Developmental Psychology
2014, Vol. 50, No. 4, 1037–1048

0012-1649/14/$12.00 DOI: 10.1037/a0034910

Rational Thinking and Cognitive Sophistication: Development, Cognitive Abilities, and Thinking Dispositions

Maggie E. Toplak
York University

Richard F. West
James Madison University

Keith E. Stanovich
University of Toronto

We studied developmental trends in 5 important reasoning tasks that are critical components of the operational definition of rational thinking. The tasks measured denominator neglect, belief bias, base rate sensitivity, resistance to framing, and the tendency toward otherside thinking. In addition to age, we examined 2 other individual difference domains that index cognitive sophistication: cognitive ability (intelligence and executive functioning) and thinking dispositions (actively open-minded thinking, superstitious thinking, and need for cognition). All 5 reasoning domains were consistently related to cognitive sophistication regardless of how it was indexed (age, cognitive ability, thinking dispositions). The implications of these findings for taxonomies of developmental trends in rational thinking tasks are discussed.

Keywords: rational thinking, decision making, intelligence, executive functions, development

Supplemental materials: http://dx.doi.org/10.1037/a0034910.supp

Developmental psychologists and cognitive scientists have recently shown an increasing interest in aspects of thinking that go beyond those assessed by intelligence tests (Blackwell, Trzesniewski, & Dweck, 2007; Duckworth & Seligman, 2005; Reyna, Chapman, Dougherty, & Confrey, 2012). This interest has been exemplified in the increasing attention in the developmental literature to the tasks and effects that have been demonstrated in the three-decades-long heuristics and biases research tradition (Kahneman, 2011). These tasks and effects collectively define rational thinking in modern cognitive science (Stanovich, 2011). However, they are not measured on intelligence tests, and performance on many of them shows strikingly modest correlations with intelligence (Stanovich, 2009; Stanovich & West, 2008b).

During the past decade, several research programs have attempted to map the developmental trajectory of some of these rational thinking tasks (Chiesi, Primi, & Morsanyi, 2011; De Neys & Vanderputte, 2011; Klaczynski, 2001; Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Morsanyi & Handley, 2008; Reyna & Ellis, 1994; Weller, Levin, & Denberg, 2011). Taken as a whole, however, the literature has been somewhat confusing. Amalgamating across tasks, there has been no fixed trend at all—no overall trend for rational thinking to be uniformly increasing or decreasing with age (Stanovich, Toplak, & West, 2008). Even when focused on a single effect (framing or base rate sensitivity, for example), the research is often inconsistent.

Developmental Trends and Stimulus Equivalence

One of the reasons for the lack of convergence across developmental studies of rational thinking is that developmentally appropriate stimuli in these domains are difficult to construct. Consider the example of an experiment by Jacobs and Potenza (1991) on base rate sensitivity. They found a significant positive developmental trend (more normative responses with age across 1st, 3rd, 6th grade students, and college students) in their so-called object condition, but in their so-called social condition the developmental trend was in the opposite direction—more reliance on the vivid information and less reliance on the more diagnostic statistical information by the *older* children. A consideration of the nature of the social problems reveals why this was the case. Here is an example of a social problem:

This article was published Online First November 4, 2013.

Maggie E. Toplak, Department of Psychology, York University, Toronto, Ontario, Canada; Richard F. West, Department of Graduate Psychology, James Madison University; Keith E. Stanovich, Department of Applied Psychology and Human Development, University of Toronto, Toronto, Ontario, Canada.

Preparation of this article was supported by grants from the Social Sciences and Humanities Research Council of Canada to Maggie E. Toplak and from the John Templeton Foundation to Keith E. Stanovich and Richard F. West. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. We thank the examiners who tested the children in this study: Alexandra Basile, Armita Hosseini, Jessica Sciaraffa, and Jill Shuster. We also thank the undergraduate students who scored the otherside thinking task: Saqina Abedi and Camelia Amiri. Finally, we thank the school boards and families who participated in this study.

Correspondence concerning this article should be addressed to Maggie E. Toplak, 126 BSB, Department of Psychology, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada. E-mail: mtoplak@yorku.ca

> In Juanita's class, 10 girls are trying out to be cheerleaders and 20 are trying out for the band. Juanita is very popular and very pretty. She is always telling jokes and loves to be around people. Do you think Juanita is trying out to be a cheerleader or for the band?

Here, the statistical information points in the direction of band, but the personal information points in the direction of cheerleader. But to understand the diagnosticity of the indicant information in this problem, one must have knowledge of a social stereotype: that popular girls are drawn more to cheerleading than to band. Knowledge of this stereotype might well increase with age and thus be less available to the younger children. Since being drawn to the stereotype is an indication of a cognitive bias (base rate neglect) in this paradigm, younger children are prevented from showing the bias because they do not know the stereotype. Older children, however, will be familiar with this stereotype. This will artificially create a pattern looking like a non-normative bias increasing with age. In fact, it is not really a *processing bias* increasing with age, but the knowledge that makes showing the bias *possible* that increases with age.

Davidson (1995) reported an analogous finding of a developmental trend of increasing conjunction errors with age. However, the stimuli in this study had the same problem as those in the social condition of Jacobs and Potenza's (1991) study—they depended on the knowledge of a stereotype that might increase with age. For example, one item went as follows:

> Mrs. Hill is not in the best health and she has to wear glasses to see. Her hair is gray and she has wrinkles. She walks kind of hunched over. Do you think that Mrs. Hill is: a waitress in a local restaurant; an old person who has grandchildren and a waitress at a local restaurant; etc.

Knowledge of the stereotype might lead children to make a conjunction error, but the knowledge of the stereotype is undoubtedly more extensive among sixth graders (Davidson's, 1995, oldest group) than among second graders (Davidson's, 1995, youngest group). Another way to state the problem with Davidson's stimuli is to say that the potential of the stimuli to trigger a representativeness judgment is confounded with age. Studies with less confounded stimuli have found base rate neglect and conjunction errors to decrease with age (De Neys & Vanderputte, 2011; Klaczynski, 2001; Kokis et al., 2002).

Such considerations explain some of the U-shaped developmental functions that have been found for heuristics and biases tasks (see Reyna & Brainerd, 1994, 1995; Reyna & Ellis, 1994). The U-shaped pattern can (among other reasons) arise when the responding of the youngest group is dominated by randomness because they do not even have the knowledge that is a prerequisite to the biased response. With age, children acquire the knowledge that enables a non-normative, biased response. Thus, at the youngest ages, the developmental trend may well be for normative responding to decline, because random responding is more likely to land on the normative response than is a systematically biased process. Finally, at the oldest ages, normative strategies either override the bias or are practiced to such automaticity that they are emitted automatically (Kahneman & Klein, 2009; Stanovich, West, & Toplak, 2011a).

In other words, such task situations often spawn the developmental sequence: random responding followed by smart errors (see Morsanyi & Handley, 2008)—finally followed by normative responding. Situations like this sometimes occur when the process of attribute substitution is taking place. According to Kahneman and Frederick (2002), people engage in attribute substitution when they substitute an easy-to-evaluate characteristic for a harder one, even if the easier one is less accurate. For example, people will substitute the less effortful attributes of vividness, stereotype, or affect for the more effortful retrieval of relevant facts (Kahneman, 2003; Li & Chapman, 2009; Slovic & Peters, 2006; Wang, 2009). Stereotypes here represent prototypical exemplars or categories that we hold in memory (Kahneman, 2011). Note, however, that the presence of knowledge or stereotypes is necessary for attribute substitution to work properly. Assume a person engages in attribute substitution by replacing a comprehensive analysis of diagnostic cues with a stereotype. But to engage in this type of attribute substitution, one must have stored enough exemplar information to abstract a stereotype for the situation. A child without the relevant stereotype knowledge will not be able to engage in this type of heuristic attribute substitution and would have to respond on some other basis.[1]

Convergence Across Different Measures of Cognitive Sophistication

The problem of stimulus equivalence might be put into context by realizing that these developmental studies are really trying to answer a more generic question. That question concerns whether people who display more complex and sophisticated cognition also tend to disproportionately give the normative response (i.e., to think more rationally). If the issue is phrased in this more generic manner, then development becomes just one of several ways of indexing cognitive sophistication. Importantly, the issue of stimulus equivalence is much less of a problem in studies that index cognitive sophistication in a manner other than by age. Thus, other measures of indexing cognitive sophistication can be used as converging evidence in triangulating the developmental trends.

We identify three general ways of indexing cognitive sophistication. The first is developmentally. It is straightforward to assume that adolescents are more cognitively sophisticated than young children and that, in turn, adults are more cognitively advanced than adolescents. The question then becomes whether cognitive developmental level, as indexed by age, is correlated with performance on rational thinking tasks. For a group of subjects of the same age, however, there are two different indices of cognitive sophistication: those that measure cognitive abilities and those that measure thinking dispositions. The natural index of cognitive ability is intelligence test performance or other highly related cognitive ability indices such as measures of executive functioning (Miyake & Friedman, 2012; Reyna, 1995; Salthouse, Atkinson, & Berish, 2003).

Intelligence, however, is not an exhaustive measure of cognitive functioning. For one thing, intelligence tests fail to tap important metacognitive strategies and cognitive styles that are critical components of what has been termed the reflective mind (Sinatra & Pintrich, 2003; Stanovich, 2009, 2011; Sternberg, 2003). These components of cognition travel under a variety of names in psy-

[1] However, Reyna and Brainerd (2011; Brainerd, Reyna, & Zember, 2011) have argued for the existence of several developmental reversals that are predicted by fuzzy-trace theory that cannot be accounted for by developmental differences in knowledge or by random responses.

chology—thinking dispositions or cognitive styles being the two most popular. Examples of some thinking dispositions that have been investigated by psychologists are: actively open-minded thinking, need for cognition, consideration of future consequences, need for closure, reflectivity, superstitious thinking, and dogmatism (Cacioppo, Petty, Feinstein, & Jarvis, 1996; Kruglanski & Webster, 1996; Liberali, Reyna, Furlan, Stein, & Pardo, 2012; Stanovich, 2011; Stanovich & West, 2007; Strathman, Gleicher, Boninger, & Edwards, 1994).

In the present study, we used all three indices of cognitive sophistication in an attempt to clarify developmental trends. The data on thinking dispositions and cognitive ability have tended to converge in empirical investigations. Specifically, the same rational thinking tasks that correlate with cognitive ability also correlate with thinking dispositions, and those that fail to correlate with cognitive ability also fail to correlate with thinking dispositions (Bruine de Bruin, Parker, & Fischhoff, 2007; Chiesi et al., 2011; Finucane & Gullion, 2010; Kokis et al., 2002; Parker & Fischhoff, 2005; Stanovich & West, 2007; Toplak, West, & Stanovich, 2011; West, Toplak, & Stanovich, 2008). As we have just argued, development is not the only way to measure increases in cognitive sophistication. And, when looking for trends with cognitive sophistication, development might be the most difficult one to study because of the stimulus equivalence issue discussed above. Thus, the data from nondevelopmental measures of cognitive sophistication become especially important as converging evidence.

In the present study, we examined denominator neglect (or ratio bias; see Reyna & Brainerd, 2008), belief bias in syllogistic reasoning (Evans, Barston, & Pollard, 1983; Markovits & Nantel, 1989), and base rate sensitivity (Bar-Hillel, 1980, 1990; Kahneman & Tversky, 1973). All three represent rational thinking paradigms that have shown relatively consistent associations with cognitive ability and thinking dispositions (De Neys, 2006; Kokis et al., 2002; Handley, Capon, Beveridge, Dennis, & Evans, 2004; Simoneau & Markovits, 2003; Stanovich & West, 1997, 1998). Thus, we might expect developmental trends in these tasks (which would indicate a convergence with the other two indicators of cognitive sophistication).

We examined another rational thinking skill in our study that has been examined using a variety of different paradigms. When people evaluate evidence, generate evidence, and test hypotheses they give inadequate attention to the beliefs, opinions, and evidence on the other side of the issue (Baron, 1995; Klaczynski & Robinson, 2000; Stanovich & West, 2007, 2008a; Westen, Blagov, Kilts, & Hamann, 2006). Otherside thinking is a recurrent theme in the critical thinking literature where the idea is that a critical thinker should be able to decouple their prior opinions from the evaluation of evidence and arguments and to consider evidence in opposition to their own view (e.g., Paul, 1984, 1987; Sá, Kelley, Ho, & Stanovich, 2005; Sternberg, 1997, 2001). It is what prevents the myside thinking bias described in the cognitive science literature (Baron, 2008; Stanovich & West, 2007). We adapted a paradigm used in several previous studies (Perkins, 1985; Perkins, Farady, & Bushey, 1991; Toplak & Stanovich, 2003) to examine the development of otherside thinking.

Finally, we studied a rational thinking task where the developmental results have been particularly confusing. The literature on framing effects in adults is vast (Kahneman & Tversky, 1984, 2000; Kühberger, 1998; Levin, Gaeth, Schreiber, & Lauriola, 2002; Maule & Villejoubert, 2007). However, the developmental literature is quite small. Obviously, the complexity of the problems has to be vastly reduced and made appropriate for children. Outcomes in developmental studies become small prizes that the children receive instead of the imaginary deaths or real money that is used in adult studies. Several investigators have creatively adapted framing paradigms for children, but the results of these experiments have been inconsistent. Levin and colleagues (Levin & Hart, 2003; Levin, Hart, Weller, & Harshman, 2007) found no developmental trend for framing effects. Children (6- to 8-year-olds) were more risk averse for gains than for losses in the manner that prospect theory predicts, but the magnitude of the framing effects that they displayed was the same as that found for adults. Framing effects were found for 5-year-olds, 6-year-olds, and 9- to 10-year-olds in three studies by Schlottmann and Tring (2005), and in 14.8-year-olds in Chien, Lin, and Worthley (1996). In the former study, the framing effects were roughly similar in magnitude.

The results of these studies were not completely convergent with those in a study by Reyna and Ellis (1994; see also Reyna et al., 2011; Reyna & Farley, 2006). The data patterns in Reyna and Ellis's (1994) study were complex, however, and variable over the ages studied. Framing interacted with level of risk and magnitude of reward at certain ages. Briefly though, preschoolers' responses were consistent across frames—they displayed no framing effect. Second graders displayed a small reverse framing effect—they were more risk averse for losses. Fifth-graders displayed a small reverse framing effect at the highest reward magnitude. Fifth-graders also displayed the standard framing effect (more risk seeking for losses) only for the lowest reward magnitude (see also, Reyna & Farley, 2006, in which additional replications of these patterns predicted by fuzzy-trace theory are reported). In the present study, we hope to partially resolve the conflicting findings from these different research groups by examining developmental trends together with cognitive ability and thinking dispositions in order to look for patterns of convergence among these various measures of cognitive sophistication. Our study is rare in looking at all three and allowing an examination of convergence.

The measures of cognitive ability used in our study encompassed intelligence indices, but we also included measures of executive functioning because the latter are strongly associated with fluid intelligence (Duncan et al., 2008; Duncan, Schramm, Thompson, & Dumontheil, 2012; Kane & Engle, 2002, 2003; Kane, Hambrick, & Conway, 2005; Salthouse et al., 2003). Recent work on the inhibitory and set-shifting properties of executive functioning tasks makes this class of processes a potentially theoretically interesting correlate of performance on rational thinking tasks (Aron, 2008; Best, Miller, & Jones, 2009; Friedman et al., 2007; Hasher, Lustig, & Zacks, 2007; Miyake & Friedman, 2012; Zelazo, 2004). As Kahneman (2003, 2011) has pointed out, many heuristics and biases problems require suppressing a prepotent "natural" response to the task. Such suppression could well be related to the types of set-shifting and inhibitory processes that are directly and indirectly assessed on measures of executive functioning. We thus, included three executive functioning tasks (a set-shifting task, an inhibition task, and a working memory task) in our study to complement the intelligence measures of cognitive ability. Finally, as measures of thinking dispositions, we sampled dispositions that were quite varied but that had been examined in previous studies of rational thinking: actively open-minded thinking, need for cognition, and superstitious thinking (Chatterjee, Heath, Milberg, & France, 2000; Kokis et al., 2002; Smith & Levin, 1996). The indicators of cognitive ability and thinking dispositions were examined as predictors of five rational thinking tasks (denominator neglect, belief bias syllogisms, base rate sensitivity, resistance to framing, and otherside thinking) in a developmental sample.

Method

Participants

The participants were 204 students (110 males and 94 females) in Grades 2, 3, 4, 5, 6, 7, 8, and 9 in regular classes from suburban and rural schools. The communities from which the participants were recruited included middle to upper income communities with families from mainly European descent. Four participants from the original sample of 208 were eliminated for having pro-rated IQ scores of less than 80. The participants were divided into three groups distinguished by grade level: Grades 2–3 ($n = 54$), Grades 4–5 ($n = 82$), and Grades 6–9 ($n = 68$). Table 1 displays the mean ages of each of the three groups.

Tasks and Variables

All of the tasks are summarized here, but they are described in more detail in the online supplemental materials. Table 1 displays mean performance in each group on all of these measures.

Cognitive Ability Measures

Verbal and nonverbal intelligence. The Vocabulary and Matrix Reasoning subtests from the Wechsler Abbreviated Scales of Intelligence (WASI; Wechsler, 1999) were used as indices of verbal and nonverbal ability. The raw scores on these subtests were converted into z-scores and summed to create a composite measure of intellectual ability for use in the analyses. A higher score indicated higher intelligence.

Executive function measures. Three measures of executive function were selected—a measure of set-shifting, a measure of inhibition, and a measure of working memory. The Trailmaking Test (Reitan, 1955, 1958) was used to measure set-shifting. Completion time is reported in Table 1. However, reaction time was significantly skewed, so a logarithmic transformation was used in subsequent analyses. Z-scores were computed from these logarithmically-transformed scores and the z-scores were then reflected so that higher scores indicated better set-shifting ability.

The Stroop Task was used to measure inhibition. The interference score is reported in Table 1. However, reaction time was significantly skewed, so a logarithmic transformation was used in subsequent analyses. Z-scores were computed from these logarithmically-transformed scores and the z-scores were then reflected so that higher scores indicated better ability to inhibit.

The working memory task selected was adapted from Gottardo, Stanovich, and Siegel (1996) and involved a variation of a sentence span task developed by Daneman and Carpenter (1980). Children listened to sets of two to five statements and indicate

Table 1
Mean Age and Performance on Intellectual Ability, Executive Function, Thinking Disposition Measures, and Rational Thinking Tasks

Measure	Grades 2–3 ($n = 54$)	Grades 4–5 ($n = 82$)	Grades 6–9 ($n = 68$)	$F(2, 201)$	Partial eta squared
Age (in months)	$104.5 (4.9)_a$	$120.5 (8.1)_b$	$152.5 (13.2)_c$	413.8^{***}	.81
Intelligence measures					
WASI Vocabulary raw score	$35.2 (5.4)_a$	$41.3 (6.2)_b$	$46.0 (7.5)_c$	42.0^{***}	.30
WASI Matrix Reasoning raw score	$20.1 (5.6)_a$	$22.9 (4.7)_b$	$25.1 (4.3)_c$	16.1^{***}	.14
WASI composite *z*-score	$-1.5 (1.4)_a$	$-0.2 (1.5)_b$	$0.9 (1.6)_c$	38.0^{***}	.27
Executive function measures					
Trailmaking Part B time	$134.3 (47.5)_a$	$121.0 (52.5)_a$	$85.5 (34.7)_b$	19.3^{***}	.16
Stroop Interference raw time	$43.9 (11.0)_a$	$38.5 (14.4)_a$	$28.7 (11.7)_b$	23.0^{***}	.19
Sentence Span Accuracy raw score	$18.9 (3.5)_a$	$19.6 (3.0)_a$	$21.0 (2.6)_b$	7.9^{***}	.07
Executive function composite *z*-score	$-1.3 (2.1)_a$	$-0.4 (2.1)_b$	$1.5 (2.1)_c$	29.9^{***}	.23
Thinking disposition measures					
AOT raw score	$44.1 (3.9)_a$	$45.9 (4.1)_b$	$48.1 (4.4)_c$	13.6^{***}	.12
Need for Cognition raw score	19.9 (3.0)	19.5 (3.0)	20.0 (2.4)	0.6	.01
Superstitious Thinking raw core	14.2 (2.4)	13.9 (2.9)	13.3 (2.9)	1.6	.02
AOT-ST composite *z*-score	$-0.6 (1.4)_a$	$-0.1 (1.6)_a$	$0.6 (1.7)_b$	9.0^{***}	.08
Rational thinking tasks					
Denominator neglect task	$3.3 (2.4)_a$	$4.0 (2.2)_{ab}$	$4.6 (2.1)_b$	4.8^{**}	.05
Belief bias syllogisms	$2.7 (1.5)_a$	$3.2 (1.6)_a$	$4.2 (1.6)_b$	15.6^{***}	.13
Resistance to framing	$-6.0 (6.0)_a$	$-3.4 (3.6)_b$	$-3.4 (4.6)_b$	6.2^{**}	.06
Base rate sensitivity	$2.4 (1.5)_a$	$3.3 (1.4)_b$	$3.6 (1.3)_b$	12.3^{***}	.11
Otherside thinking	$1.4 (0.8)_a$	$1.6 (1.1)_{a,b}$	$2.0 (1.0)_b$	5.9^{**}	.06
Rational thinking composite *z*-score	$-1.9 (2.7)_a$	$0.02 (2.7)_b$	$1.5 (2.7)_c$	24.1^{***}	.19
Parent rating of child's decision making					
Making good choices and decisions[a]	4.9 (0.9)	4.8 (1.2)	5.3 (1.2)	2.8	.03

Note. *F*-ratios refer to analysis of variance test of differences between grade groups; means with different letter subscripts indicate significant differences in post hoc analyses. WASI = Wechsler Abbreviated Scales of Intelligence; AOT = Actively Open-Minded Thinking scale; AOT-ST = composite of the AOT and Superstitious Thinking measures.
[a] Group sample sizes were slightly reduced for this measure/item $n = 191$ and $F(2, 188)$; Grades 2–3 ($n = 53$); Grades 4–5 ($n = 82$); Grades 6–9 ($n = 56$).
$^{**} p < .01$. $^{***} p < .001$.

whether each statement was true or false. After responding to each of the sentences in a set, the child was asked to recall the final word of each sentence in the set. Recall accuracy was the dependent measure. A higher score indicated better working memory ability.

The three executive function measures displayed moderate correlations: Stroop and Trailmaking ($r = .52$, $p < .001$), Stroop and sentence span ($r = .38$, $p < .001$), and Trailmaking and sentence span ($r = .47$, $p < .001$). An executive function composite score was calculated based on summing the *z*-scores of the three executive function measures. A higher score indicated better executive function performance.

Thinking Dispositions

Participants completed a self-report questionnaire that amalgamated items from three different thinking dispositions measures. They were asked to rate their agreement with each question using the following 4-point scale: *Strongly Agree* (1), *Agree* (2), *Disagree* (3), and *Strongly Disagree* (4). Questions from the three scales were presented in mixed order.

Actively Open-Minded Thinking Scale. A child version of the Actively Open-Minded Thinking scale (AOT) was used in this investigation (Kokis et al., 2002). A higher score indicated more actively open-minded thinking. The split-half reliability (Spearman–Brown corrected) of the scale was .60, and Cronbach's alpha was .59.

Superstitious Thinking Scale. A child version of the Superstitious Thinking Scale adapted from Kokis et al. (2002) was used. The split-half reliability (Spearman-Brown corrected) of the scale was .68, and Cronbach's alpha was .58. A higher score indicated more superstitious thinking.

Need for Cognition Scale. A child version of the Need for Cognition Scale was adapted from Kokis et al. (2002). The split-half reliability (Spearman-Brown corrected) of the scale was .63, and Cronbach's alpha was .55. A higher score indicated higher need for cognition.

Only one correlation between these three thinking dispositions was statistically significant, that between the AOT scale and Superstitious Thinking ($r = –.36$, $p < .001$). The Need for Cognition measure did not correlate with either the AOT or the Superstitious Thinking measure (.03 and .05, respectively). We formed a composite of the AOT and Superstitious Thinking measures (AOT-ST). For the purposes of the AOT-ST composite, the score on the Superstitious Thinking Scale was reflected, so that a higher score indicated less superstitious thinking. Both of these scores were converted into z-scores and summed for the composite AOT-ST scale.

Rational Thinking Tasks

Denominator neglect. This task was adapted from a previous children's version of the task (see Kokis et al., 2002). A total score was derived across six trials. Bowl 1 was the correct choice on each trial. The mean score across the six trials was 4.01 (SD = 2.29).

Belief bias syllogisms. Eight syllogisms were administered to participants, in which the content of the premises conflicted with the believability of the conclusion. Only inconsistent syllogisms (those where belief and validity were in conflict) were used, because previous developmental research has found that consistent syllogisms display ceiling performance (Kokis et al., 2002). The mean score across the eight items was 3.38 (SD = 1.68).

Resistance to framing. Three attribute framing situations were designed to be developmentally appropriate alternatives to tasks used with adult samples to study framing effects (Bruine de Bruin et al., 2007). In the within-subjects design in this study, each framing situation had a positive and negative frame that was presented in separate parts of the battery. Scoring was based on subtracting the positive frame rating from the corresponding negative frame rating of a situation. A higher score indicated more resistance to framing. The mean score across the three items was −4.10 (SD = 4.75). This mean score was significantly different from zero, $t(203) = –12.33$, $p < .001$, suggesting that participants were displaying a framing effect.

Base rate sensitivity. Five problems from Kokis et al. (2002) were used in which base rate information conflicted with individuating information. Reliance on probabilistic aggregate information was scored as correct, whereas reliance on concrete/personal information was scored as incorrect. A composite based on these five problems was generated (mean score = 3.16, SD = 1.44).

Otherside thinking task. A task that assessed the tendency toward otherside thinking was adapted for children based on the methods used by Toplak and Stanovich (2003). This task was administered by the examiner using an interview, and the examiner recorded responses provided by the participant. An issue about cell phones was used: "Do you think kids should have cell phones?" Participants were asked to provide their prior belief on this issue, followed by providing reasons consistent, then inconsistent, with their prior belief on this issue. A scoring scheme was developed for distinguishing between the conceptually unique reasons participants listed as being consistent with the pro and con sides on the issue. The mean number of myside reasons on the cell phone issue was 2.23 (SD = 0.94). The key measure, however, was the mean number of otherside reasons, which was 1.71 (SD = 0.99).

Parent Rating of Child Decision Making

Parents were asked to rate their children's decision making compared to other children of the same age based on observations over the past month on the following item: "Making good choices and decisions." Parents used a seven point scale ranging from: far below average, below average, slightly below average, average, slightly above average, above average, to far above average. A higher score indicated better decision-making performance. The sample size was 191 for this question. On a 7-point rating scale, the mean rating was 5.0 (SD = 1.1).

Procedure

Participants were tested individually by trained examiners. Most participants were able to independently read the rational thinking and thinking dispositions questionnaires. Some of the younger participants were slower and their reading was effortful. In these instances, the examiner orally read the questions to the participant. Parent consent and child assent were obtained for each participant. Task order was the same for each participant, and the order was as follows: demographics form; denominator neglect; belief bias syllogisms; WASI Vocabulary; WASI Matrices; resistance to framing part 1; otherside thinking task; Stroop; Trailmaking; sentence span; resistance to framing part 2; base rate sensitivity; and thinking dispositions questionnaire. Parents made the rating of their child's decision-making performance on the signed consent form.

Results

Table 1 displays the differences between the age groups on all the variables in the study. There were highly significant developmental differences in the measures of cognitive ability: both in the intelligence measures and in the executive functioning measures. The WASI Vocabulary and WASI Matrix measures of intelligence were highly related to age ($p < .001$ in both cases and variance explained .30 and .14, respectively). Likewise, the three measures of executive functioning were all significantly related to age group at the .001 level and explained from .07 to .19 of the variance. Table 1 indicates that the executive functioning composite score displayed a significant age effect, $F(2, 201) = 29.9$, $p < .001$, partial eta squared = .23. With respect to the thinking dispositions, the Actively Open-Minded Thinking scale displayed a significant age effect ($p < .001$, partial eta squared = .12), but the Need for Cognition and Superstitious Thinking scales did not. The AOT-ST composite displayed a significant age effect ($p < .001$), but age explained less variance on the composite than on the AOT task itself (.08 versus .12).

Appearing next in Table 1 are the means across the age groups for the rational thinking tasks. All five rational thinking tasks displayed significant age effects: denominator neglect, belief bias syllogisms, framing effects, base rate sensitivity, and otherside thinking. Performance increased with age in each case—that is, older children were less likely to display each of the cognitive biases. Developmental level explained .05, .06, and .06 of the variance for the first, third, and fifth of these ($p < .01$), but it explained substantially more of the variance (.13 and .11) in belief bias syllogisms and base rate sensitivity ($p < .001$ in both cases). A rational thinking composite score was formed by combining the scores on the five tasks. The scores were converted to z-scores and averaged in order to derive the composite. Table 1 indicates that this rational thinking composite score displayed a statistically significant effect of age, $F(2, 201) = 24.1$, $p < .001$, partial eta squared = .19. Finally, the effect of age on parents' report of their child's decision making approached, but did not achieve statistical significance, $F(2, 188) = 2.80$, $p = .062$, partial eta squared = .03.

Table 2 presents the zero-order correlations among the major variables in the study. Because of the fairly large sample size in the study, all correlations over .137 are significant at the .05 level (two-tailed). The five tasks in the rational thinking composite displayed positive manifold. All 10 correlations among the tasks were positive and eight were statistically significant. All of the tasks displayed statistically significant correlations with age, the WASI composite, the executive functioning composite, and the AOT-ST composite.

The rational thinking composite score was moderately correlated with age (.46) and even more strongly related to the measures of cognitive ability: WASI composite (.65) and executive functioning composite (.53). The rational thinking composite score displayed a significant .35 correlation with AOT-ST composite scale (it correlated with each of the components separately as well), and it displayed a significant but small correlation with the Need for Cognition scale (.15). Although the parents' rating of child's decision making was not related to age, it did display significant correlations with the WASI composite (.18), executive functioning composite (.23), Need for Cognition (.18), AOT-ST composite (.26), and the rational thinking composite (.20).

For readers who wish to contextualize these correlations using a familywise correction, we would suggest that a family of hypotheses in this experiment might be comprised of one rational think-

Table 2
Correlations Among All the Major Variables

Indexes and tasks	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Age	—												
Cognitive ability measures													
2. WASI composite *z*-score	.52	—											
3. Executive function composite	.47	.63	—										
Thinking dispositions													
4. Actively Open-Minded Thinking	.31	.40	.24	—									
5. Need for Cognition	.01	.18	.20	.03	—								
6. Superstitious Thinking	−.16	−.25	−.16	−.36	.05	—							
7. AOT-ST composite	.28	.40	.24	(.83)	−.01	(−.83)	—						
Rational thinking tasks													
8. Denominator neglect task	.27	.36	.30	.09	.20	−.16	.15	—					
9. Belief bias syllogisms	.37	.42	.33	.31	−.04	−.23	.33	.27	—				
10. Resistance to framing	.22	.40	.30	.19	.10	−.11	.18	.25	.17	—			
11. Base rate sensitivity	.29	.51	.34	.23	.06	−.21	.27	.26	.30	.23	—		
12. Otherside thinking	.25	.28	.31	.15	.13	−.07	.14	.12	.23	.07	.16	—	
13. Rational thinking composite	.46	.65	.53	.32	.15	−.26	.35	(.63)	(.65)	(.57)	(.64)	(.52)	—
Parent rating													
14. Parent rating–decision making	.12	.18	.23	.19	.18	−.24	.26	.08	.16	.17	.09	.10	.20

Note. $r = .137$, $p < .05$; $r = .180$, $p < .01$; $r = .229$, $p < .001$, two-tailed (for parent rating: $r = .142$, $p < .05$; $r = .186$, $p < .01$; $r = .236$, $p < .001$). Correlations in parentheses reflect part-whole relationships. WASI = Wechsler Abbreviated Scales of Intelligence; AOT-ST = composite of the Actively Open-Minded Thinking scale and Superstitious Thinking measures.

ing task crossed with each of the measures of cognitive sophistication (age, WASI composite, executive function composite, Actively Open-Minded Thinking, Need for Cognition, Superstitious Thinking). The critical value for a significant two-tailed correlation at the .05 level would be adjusted from .137 to .183.

As Table 2 indicates, the rational thinking composite (and each of the five tasks in it) was related to age and to the WASI and executive function composite. However, age had a substantial correlation with both cognitive ability measures and a somewhat more modest correlation with some thinking dispositions. It is important to know whether the individual difference measures correlate with rational thinking skill separately from age. The following analyses explore this question.

Table 3 displays the correlations with the rational thinking composite within each of the age groupings defined previously. The restriction of range in each of the groups has markedly reduced the correlation with age (to nonsignificance in two of the three cases). However, the correlations with the cognitive ability measures (WASI and executive functioning) remained almost as large in the age-restricted samples as they were in the entire sample. The same was true of the AOT-ST composite. Finally the parent rating of child decision making was positively correlated with the rational thinking composite in all three cases (and significantly so in one).

At the top of Table 4 are the results of a simultaneous regression in which age, the AOT-ST composite, WASI composite, and executive function composite were used to predict the score on the rational thinking composite. These four predictors accounted for substantial variance (multiple $R = .69$, multiple $R^2 = .47$). The dominant predictor was the WASI composite, which explained 9.9% unique variance and had a standardized beta weight of .447, $t(199) = 6.10$, $p < .001$. Age and the executive function composite explained small but statistically significant proportions of unique variance—age: $\beta = .123$, unique variance explained = 1.0%, $t(199) = 1.97$, $p = .05$; executive function composite: $\beta = .162$, unique variance explained = 1.5%, $t(199) = 2.38$, $p < .05$.

The WASI composite and executive functioning composite had a strong association ($r = .63$ in Table 2). Thus, the WASI composite might be stealing quite a bit of variance from the executive functioning variable in the regression analysis, likely obscuring the latter's separate predictive power. In the next simultaneous regression in Table 4, the WASI composite was removed. This lowered the multiple R from .68 to .61 but, as expected, it resulted in a significant beta weight of .369 for the executive function composite: $t(199) = 5.76$, $p < .001$, 10.4% unique variance explained. Although the executive functioning composite was the dominant predictor, both age (.233) and the AOT-ST composite (.197) displayed significant beta weights ($p < .001$) in this analysis.

The remaining simultaneous regressions displayed in Table 4 look at each of the five rational thinking tasks as criterion variables in turn. All four of the predictors (age, AOT-ST composite, WASI composite, executive function composite) were used in the regression equations. For three of the five tasks (denominator neglect, base rate sensitivity, framing), the WASI composite was the only significant unique predictor. The three other variables did not predict significant unique variance in these three tasks. The results from the belief bias task were different, however. There, three variables had significant beta weights (age, AOT-ST, and WASI composite) and each predicted roughly equal amounts of unique variance in belief bias syllogisms (2.1%, 2.6%, and 2.1%, respectively). Finally, otherside thinking was predicted uniquely only by the executive function composite: $\beta = .201$, unique variance explained = 2.3%, $t(199) = 2.30$, $p < .05$.

Discussion

The results of this study are, perhaps surprisingly, rather easy to describe in a global sense. All three measures of cognitive sophistication converged in their relationships (or lack of relationship) with the rational thinking measures studied here. All of the tasks were significantly related to age and the other measures of cognitive sophistication. This pattern was indicated in both the discrete analyses displayed in Table 1 and in the continuous correlational analyses displayed in Table 2.

In the discrete analyses (see Table 1) the five rational thinking tasks were each significant at least at the .01 level and age explained from .05 to .13 percent of the variance in them. Performance improved with age. In Table 2, the five rational thinking tasks displayed correlations with age ranging from .22 to .37 (all statistically significant). These same tasks displayed similar correlations with the executive function composite (correlations from .30 to .34) and even higher correlations with the WASI composite (correlations from .28 to .51). All five correlations with the AOT-ST composite were statistically significant (correlations from .14 to .33). In short, these five rational thinking tasks displayed similar relationships with age, cognitive ability, and thinking dispositions. One additional, but more subtle indicator of convergence, is parent's rating of their child's decision-making

Table 3
Rational Thinking Composite and Individual Difference Correlations by Developmental Level

Grade	n	Age	WASI composite	Executive function composite	AOT-ST composite	Need for cognition	Parent-rating decision[d]
Grades 2–3[a]	54	.03	.35	.45	.17	.09	.15
Grades 4–5[b]	82	.14	.50	.39	.24	.14	.26
Grades 6–9[c]	68	.34	.76	.40	.36	.25	.07

Note. WASI = Wechsler Abbreviated Scales of Intelligence; AOT-ST = composite of the Actively Open-Minded Thinking scale and Superstitious Thinking measures.
[a] Values greater than .269 are significant at the $p < .05$ level, one-tailed. [b] Values greater than .218 are significant at the $p < .05$ level, one-tailed. [c] Values greater than .239 are significant at the $p < .05$ level, one-tailed. [d] Parent-rating decisions: Grades 2–3 ($n = 53$), Grades 4–5 ($n = 82$), and Grades 6–9 ($n = 56$).

Table 4
Regression Results

Variable	Standardized β	t(199 or 200)	Unique variance explained
Criterion variable = Rational Thinking composite z-score			
Age	.123	1.97*	.010
AOT-ST composite z-score	.101	1.78	.008
WASI composite z-score	.447	6.10***	.099
Executive function composite z-score	.162	2.38*	.015
Overall regression: $F(4, 199) = 43.97$***			
Multiple $R = .69$			
Multiple $R^2 = .47$			
Age	.233	3.59***	.041
AOT-ST composite z-score	.197	3.33***	.035
Executive function composite z-score	.369	5.76***	.104
Overall regression: $F(3, 200) = 39.15$***			
Multiple $R = .61$			
Multiple $R^2 = .37$			
Criterion variable = Denominator neglect task			
Age	.087	1.09	.005
AOT-ST composite z-score	.006	0.09	.000
WASI composite z-score	.246	2.64**	.030
Executive function composite z-score	.103	1.19	.006
Overall regression: $F(4, 199) = 8.27$***			
Multiple $R = .38$			
Multiple $R^2 = .14$			
Criterion variable = Belief bias syllogistic reasoning			
Age	.176	2.36*	.021
AOT-ST composite z-score	.175	2.58*	.026
WASI composite z-score	.208	2.36*	.021
Executive function composite z-score	.075	0.92	.003
Overall regression: $F(4, 199) = 15.16$***			
Multiple $R = .48$			
Multiple $R^2 = .23$			
Criterion variable = Base rate sensitivity			
Age	.018	0.25	.000
AOT-ST composite z-score	.073	1.10	.004
WASI composite z-score	.458	5.32***	.104
Executive function composite z-score	.028	0.35	.000
Overall regression: $F(4, 199) = 18.33$***			
Multiple $R = .52$			
Multiple $R^2 = .27$			
Criterion variable = Resistance to framing			
Age	−.004	−0.05	.000
AOT-ST composite z-score	.028	0.39	.001
WASI composite z-score	.338	3.67***	.057
Executive function composite z-score	.081	0.95	.004
Overall regression: $F(4, 199) = 9.66$***			
Multiple $R = .40$			
Multiple $R^2 = .16$			
Criterion variable = Otherside thinking			
Age	.093	1.16	.006
AOT-ST composite z-score	.022	0.30	.0004
WASI composite z-score	.100	1.06	.005
Executive function composite z-score	.201	2.30*	.023
Overall regression: $F(4, 199) = 6.61$***			
Multiple $R = .34$			
Multiple $R^2 = .12$			

Note. AOT-ST = composite of the Actively Open-Minded Thinking scale and Superstitious Thinking measures; WASI = Wechsler Abbreviated Scales of Intelligence.
* $p < .05$. ** $p < .01$. *** $p < .001$.

competence which also displayed a positive association with the rational thinking composite. Looking at cognitive ability and thinking dispositions within more homogeneous age groups (see Table 3) produced largely converging conclusions. Measures of cognitive ability and thinking dispositions were moderately correlated with the rational thinking composite within each of the age groups.

When all of the measures of cognitive sophistication were used as predictors of the rational thinking composite, the WASI composite was the dominant predictor. Largely, it appears that age is just a proxy for cognitive ability, because the former does not account very much variance (1.0%) once the latter is in the regression equation (except for the belief bias task). However, when controlling for age, the WASI composite remains a significant predictor (see Tables 3 and 4).

To understand the connection between performance on these rational thinking tasks and cognitive sophistication, we will build here on a general framework for understanding individual differences in heuristics and biases task performance introduced by Stanovich and West (2008b). That framework posits that performance on rational thinking tasks correlates with cognitive ability under two different circumstances. Both circumstances follow from the fact that dual process models (Evans, 2008; Evans & Stanovich, 2013; Kahneman, 2011; Stanovich, 2004, 2011) posit that an important function of Type 2 processes is to take early representations triggered by Type 1 processing offline and to substitute better responses. This is because most of the tasks in the heuristics and biases literature were deliberately designed to pit an automatically-triggered, natural response against a normative response generated by Type 2 processing (Kahneman, 2011). To successfully complete the task, a person must inhibit the natural response by decoupling it and then must substitute an alternative response based on stored mindware. Mindware is a term coined by Perkins (1995) to refer to the rules, knowledge, procedures, and strategies that a person can retrieve from memory in order to aid decision making. Decoupling occurs when the automatic input-output connections of the brain are suspended while cognitive simulations of hypotheticals are carried out (Stanovich, 2009, 2011).

Correlations of performance on these tasks and cognitive sophistication thus arise for two reasons. First, for successful Type 2 override to occur, a person must have the mindware available. That is, more normative declarative knowledge must be available to substitute for an automatic response that happens to be nonnormative. If the task requires specific mindware (probabilistic thinking skills, falsifiability tendencies, sensitivity to contradiction, etc.) that is not present, then people will rely on the Type 1 response. Mindware is acquired with development, acquired more rapidly by those higher in cognitive ability, and acquired more flexibly and efficiently by those high in dispositions such as actively open-minded thinking. Thus, all the forms of cognitive sophistication result in more relevant mindware being available and less chance of a biased response.

The second way in which cognitive sophistication (particularly cognitive ability) comes to be correlated with performance on rational thinking tasks concerns the situation where the necessity for override has been detected and the relevant mindware is available. The conflict still must be resolved, and this may require cognitive capacity, especially if cognitive decoupling must take place for a considerable period of time. Recent work on inhibition and executive functioning has indicated that such cognitive decoupling is very capacity demanding and that it is strongly related to individual differences in fluid intelligence (Duncan et al., 2012; Kane et al., 2005; Salthouse et al., 2003; Unsworth & Engle, 2007).

In this respect, our framework converges with that of fuzzy trace theory, which has long emphasized inhibition of task interference as an account of associations between development and task performance (see Reyna, 1995; Reyna & Brainerd, 1995). Interference is present in all of the tasks showing associations with cognitive sophistication in our study. The development of the monitoring/inhibition component of fuzzy trace theory would limit the impact of the interference and hence enhance performance according to that theory (Reyna, 1995), thus creating associations with development in all of the tasks. This inhibition of interference would be dependent upon the decoupling mechanism in our theoretical framework.

Several of our tasks seem to necessitate the interference/inhibition of fuzzy trace theory (Reyna, 1995) or the decoupling operation in our version of dual process theory (Stanovich, 2011). Within-subjects framing paradigms like that used in the present study exemplify this logic of inhibition/decoupling. The appearance of the second problem provides a strong signal that an issue of consistency is at stake. The tendency to provide responses triggered by the positive or negative frame must be suppressed (by capacity demanding decoupling operations) in order to respond more in line with the formal equivalence of the two versions. This is what probably accounts for the linkage between performance and cognitive ability observed on this task.

It is likewise with the denominator neglect task. It is a within-subjects paradigm in which the two conditions are directly compared by the participant. The visual attractiveness of the more numerous (but lower probability) bowl must be suppressed for the small (but higher probability) bowl to be chosen. The decoupling involved in this inhibitory operation probably accounts for the linkage between performance and cognitive ability observed on this task.

It may well be that a parallel explanation is also what accounts for the correlation observed on the base rate sensitivity task. The two conflicting pieces of information—the aggregate base rate information and the personal experience—are presented, within-subjects, right in front of the participant. However, in this task the other class of factor creating correlations with cognitive ability might be in play. Specifically, there may be a mindware gap (see Stanovich, West, & Toplak, 2011b) that is negatively correlated with age. Children need to learn the importance of aggregate information. Participants higher in cognitive ability might be more apt to learn this information. Thus, there may be two reasons for a cognitive ability correlation with base rate sensitivity: a decoupling requirement and a mindware acquisition requirement. It is perhaps because of these dual reasons that base rate sensitivity shows a higher correlation with the WASI composite than the other tasks in the rationality composite.

In summary, with respect to five important tasks and effects in the rational thinking literature—denominator neglect, belief bias syllogisms, base rate sensitivity, resistance to framing, and otherside thinking—we have observed convergent results. All three measures of cognitive sophistication (age, cognitive ability, and thinking dispositions) correlated with these tasks and effects. The trends we have observed here are consistent with some earlier attempts at taxonomies of rational thought (Reyna, Lloyd, & Brainerd, 2003; Stanovich, 2009, 2011; Stanovich et al., 2008). Nevertheless, tracing all of the important developmental trends will be a lengthy task because rational thought spans many domains. Understanding the development of rationality will clearly be a tall order. It will be worth the effort, and not just for scientific reasons. Assumptions about the nature and development of rationality are implicated in judgments of legal responsibility. Reyna and Farley (2006) have emphasized how background assumptions about adolescent rationality frame efforts to change adolescent risk behavior.

References

Aron, A. R. (2008). Progress in executive-function research: From tasks to functions to regions to networks. *Current Directions in Psychological Science, 17,* 124–129. doi:10.1111/j.1467-8721.2008.00561.x

Bar-Hillel, M. (1980). The base-rate fallacy in probability judgments. *Acta Psychologica, 44,* 211–233. doi:10.1016/0001-6918(80)90046-3

Bar-Hillel, M. (1990). Back to base rates. In R. M. Hogarth (Ed.), *Insights into decision making: A tribute to Hillel J. Einhorn* (pp. 200–216). Chicago, IL: University of Chicago Press.

Baron, J. (1995). Myside bias in thinking about abortion. *Thinking and Reasoning, 1,* 221–235. doi:10.1080/13546789508256909

Baron, J. (2008). *Thinking and deciding* (4th ed.). Cambridge, MA: Cambridge University Press.

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review, 29,* 180–200. doi:10.1016/j.dr.2009.05.002

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78,* 246–263. doi:10.1111/j.1467-8624.2007.00995.x

Brainerd, C. J., Reyna, V. F., & Zember, E. (2011). Theoretical and forensic implications of developmental studies of the DRM illusion. *Memory & Cognition, 39,* 365–380. doi:10.3758/s13421-010-0043-2

Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology, 92,* 938–956. doi:10.1037/0022-3514.92.5.938

Cacioppo, J. T., Petty, R. E., Feinstein, J., & Jarvis, W. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin, 119,* 197–253. doi:10.1037/0033-2909.119.2.197

Chatterjee, S., Heath, T. B., Milberg, S. J., & France, K. R. (2000). The differential processing price in gains and losses: The effects of frame and need for cognition. *Journal of Behavioral Decision Making, 13,* 61–75. doi:10.1002/(SICI)1099-0771(200001/03)13:1<61::AID-BDM343>3.0.CO;2-J

Chien, Y. C., Lin, C., & Worthley, J. (1996). Effect of framing on adolescents' decision making. *Perceptual and Motor Skills, 83,* 811–819. doi:10.2466/pms.1996.83.3.811

Chiesi, F., Primi, C., & Morsanyi, K. (2011). Developmental changes in probabilistic reasoning: The role of cognitive capacity, instructions, thinking styles, and relevant knowledge. *Thinking & Reasoning, 17,* 315–350. doi:10.1080/13546783.2011.598401

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19,* 450–466. doi:10.1016/S0022-5371(80)90312-6

Davidson, D. (1995). The representativeness heuristic and the conjunction fallacy effect in children's decision making. *Merrill-Palmer Quarterly, 41,* 328–346.

De Neys, W. (2006). Dual processing in reasoning—Two systems but one reasoner. *Psychological Science, 17,* 428–433. doi:10.1111/j.1467-9280.2006.01723.x

De Neys, W., & Vanderputte, K. (2011). When less is not always more: Stereotype knowledge and reasoning development. *Developmental Psychology, 47,* 432–441. doi:10.1037/a0021313

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16,* 939–944. doi:10.1111/j.1467-9280.2005.01641.x

Duncan, J., Parr, A., Woolgar, A., Thompson, R., Bright, P., Cox, S., . . . Nimmo-Smith, I. (2008). Goal neglect and Spearman's *g*: Competing parts of a complex task. *Journal of Experimental Psychology: General, 137,* 131–148. doi:10.1037/0096-3445.137.1.131

Duncan, J., Schramm, M., Thompson, R., & Dumontheil, I. (2012). Task rules, working memory, and fluid intelligence. *Psychonomic Bulletin & Review, 19,* 864–870. doi:10.3758/s13423-012-0225-y

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

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

Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8,* 223–241. doi:10.1177/1745691612460685

Finucane, M. L., & Gullion, C. M. (2010). Developing a tool for measuring the decision-making competence of older adults. *Psychology and Aging, 25,* 271–288. doi:10.1037/a0019106

Friedman, N. P., Haberstick, B. C., Willcutt, E. G., Miyake, A., Young, S. E., Corley, R. P., & Hewitt, J. K. (2007). Greater attention problems during childhood predict poorer executive functioning in late adolescence. *Psychological Science, 18,* 893–900. doi:10.1111/j.1467-9280.2007.01997.x

Gottardo, A., Stanovich, K. E., & Siegel, L. S. (1996). The relationships between phonological sensitivity, syntactic processing, and verbal working memory in the reading performance of third-grade children. *Journal of Experimental Child Psychology, 63,* 563–582. doi:10.1006/jecp.1996.0062

Handley, S. J., Capon, A., Beveridge, M., Dennis, I., & Evans, J. S. B. T. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking & Reasoning, 10,* 175–195. doi:10.1080/13546780442000051

Hasher, L., Lustig, C., & Zacks, R. (2007). Inhibitory mechanisms and the control of attention. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in working memory* (pp. 227–249). doi:10.1093/acprof:oso/9780195168648.003.0009

Jacobs, J. E., & Potenza, M. (1991). The use of judgment heuristics to make social and object decisions: A developmental perspective. *Child Development, 62,* 166–178. doi:10.1111/j.1467-8624.1991.tb01522.x

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58,* 697–720. doi:10.1037/0003-066X.58.9.697

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

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). doi:10.1017/CBO9780511808098.004

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist, 64,* 515–526. doi:10.1037/a0016755

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251. doi:10.1037/h0034747

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39,* 341–350. doi:10.1037/0003-066X.39.4.341

Kahneman, D., & Tversky, A. (Eds.). (2000). *Choices, values, and frames*. Cambridge, England: Cambridge University Press.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9,* 637–671. doi:10.3758/BF03196323

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47–70. doi:10.1037/0096-3445.132.1.47

Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131,* 66–71. doi:10.1037/0033-2909.131.1.66

Klaczynski, P. A. (2001). Analytic and heuristic processing influences on adolescent reasoning and decision making. *Child Development, 72,* 844–861. doi:10.1111/1467-8624.00319

Klaczynski, P. A., & Robinson, B. (2000). Personal theories, intellectual ability, and epistemological beliefs: Adult age differences in everyday reasoning tasks. *Psychology and Aging, 15,* 400–416. doi:10.1037/0882-7974.15.3.400

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. doi:10.1016/S0022-0965(02)00121-2

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing the mind: "Seizing" and "freezing". *Psychological Review, 103,* 263–283. doi:10.1037/0033-295X.103.2.263

Kühberger, A. (1998). The influence of framing on risky decisions: A meta-analysis. *Organizational Behavior and Human Decision Processes, 75,* 23–55. doi:10.1006/obhd.1998.2781

Levin, I. P., Gaeth, G. J., Schreiber, J., & Lauriola, M. (2002). A new look at framing effects: Distribution of effect sizes, individual differences, and independence of types of effects. *Organizational Behavior and Human Decision Processes, 88,* 411–429. doi:10.1006/obhd.2001.2983

Levin, I. P., & Hart, S. S. (2003). Risk preferences in young children: Early evidence of individual differences in reaction to potential gains and losses. *Journal of Behavioral Decision Making, 16,* 397–413. doi:10.1002/bdm.453

Levin, I. P., Hart, S. S., Weller, J. A., & Harshman, L. A. (2007). Stability of choices in a risky decision-making task: A 3-year longitudinal study with children and adults. *Journal of Behavioral Decision Making, 20,* 241–252. doi:10.1002/bdm.552

Li, M., & Chapman, G. B. (2009). "100% of anything looks good": The appeal of one hundred percent. *Psychonomic Bulletin & Review, 16,* 156–162. doi:10.3758/PBR.16.1.156

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making, 25,* 361–381. doi:10.1002/bdm.752

Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory & Cognition, 17,* 11–17. doi:10.3758/BF03199552

Maule, J., & Villejoubert, G. (2007). What lies beneath: Reframing framing effects. *Thinking & Reasoning, 13,* 25–44. doi:10.1080/13546780600872585

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science, 21,* 8–14. doi:10.1177/0963721411429458

Morsanyi, K., & Handley, S. J. (2008). How smart do you need to be to get it wrong? The role of cognitive capacity in the development of heuristic-based judgment. *Journal of Experimental Child Psychology, 99,* 18–36. doi:10.1016/j.jecp.2007.08.003

Parker, A. M., & Fischhoff, B. (2005). Decision-making competence: External validation through an individual differences approach. *Journal of Behavioral Decision Making, 18,* 1–27. doi:10.1002/bdm.481

Paul, R. W. (1984). Critical thinking: Fundamental to education for a free society in North America: A new theory of knowledge, learning, and literacy. *Educational Leadership, 42,* 4–14. doi:10.1007/BF00128149

Paul, R. W. (1987). Critical thinking and the critical person. In D. N. Perkins, J. Lockhead, & J. Bishop (Eds.), *Thinking: The second international conference* (pp. 373–403). doi:10.1177/019263659107553325

Perkins, D. N. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology, 77,* 562–571. doi:10.1037/0022-0663.77.5.562

Perkins, D. N. (1995). *Outsmarting IQ: The emerging science of learnable intelligence*. New York, NY: Free Press.

Perkins, D. N., Farady, M., & Bushey, B. (1991). Everyday reasoning and the roots of intelligence. In J. Voss, D. Perkins, & J. Segal (Eds.), *Informal reasoning and education* (pp. 83–105). Hillsdale, NJ: Erlbaum.

Reitan, R. M. (1955). The relation of the Trail Making Test to organic brain damage. *Journal of Consulting Psychology, 19,* 393–394. doi:10.1037/h0044509

Reitan, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor Skills, 8,* 271–276. doi:10.2466/PMS.8.7.271-276

Reyna, V. F. (1995). Interference effects in memory and reasoning: A fuzzy-trace theory analysis. In F. N. Dempster & C. J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 29–59). doi:10.1016/B978-012208930-5/50003-9

Reyna, V. F., & Brainerd, C. J. (1994). The origins of probability judgment: A review of data and theories. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 239–272). New York, NY: Wiley.

Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences, 7,* 1–75. doi:10.1016/1041-6080(95)90031-4

Reyna, V. F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences, 18,* 89–107. doi:10.1016/j.lindif.2007.03.011

Reyna, V. F., & Brainerd, C. J. (2011). Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. *Developmental Review, 31,* 180–206. doi:10.1016/j.dr.2011.07.004

Reyna, V., Chapman, S., Dougherty, M., & Confrey, J. (Eds.). (2012). *The adolescent brain: Learning, reasoning, and decision making*. doi:10.1037/13493-000

Reyna, V. F., & Ellis, S. (1994). Fuzzy-trace theory and framing effects in children's risky decision making. *Psychological Science, 5,* 275–279. doi:10.1111/j.1467-9280.1994.tb00625.x

Reyna, V. F., Estrada, S., DeMarinis, J., Myers, R., Stanisz, J., & Mills, B. (2011). Neurobiological and memory models of risky decision making in adolescents versus young adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1125–1142. doi:10.1037/a0023943

Reyna, V. F., & Farley, F. (2006). Risk and rationality in adolescent decision making: Implications for theory, practice, and public policy. *Psychological Science in the Public Interest, 7,* 1–44. doi:10.1111/j.1529-1006.2006.00026.x

Reyna, V. F., Lloyd, F. J., & Brainerd, C. J. (2003). Memory, development, and rationality: An integrative theory of judgment and decision making. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 201–245). doi:10.1017/CBO9780511609978.009

Sá, W., Kelley, C., Ho, C., & Stanovich, K. E. (2005). Thinking about personal theories: Individual differences in the coordination of theory and evidence. *Personality and Individual Differences, 38,* 1149–1161. doi:10.1016/j.paid.2004.07.012

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132,* 566–594. doi:10.1037/0096-3445.132.4.566

Schlottmann, A., & Tring, J. (2005). How children reason about gains and losses: Framing effects in judgment and choice. *Swiss Journal of Psychology, 64,* 153–171. doi:10.1024/1421-0185.64.3.153

Simoneau, M., & Markovits, H. (2003). Reasoning with premises that are not empirically true: Evidence for the role of inhibition and retrieval. *Developmental Psychology, 39,* 964–975. doi:10.1037/0012-1649.39.6.964

Sinatra, G. M., & Pintrich, P. R. (Eds.). (2003). *Intentional conceptual change*. Mahwah, NJ: Erlbaum.

Slovic, P., & Peters, E. (2006). Risk perception and affect. *Current Directions in Psychological Science, 15,* 322–325. doi:10.1111/j.1467-8721.2006.00461.x

Smith, S. M., & Levin, I. P. (1996). Need for cognition and choice framing effects. *Journal of Behavioral Decision Making, 9,* 283–290. doi:10.1002/(SICI)1099-0771(199612)9:4<283::AID-BDM241>3.0.CO;2-7

Stanovich, K. E. (2004). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago, IL: University of Chicago Press. doi:10.7208/chicago/9780226771199.001.0001

Stanovich, K. E. (2009). *What intelligence tests miss: The psychology of rational thought*. New Haven, CT: Yale University Press.

Stanovich, K. E. (2011). *Rationality and the reflective mind*. New York, NY: Oxford University Press.

Stanovich, K. E., Toplak, M. E., & West, R. F. (2008). The development of rational thought: A taxonomy of heuristics and biases. *Advances in Child Development and Behavior, 36,* 251–285. doi:10.1016/S0065-2407(08)00006-2

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. doi:10.1037/0022-0663.89.2.342

Stanovich, K. E., & West, R. F. (1998). Individual differences in rational thought. *Journal of Experimental Psychology: General, 127,* 161–188. doi:10.1037/0096-3445.127.2.161

Stanovich, K. E., & West, R. F. (2007). Natural myside bias is independent of cognitive ability. *Thinking & Reasoning, 13,* 225–247. doi:10.1080/13546780600780796

Stanovich, K. E., & West, R. F. (2008a). On the failure of cognitive ability to predict myside bias and one-sided thinking bias. *Thinking & Reasoning, 14,* 129–167. doi:10.1080/13546780701679764

Stanovich, K. E., & West, R. F. (2008b). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94,* 672–695. doi:10.1037/0022-3514.94.4.672

Stanovich, K. E., West, R. F., & Toplak, M. E. (2011a). The complexity of developmental predictions from dual process models. *Developmental Review, 31,* 103–118. doi:10.1016/j.dr.2011.07.003

Stanovich, K. E., West, R. F., & Toplak, M. E. (2011b). Intelligence and rationality. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 784–826). doi:10.1017/CBO9780511977244.040

Sternberg, R. J. (1997). *Thinking styles*. doi:10.1017/CBO9780511584152

Sternberg, R. J. (2001). Why schools should teach for wisdom: The balance theory of wisdom in educational settings. *Educational Psychologist, 36,* 227–245. doi:10.1207/S15326985EP3604_2

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized*. doi:10.1017/CBO9780511509612

Strathman, A., Gleicher, F., Boninger, D. S., & Edwards, S. (1994). The consideration of future consequences: Weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology, 66,* 742–752. doi:10.1037/0022-3514.66.4.742

Toplak, M. E., & Stanovich, K. E. (2003). Associations between myside bias on an informal reasoning task and amount of post-secondary education. *Applied Cognitive Psychology, 17,* 851–860. doi:10.1002/acp.915

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. doi:10.3758/s13421-011-0104-1

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114,* 104–132. doi:10.1037/0033-295X.114.1.104

Wang, L. (2009). Money and fame: Vividness effects in the National Basketball Association. *Journal of Behavioral Decision Making, 22,* 20–44. doi:10.1002/bdm.609

Wechsler. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: The Psychological Corporation.

Weller, J. A., Levin, I. P., & Denburg, N. L. (2011). Trajectory of risky decision making for potential gains and losses from ages 5 to 85. *Journal of Behavioral Decision Making, 24,* 331–344. doi:10.1002/bdm.690

West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology, 100,* 930–941. doi:10.1037/a0012842

Westen, D., Blagov, P., Kilts, C., & Hamann, S. (2006). Neural bases of motivated reasoning: An fMRI study of emotional constraints on partisan political judgment in the 2004 U.S. presidential election. *Journal of Cognitive Neuroscience, 18,* 1947–1958. doi:10.1162/jocn.2006.18.11.1947

Zelazo, P. D. (2004). The development of conscious control in childhood. *Trends in Cognitive Sciences, 8,* 12–17. doi:10.1016/j.tics.2003.11.001

Received November 21, 2012
Revision received September 9, 2013
Accepted September 13, 2013 ■