Routledge.
- Shtulman, A., & McCallum, K. (2014). Cognitive reflection predicts science understanding. In *Proceedings of the 36th Conference of the Cognitive Science Society* (pp. 2937–2942). <https://cogsci.mindmodeling.org/2014/papers/507/>
- Sobkow, A., Olszewska, A., & Sirota, M. (2022). The factor structure of cognitive reflection, numeracy, and fluid intelligence: The evidence from the polish adaptation of the verbal CRT. *Journal of Behavioral Decision Making*. <https://doi.org/10.1002/bdm.2297>
- Stanley, M. L., Barr, N., Peters, K., & Seli, P. (2021). Analytic-thinking predicts hoax beliefs and helping behaviors in response to the COVID-19 pandemic. *Thinking & Reasoning*, *27*(3), 464–477. <https://doi.org/10.1080/13546783.2020.1813806>
- Stanovich, K., West, R., & Toplak, M. (2016). *The rationality quotient: Toward a test of rational thinking*. MIT Press.
- Suzuki, A., Megumi, A., & Yasumura, A. (2021). Developmental changes in cognitive bias. *Psychology*, *12*, 293–303. <https://doi.org/10.4236/psych.2021.122019>
- Tardiff, N., Bascandziew, I., Carey, S., & Zaitchik, D. (2020). Specifying the domain-general resources that contribute to conceptual construction: Evidence from the child's acquisition of vitalist biology. *Cognition*, *195*(February), 104090. <https://doi.org/10.1016/j.cognition.2019.104090>
- Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making*, *11*(1), 99–113. <https://sajdm.org/~baron/journal/15/151029/jdm151029.html>
- Toplak, M. E. (2021). *Cognitive sophistication and the development of judgment and decision-making*. Academic Press.
- 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. (2014). Rational thinking and cognitive sophistication: Development, cognitive abilities, and thinking dispositions. *Developmental Psychology*, *50*(4), 1037–1048. <https://doi.org/10.1037/a0034910>
- U.S. Census Bureau. (2021). *QuickFacts*. <https://www.census.gov/quickfacts/pasadenacitycalifornia>
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, *4*(1), 45–69. [https://doi.org/10.1016/0959-4752\(94\)90018-3](https://doi.org/10.1016/0959-4752(94)90018-3)
- Vosniadou, S., Pnevmatikos, D., Makris, N., Lepenioti, D., Eikospentaki, K., Chountala, A., & Kyrianakis, G. (2018). The recruitment of shifting and inhibition in on-line science and mathematics tasks. *Cognitive Science*, *42*(6), 1860–1886. <https://doi.org/10.1111/cogs.12624>
- Walker, C. M., & Nyhout, A. (2020). Asking “why?” and “what if?": The influence of questions on children's inferences. In L. Butler, S. Ronfard, & K. Corriveau (Eds.), *The questioning child: Insights from psychology & education* (pp. 252–280). Cambridge University Press.
- Wilkinson, H. R., Smid, C., Morris, S., Farran, E. K., Dumontheil, I., Mayer, S., Tolmie, A., Bell, D., Porayska-Pomsta, K., Holmes, W., Mareschal, D., & Thomas, M. S. (2020). Domain-specific inhibitory control training to improve children's learning of counterintuitive concepts in mathematics and science. *Journal of Cognitive Enhancement*, *4*, 296–314. <https://doi.org/10.1007/s41465-019-00161-4>
- Young, A., & Shtulman, A. G. (2020a). Children's cognitive reflection predicts conceptual understanding in science and mathematics. *Psychological Science*, *31*(11), 1396–1408. <https://doi.org/10.1177/0956797620954449>

- Young, A. G., Powers, A., Pilgrim, L., & Shtulman, A. (2018). Developing a cognitive reflection test for school-age children. In *Proceedings of the 40th Conference of the Cognitive Science Society* (pp. 1232–1237). <https://cogsci.mindmodeling.org/2018/papers/0241/>
- 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>

How to cite this article: Shtulman, A., & Young, A. G. (2023). The development of cognitive reflection. *Child Development Perspectives*, *17*, 59–66. <https://doi.org/10.1111/cdep.12476>