Journal of Educational Psychology
2008, Vol. 100, No. 4, 930–941

Copyright 2008 by the American Psychological Association
0022-0663/08/$12.00 DOI: 10.1037/a0012842

Heuristics and Biases as Measures of Critical Thinking: Associations with Cognitive Ability and Thinking Dispositions

Richard F. West
James Madison University

Maggie E. Toplak
York University

Keith E. Stanovich
University of Toronto

In this article, the authors argue that there are a range of effects usually studied within cognitive psychology that are legitimately thought of as aspects of critical thinking: the cognitive biases studied in the heuristics and biases literature. In a study of 793 student participants, the authors found that the ability to avoid these biases was moderately correlated with a more traditional laboratory measure of critical thinking—the ability to reason logically when logic conflicts with prior belief. The correlation between these two classes of critical thinking skills was not due to a joint connection with general cognitive ability because it remained statistically significant after the variance due to cognitive ability was partialed out. Measures of thinking dispositions (actively open-minded thinking and need for cognition) predicted unique variance in both classes of critical thinking skills after general cognitive ability had been controlled.

Keywords: critical thinking, heuristics and biases, intelligence, thinking dispositions

In the critical thinking literature, the ability to evaluate evidence and arguments independently of one's prior beliefs and opinions is a skill that is strongly emphasized (Baron, 1991, 2000; Ennis, 1987, 1996; Perkins, 1995; Sternberg, 1997, 2001, 2003). Virtually all measures of critical thinking try to assess the ability to avoid reasoning that is too biased by prior opinion and prior belief (e.g., Ennis, Millman, & Tomko, 1985; Facione, 1992; Norris & Ennis, 1989; Watson & Glaser, 1980).

The Watson-Glaser Critical Thinking Assessment (WGCTA; Watson & Glaser, 1980) is widely used and is representative of the discourse-logic-based critical thinking measures that have played a prominent role in critical thinking assessment. The WGCTA is an 80-item multiple choice test that comprises five subsets of items (Inference, Recognition of Assumptions, Deduction, Interpretation, and Evaluation of Arguments). Each test item generally consists of a series of statements about which the validity of various conclusions must be judged. Four of the WGCTA's five subscales emphasize the propositional logic of necessity rather than sufficiency. The test's authors acknowledge that strong opinions and beliefs have the potential to adversely influence people's ability to think critically, and the WGCTA includes both neutral items and items on issues that are assumed likely to provoke a pressure to be biased. Thus, four of the five subscales admonish the test-taker to regard even factually questionable statements as correct for the purposes of the test (e.g., "For the purposes of this test, consider the statements in each exercise as true without exception"; "Try not to let your prejudices influence your judgment—just stick to the given statements [premises] and judge each conclusion as to whether it necessarily follows from the premises" [Harcourt Assessment, 2006, p. 4]). This emphasis on the importance of avoiding unbiased reasoning is also a common feature of the other discourse-logic-based critical thinking measures (e.g., Cornell Critical Thinking Tests, Ennis et al., 1985; California Critical Thinking Skills Test, Facione, 1992), and shares a common motivational underpinning with a number of tasks examined by cognitive scientists. Historically, Piaget's (1972) conceptualization of formal operational thought places such mechanisms of decontextualization—freeing from irrelevant context—in positions of paramount importance, because according to his view, "one of the essential characteristics of formal thought appears to us to be the independence of its form from reality content" (p. 10).

In the laboratory, the ability to reason in an unbiased manner has been operationalized with a few well-known paradigms. The belief bias syllogism task, one of the most thoroughly explored of these paradigms, shares a number of important features with discourse-logic-based critical thinking measures. This paradigm assesses the tendency for judgments of logical validity to be contaminated by prior knowledge of the world—for example, when the validity of a syllogism and the facts expressed in the conclusion of the syllogism conflict (e.g., "All flowers have petals; roses have petals; therefore, roses are flowers"—which is invalid under necessity). The inability to decouple prior knowledge from reasoning processes has been termed the *belief bias effect* (Evans, Barston, & Pollard, 1983). It has been the subject of extensive study in the cognitive science literature, and several formal models of how belief bias operates to disrupt syllogis-

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

This research was supported by a grant from the Social Sciences and Humanities Research Council of Canada to Maggie E. Toplak and by grants from the Social Sciences and Humanities Research Council of Canada and the Canada Research Chairs program to Keith E. Stanovich.

Correspondence concerning this article should be addressed to Richard F. West, Department of Graduate Psychology, MSC 7401, James Madison University, Harrisonburg, VA 22807. E-mail: westrf@jmu.edu

tic reasoning have been proposed (De Neys, 2006; Evans & Curtis-Holmes, 2005; Evans & Feeney, 2004; Garnham & Oakhill, 2005; Klauer, Musch, & Naumer, 2000).

Belief bias has also been revealed in paradigms in which participants must evaluate the quality of empirical evidence in a manner not contaminated by their prior opinion on the issue in question. In several studies, Klaczynski (1997) and colleagues (Klaczynski & Gordon, 1996; Klaczynski & Lavallee, 2005; Klaczynski & Robinson, 2000) presented participants with flawed hypothetical experiments that led to conclusions that were either consistent or inconsistent with prior positions and opinions. Participants then critiqued the flaws in the experiments (which were most often badly flawed). Participants found many more flaws when the experiment's conclusions were inconsistent with their prior opinions than when the experiment's conclusions were consistent with their prior opinions and beliefs.

It is, of course, important to show that the ability to reason independently of prior opinion is not entirely coexistent with general cognitive ability (intelligence), and there has been some preliminary evidence indicating that this is in fact the case. The tendency toward biased reasoning in the experiment evaluation paradigms (as well as related paradigms) shows considerable dissociation from cognitive ability (Kardash & Scholes, 1996; Klaczynski & Gordon, 1996; Klaczynski & Lavallee, 2005; Klaczynski & Robinson, 2000; Macpherson & Stanovich, 2007). Belief bias in syllogistic reasoning has shown a significant correlation with cognitive ability, but it is modest in size and does not exhaust the reliable variance in the magnitude of the bias displayed. We know the latter because various thinking dispositions (actively open-minded thinking, need for cognition) have been found to predict belief bias after the variance in cognitive ability has been partialed out (Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Sá, West, & Stanovich, 1999; Stanovich & West, 1998).

The ability to reason independently of prior belief is only one component of critical thinking. Many theorists view critical thinking as a subspecies of rational thinking or at least as closely related to rational thinking (Kuhn, 2005; Moshman, 2004, 2005, in press; Reyna, 2004; Siegel, 1988, 1997). Cognitive scientists recognize two types of rationality: instrumental and epistemic. To think rationally means adopting appropriate goals, taking the appropriate action given one's goals and beliefs, and holding beliefs about the world that are commensurate with available evidence. These characteristics of rational thinking are precisely the features that a number of leading critical thinking theorists have highlighted in recent descriptions of critical thinking. Thus, Ennis (1996) describes critical thinking as "a process, the goal of which is to make reasonable decisions about what to believe and what to do" (p. xvii), Halpern (2008) states that "critical thinking is the use of those cognitive skills or strategies that increase the probability of desirable outcomes" (p. 3), and Facione (2007) emphasizes the importance of making purposeful, reflective judgments "about what to believe or what to do—precisely the kind of judgment which is the focus of critical thinking" (p.13).

If one accepts the theoretical linkage between critical thinking and rational thought, then there may well be other aspects of critical thought beyond the avoidance of egocentric processing that has been emphasized in previous work in critical thinking. However, we would also want any new domain of critical thinking to at least partially dissociate from measures of cognitive ability (as does the avoidance of belief bias). In this study, we examined one candidate class of biases that qualify as indices of critical/rational thought and that may well dissociate from cognitive ability because tests of the latter do not assess it directly.

A prime candidate for a cognitive domain not assessed by tests of intelligence or tests of critical thinking is found in the panoply of effects studied in the heuristics and biases literature, which has a 30-year history in cognitive psychology (Evans, 1989, 2007; Evans & Over, 1996; Gilovich, Griffin, & Kahneman, 2002; Kahneman, 2003; Kahneman & Tversky, 1973, 1996, 2000; Over, 2004; Tversky & Kahneman, 1974, 1983, 1986). Many of the heuristics and biases studied relate to important aspects of rational and critical thought: causal reasoning, probabilistic reasoning, hypothetical thought, theory justification, assessment of the covariation of events, scientific reasoning, disjunctive reasoning, the tendency to think statistically, and the tendency to think of alternative explanations. These areas are legitimately classified as part of critical thinking. Research in the field of cognitive psychology has shown that these thinking characteristics can be measured and that they relate to important real-world decisions in domains such as personal finance, employment, health, and public policy (Baron, Bazerman, & Shonk, 2006; Hastie & Dawes, 2001; Hilton, 2003; Kahneman & Tversky, 2000; Lichtenstein & Slovic, 2006; Myers, 2002; Reyna & Farley, 2006; Reyna & Lloyd, 2006; Sunstein, 2002, 2005).

In short, on theoretical grounds, the traditional heuristics and biases studied by cognitive psychologists should be considered to be part of a broadened concept of critical thinking. We wish here to introduce theorists to this class of thinking skills because they are largely untapped by currently used critical thinking tests and to provide a preliminary indication of their likely empirical relationships to related constructs. In this investigation, we began some of the empirical groundwork for a theoretical integration of the heuristics and biases framework with more traditional approaches to critical thinking. We examined their relationship to an empirical marker of classic definitions of critical thinking (belief bias in logical reasoning) as well as their relationship to cognitive ability and thinking dispositions.

Although a heuristic process can result in behavior that is appropriate for a given purpose (e.g., unreflectively looking both ways before crossing a street), the set of tasks that we explored highlight important situations in which heuristics and biases are negatively associated with good critical thinking, and result in poor judgments and decisions about what to believe and what to do. Thus, in this article, we posit that the override of heuristics and the avoidance of biases are related to critical thinking.

We should emphasize that the large collection of skills and tasks from the heuristics and biases literature that we examined are not represented on traditional critical thinking tests like the WGCTA. For example, the WGCTA and similar genre of critical thinking tests do not assess the use of base rates in probabilistic reasoning; they do not tap knowledge of the law of large numbers or the ability of people to make regressive probabilistic predictions. They do not assess the gambler's fallacy, nor, in general, do they assess covariation detection, denominator neglect, or Bayesian probabilistic updating. In addition, they do not assess a critical aspect of rationality, *descriptive invariance*, that is the property that causes framing effects when violated (for reviews of the large body of literature on each of these effects, see Gilovich et al., 2002; Nickerson, 2004; Reyna, Lloyd, & Brainerd, 2003; Stanovich, 1999, 2004, 2008b). All of these effects and biases were examined in the present study.

In addition to examining the relationship between these varied measures of the ability to avoid heuristics and cognitive biases, we used the syllogistic reasoning task as an index of the type of discourse-logic-based reasoning skills that are assessed with traditional critical thinking tests like the WGCTA—the critical ability to reason logically when logic conflicts with prior belief. We examined whether cognitive ability was coextensive with variation on these two classes of critical thinking skills and whether cognitive ability mediated any connection between the two critical thinking skills. Additionally, we examined two thinking dispositions that have been found to be independent predictors of belief bias (need for cognition and actively open-minded thinking) in order to ascertain whether they were likewise independent predictors of the ability to avoid cognitive biases.

Method

Participants and Procedure

The participants were 793 undergraduate students (264 men and 529 women) recruited through an introductory psychology subject pool at a medium-sized state university. Their mean age was 19.0 years ($SD = 2.1$). The majority of these students were freshmen (486 students) or sophomores (213 students), and a minority were juniors (65 students) and seniors (27 students). The year in college was missing for 2 participants. Almost 89% of them identified themselves as White (704 students), and a minority identified themselves as African American (20 students), Asian American (42 students), or "other" race/ethnicity (27 students).

Participants completed the battery of tasks during a single, 2-hr session. After completing the informed consent and demographic information sheet, participants completed, in order, the thinking disposition items, the syllogistic reasoning task, and the heuristics and biases tasks.

Tasks and Variables

Syllogistic Reasoning Problems with Belief Bias

Twelve syllogistic reasoning problems, largely drawn from Markovits and Nantel (1989), were completed by the participants. Each problem was worded such that the validity judgment was in conflict with the believability of the conclusion. There were two types of these so-called inconsistent syllogisms. One type of inconsistent syllogism had a believable conclusion but an invalid format (e.g., "All flowers have petals; roses have petals; therefore, roses are flowers"—which is invalid). The other type had an unbelievable conclusions in a logically valid format (e.g., "All things with four legs are dangerous; poodles are not dangerous; therefore, poodles do not have four legs"—which is valid). Therefore, the believability of the content was inconsistent with the logical format of the syllogism in both types. Problems of this type have typically been thought to mirror the critical thinking skill of being able to put aside one's prior knowledge and reason from new premises. After each item, the participants indicated their responses by selecting one of the two alternatives: (a) Conclusion follows logically from premises, or (b) Conclusion does not follow logically from premises. Scores on the inconsistent syllogisms ranged from 0 to 12 ($M = 6.9$, $SD = 3.0$).

Heuristics and Biases Tasks

Causal Base Rate 1: Volvo problem. In this problem, adapted from Fong, Krantz, and Nisbett (1986), a couple are deciding to buy one of two otherwise equal cars. The *Consumer Reports* survey, statistics on repair records, and polls of experts favor the Volvo over the Saab. However, a friend reports experiencing a severe mechanical problem with the Volvo he owns. The participants were asked to provide advice to the couple. Preference for the Volvo indicates a tendency to rely on the large-sample information in spite of salient personal testimony. A preference for the Saab indicates reliance on the personal testimony over the opinion of experts and the large-sample information. Any degree of preference for the Volvo was scored as 1, and any degree of preference for the Saab was scored as 0.

Causal Base Rate 2: Superintendent problem. This is a slightly edited and revised version of a problem in Fong et al. (1986) that was analogous to the Volvo problem and was scored similarly. In this scenario, the superintendent of schools is urging the school board to make a curriculum shift. The superintendent must trade off the evidence of empirical studies conducted on a large sample of school districts that points in one direction versus the testimony of a school board member with a personal anecdote that points in the other direction.

Noncausal base rate problem. Base rates that have a causal relationship to the criterion behavior (Ajzen, 1977; Bar-Hillel, 1990; Tversky & Kahneman, 1979) are often distinguished from noncausal base rate problems—those involving base rates with no obvious causal relationship to the criterion behavior. In the previous two problems the statistical base rate information was causal (the *Consumer Reports* survey and the empirical studies of the school districts). In contrast, this problem is an adaptation of a noncausal base rate problem that has been much studied (Casscells, Schoenberger, & Graboys, 1978; Sloman, Over, Slovak, & Stibel, 2003; Stanovich & West, 1999, 2000). The problem was worded as follows:

> It is known that 1 dollar out of every 10,000 is counterfeit. Imagine a money-changing machine that rejects real dollar bills 5 out of every 100 times it changes money. However, it always rejects bills when they are counterfeit. If this machine rejects your dollar bill, what is the probability (expressed as a percentage ranging from 0% to 100%) that your bill is counterfeit? Choose the best answer.

The problem was followed by the choices: (a) Less than 1%, (b) About 5%, (c) About 50%, (d) About 95%, and (e) More than 95%. Alternative a is the correct response and was scored as 1 (other responses were scored as 0).

Law of large numbers: Hospital problem. This problem is a classic and much-cited problem studied by Tversky and Kahneman (1974). It is used to explore participants' understanding that, other things being equal, a larger sample size more accurately estimates a population value.

Regression to the mean. Drawn from Lehman, Lempert, and Nisbett (1988), this is a multiple-choice problem involving baseball batting averages, and only one of the alternatives shows some recognition of the possibility of regression effects and was scored as 1. Other options were scored as 0.

Gambler's Fallacy 1. In the first gambler's fallacy problem—the slot machine problem—the participant read the following:

> When playing slot machines, people win something about 1 in every 10 times. Lori, however, has just won on her first three plays. What are her chances of winning the next time she plays? Choose the best answer.

The problem was followed by the choices: (a) She has better than 1 chance in 10 of winning on her next play, (b) She has less

than 1 chance in 10 of winning on her next play, (c) She has a 1 chance in 10 that she will win on her next play. The correct response of c was scored as 1, while any other response incorrect and scored as 0.

Gambler's Fallacy 2. This problem is similar to the last and involves coin tosses.

Conjunction problem. This problem is based on Tversky and Kahneman's (1983) much-studied Linda problem, one of the most studied problems in the heuristics and biases literature. Conjunction problems assess whether people appreciate that the probability of Event A and Event B both occurring must be lower than the probability of either A or B alone occurring.

Covariation detection. This problem appeared as follows:

> A doctor had been working on a cure for a mysterious disease. Finally, he created a drug that he thought would cure people of the disease. Before he could begin to use it regularly, he had to test the drug. He selected 400 people at random who had the disease. Of the 400, he randomly assigned 300 to the treatment group and gave them the drug to see what happened. He randomly assigned 100 people to the no-treatment group and gave them a placebo (a sugar pill manufactured to look like the treatment drug) to see what happened. The table below indicates the outcome (Figure 1):

	Group	
Drug condition	Cured	Not cured
Received	200	100
Did not receive	75	25

Does the drug work?

Figure 1.

Participants were asked to choose the statement that best summarized the results shown in the table from among the following statements: (a) The evidence indicates that the drug was effective, (b) The evidence is inconclusive, or (c) The evidence indicates that the drug was not effective. The correct response, c, was scored as 1 (the other responses were scored as 0).

Methodological reasoning. Adapted from Lehman et al. (1988), this multiple-choice problem has only one alternative that indicates the ability to reason methodologically about confounded variables in everyday life. It was scored as 1 and the other responses as 0.

Bayesian Reasoning 1. This problem was the David Maxwell problem adapted from Beyth-Marom and Fischhoff (1983) and studied by Stanovich and West (1998). It is used to assess Bayesian belief updating.

Bayesian Reasoning 2. This problem was the Mark Smith problem adapted from Beyth-Marom and Fischhoff (1983) and studied by Stanovich and West (1998) It is used to assess Bayesian belief updating.

Four card selection task. Originally used by Wason (1966), the selection task has been studied extensively in the reasoning literature (e.g., Evans, Newstead, & Byrne, 1993; Evans & Over, 1996; Johnson-Laird, 1999). The problem involves reasoning about whether an "if-P-then-Q" type of rule can be falsified. Because the rule is in the form of an if-P-then-Q rule, the participant must turn over the two cards that could potentially falsify the rule—the P card and the not-Q card (in this case, the Baltimore and train cards), which is the correct answer and was scored as 1. Very few participants give this completely correct response. Thus, we scored the task leniently by including as correct an alternative task construal championed by Margolis (1987). He has argued that turning the P card only is an appropriate response if the participant has adopted a so-called "open" reading of the rule—one in which the cards represent classes rather than individual exemplars. Therefore, the selection of the P card only was considered a correct answer and also scored as 1. All other selections were scored as 0.

Framing problem. The much-researched disease framing problem was originally studied by Tversky and Kahneman (1981). This problem is presented in two parts: positive and negative framing. Descriptive invariance is correct and was scored as 1. Violation of description invariance was scored as 0.

Probabilistic reasoning: Denominator neglect. This probabilistic reasoning task is a marble game that was modeled on a task introduced by Kirkpatrick and Epstein (1992; see also Denes-Raj & Epstein, 1994; Reyna, 1991; Reyna & Brainerd, 1994, 2007). The problem read as follows:

> Assume that you are presented with two trays of black and white marbles: a large tray that contains 100 marbles and a small tray that contains 10 marbles. The marbles are spread in a single layer on each tray. You must draw out 1 marble (without peeking, of course) from either tray. If you draw a black marble, you win $2. Consider a condition in which the small tray contains 1 black marble and 9 white marbles, and the large tray contains 8 black marbles and 92 white marbles. [A drawing of two trays with their corresponding numbers of marbles arranged neatly in 10-marbles-rows appeared above the latter sentence.] From which tray would you prefer to select a marble in a real situation?

The correct response is to select the small tray and was scored as 1.

Disjunctive reasoning. This problem was drawn from Levesque (1986, 1989) and is a variant on ones used by Toplak and Stanovich (2002). The problem was presented as follows:

> There are 3 blocks in a stack, where each of the blocks is either new or old. The top block is new, and the bottom one is old. The middle block is either new or old. Is there a new block directly on top of an old block?

The text of the problem was accompanied by a graphic that displayed three blocks with their corresponding descriptions. The top block was labelled "new" the middle block was labelled "new or old", and the bottom block was labelled "old."

Participants were presented with the following alternatives: (a) Yes, (b) No, and (c) Cannot be determined. In order to solve this problem, one needs to consider the disjuncts of the middle box. That is, if the middle box is new, then the answer is "yes" because it would be on top of the bottom box, which is old. If the middle box is old, then the answer is still "yes" because the top box would now be on top of the middle box, which is old. The correct solution is thus "yes" and was scored as 1, and other responses were scored as 0.

Heuristics and biases composite score. For the purposes of several of the analyses reported below, we summed the scores on the 16 heuristics and biases tasks to form a composite score (M = 7.7, SD = 2.6, range = 1–15). By forming a composite score, we do not mean to imply that these heuristics and biases tasks form a unidimensional construct. Previous intercorrelations of smaller sets of these tasks than those used here have shown that they are

only modestly correlated (Bruine de Bruin, Parker, & Fischhoff, 2007; Klaczynski, 2001; Parker & Fischhoff, 2005; Stanovich & West, 1998, 2000). The rational thinking tendencies measured by these heuristics and biases tasks are probably multifarious (Parker & Fischhoff, 2005; Reyna et al., 2003; Stanovich 2008a, 2008b; Toplak, Liu, Macpherson, Toneatto, & Stanovich, 2007). Finally, each task, from a psychometric point of view, represents only a single item. Thus, only modest reliability for the composite score is expected, and this was the case. The split-half reliability was .52, and Cronbach's alpha was .53.

Cognitive Ability

Students were asked to indicate their verbal, mathematical, and total SAT scores on the demographics form. The mean reported verbal SAT score of the students was 584 (SD = 70), the mean reported mathematical SAT score was 590 (SD = 68), and mean total SAT score was 1,173 (SD = 102). These self-reported scores closely matched the averages for this institution at the time of testing (582, 587, and 1,169, respectively). Several studies have indicated that the correlation between self-reported SAT scores and verified SAT scores is in the range of .80—.90 (Cassady, 2001; Kuncel, Crede, & Thomas, 2005) as is the correlation between self-reported grade point average (GPA) and verified GPA (Higgins, Peterson, Pihl, & Lee, 2007). An indication of the validity of the self-reported scores is that they correlated with a third variable to the same extent as verified scores. Stanovich and West (1998) found that the correlation between scores on a vocabulary test and self-reported SAT total scores (.49) was quite similar to the .51 correlation between the vocabulary test scores and verified total SAT scores in a previous investigation using the same vocabulary measure (West & Stanovich, 1991). The total SAT score is used as an index of cognitive ability in the analyses reported here because it loads highly on psychometric general intelligence as measured by a variety of indicators (Frey & Detterman, 2004; Unsworth & Engle, 2007). The total SAT score will be used in the analyses reported here.

Thinking Dispositions

Two thinking dispositions were measured—actively open-minded thinking (Stanovich & West, 1997, 2007) and need for cognition (Cacioppo, Petty, Feinstein, & Jarvis, 1996). Items from the two scales were intermixed. We used the 41-item Actively Open-Minded Thinking (AOT) Scale (Stanovich and West, 2007). All items were scored such that higher scores represented a greater tendency toward open-minded thinking. Examples of items are "People should always take into consideration evidence that goes against their beliefs"; "Certain beliefs are just too important to abandon, no matter how good a case can be made against them" (reverse scored); and "No one can talk me out of something I know is right" (reverse scored). The responses for each item in the questionnaire were *strongly agree* (6), *moderately agree* (5), *slightly agree* (4), *slightly disagree* (3), *moderately disagree* (2), and *strongly disagree* (1). Higher scores on the scale indicate cognitive flexibility and lower scores indicate cognitive rigidity and resistance to belief change. The score on the scale was obtained by summing the 41 responses to the items (M = 170.5, SD = 18.3). The split-half reliability (Spearman–Brown corrected) of the scale was .76, and Cronbach's alpha was .84. The total score on the AOT scale was standardized, and the z score was used in statistical analyses.

The 18-item Need for Cognition Scale (Cacioppo et al., 1996) was used in this study. Sample items include "The notion of thinking abstractly is appealing to me," and "I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought." The response format for each item in the questionnaire was the same as that used for the AOT. The mean score was 69.0 (SD = 12.2). The split-half reliability (Spearman–Brown corrected) of the Need for Cognition Scale was .81, and Cronbach's alpha was .88. The total score on the scale was standardized, and the z score was used in statistical analyses.

Table 1
Heuristics and Biases Tasks: Percentage Correct, Correlations, and Partial Correlations, With SAT Total Covaried

		Correlation			Partial correlation
Task	Percent correct	SAT	Thinking dispositions[a]	Belief bias syllogisms	Syllogisms with SAT covaried[b]
Causal Base Rate 1	63.7	.211***	.128***	.215***	.139***
Causal Base Rate 2	77.3	.158***	.127***	.125***	.063
Noncausal base rate problem	47.9	.200***	.086*	.212***	.141***
Law of large numbers	41.7	.044	.057	.029	.010
Regression to the mean	39.6	.151***	.052	.155***	.099**
Gambler's Fallacy 1	50.9	.179***	.136***	.242***	.184***
Gambler's Fallacy 2	79.1	.115**	.104**	.091*	.045
Conjunction problem	27.5	.022	.033	.004	−.006
Covariation detection	31.7	.011	.063	.184***	.199***
Methodological reasoning	37.5	.062	.044	.066	.043
Bayesian Reasoning 1	61.8	.204***	.137***	.213***	.140***
Bayesian Reasoning 2	26.7	.237***	.178***	.274***	.195***
Four card selection task	19.9	.136***	.151***	.189***	.145***
Disease framing problem	62.2	.073*	.096**	.141***	.122***
Probabilistic reasoning: Denominator neglect	64.9	.191***	.092**	.149***	.073*
Disjunctive reasoning	37.7	.205***	.085*	.172***	.093**

Note. N = 793. SAT = SAT total.
[a] Thinking disposition composite. [b] Belief bias syllogisms.
* $p < .05$. ** $p < .01$. *** $p < .001$.

A composite thinking dispositions score was also constructed. The z scores for the AOT and Need for Cognition scales were summed to create a composite score of these two measures. The reliability of the Composite Thinking Dispositions Scale was .88 (Cronbach's alpha).

Results

Table 1 displays the percentage of participants who responded correctly on each of the heuristics and biases tasks. There is considerable variation in task difficulty. Consistent with the extensive literature on the difficulty of the four card selection task (e.g., Evans, 2006; Evans et al., 1993), this task was solved by less than 20% of the participants (the lowest solution rate of all the tasks). The easiest task was the second gambler's fallacy problem, which was answered correctly by almost 80% of the participants. More important is the fact that each of these tasks was answered incorrectly by at least substantial minorities. This is significant because, collectively, these tasks assess whether people adhere to some of the most fundamental strictures of rational thought (see Evans & Over, 1996; Gilovich et al., 2002; Kahneman & Tversky, 1996, 2000; Over, 2004; Samuels & Stich, 2004; Stanovich, 1999, 2004, 2008b; Stein, 1996). These results converge with those reported in a small body of work indicating that the susceptibility to these biases varies considerably (Bruine de Bruin et al., 2007; Klaczynski, 2001; Parker & Fischhoff, 2005; Stanovich & West, 1998, 2000, 2008). What predicts this variation in the susceptibility to these biases, and how does this variation relate to that in another foundational critical thinking skill—reasoning independently of prior belief? The next several analyses address these questions in various ways.

The second column in Table 1 lists the zero-order correlations between each of the 16 heuristics and biases tasks and the SAT total. Twelve of the 16 correlations in this column were significant. The four nonsignificant correlations were found for the law of large numbers, conjunction problem, covariation detection, and methodological reasoning problems. The third column lists the zero-order correlations between the heuristics and biases tasks and the thinking dispositions composite. Eleven of these 16 correlations were significant. Although the correlations between the reasoning tasks and thinking dispositions were more modest than those involving SAT, most once again reached a level of statistical significance. The five nonsignificant correlations were found for the law of large numbers, regression to the mean, the conjunction problem, covariation detection, and methodological reasoning problems. The fourth column lists the zero-order correlations between the heuristics and biases tasks and the belief bias syllogisms. Thirteen of these 16 correlations were significant. The three nonsignificant correlations were found for the law of large numbers, the conjunction problem, and methodological reasoning tasks. The final column displays the partial correlations between the heuristics and biases tasks and belief bias syllogisms after SAT was statistically covaried. Even after SAT was statistically partialed, 11 of the 16 correlations remained significant. Across the four columns displaying correlational data, only 3 of the 16 heuristics and biases tasks (law of large numbers, conjunction problem, and methodological reasoning) failed to demonstrate any reliable relationships with the other major measures.

Table 2 presents the zero-order correlations among the major variables in the study.[1] Because of the large sample size in the study, all correlations were significant at the .001 level. The two components of critical thinking—avoidance of thinking biases and syllogistic reasoning independent of prior belief—displayed a moderate correlation with each other (.437). Both variables were moderately correlated with cognitive ability (.390 and .441, respectively) and modestly correlated with the composite thinking dispositions measure (.275 and .248, respectively). Because the association between the two thinking dispositions, the AOT and the Need for Cognition z scores, was relatively moderate ($r = .401$, $p < .001$), analyses paralleling those subsequently reported for the thinking disposition composite were also conducted on the AOT and on the Need for Cognition z scores separately.[2]

Previous research (Kokis et al., 2002; Stanovich & West, 1998) has found that variation in thinking dispositions related to cogni-

Table 2
Correlations Between Cognitive Ability, Thinking Dispositions, Heuristics and Biases Composite Score, and Syllogistic Reasoning With Belief Bias

Variable	1	2	3	4	5	6
1. SAT total score	—					
2. AOT z score	.205	—				
3. Need for Cognition z score	.239	.401	—			
4. AOT and Need for Cognition composite	.265	.837	.837	—		
5. Heuristics and biases composite score	.390	.213	.247	.275	—	
6. Belief bias syllogisms	.441	.191	.225	.248	.437	—

Note. Values greater than $r = .117$ significant at $p < .001$. AOT = Actively Open-Minded Scale.

[1] One caveat concerning the associations that we observed in these studies relates to the restriction of range in our sample. Certainly, it is true that individuals with average and above-average cognitive ability are over-represented in samples composed entirely of university students. Nevertheless, the actual range in cognitive ability found among college students in the United States is quite large. In the past 30 years, the percentage of 25- to 29-years-olds in the United States who have attended college has increased by 50%. By 2002, 58% of these young adults had completed at least 1 or more years of college, and 29% had received at least a bachelor's degree (U.S. Department of Health and Human Services, 2003). However, the restriction of range in cognitive ability was somewhat greater in our sample, because our participants attended a moderately selective state university. The SAT total means of our sample were roughly .74 of a standard deviation above the national mean (College Board, 2006). The standard deviation of the distribution of scores in our sample was roughly .50 of the standard deviation in the nationally representative sample.

[2] Separate analyses involving the AOT and need for cognition thinking dispositions were also conducted in parallel to all of the analyses in this article in which the thinking dispositions composite score was used. Across the analyses, the pattern of findings for the two component dispositions closely matched those found for the composite measure. Both the AOT and need for cognition were individually significant predictors of performance in all of the regression analyses presented in Table 3, and both AOT and need for cognition accounted for comparable unique and common variance in the commonality analyses shown in Table 4. In each case, however, the thinking disposition composite measure accounted for a larger proportion of variance on the criterion variables.

tive flexibility predicts the avoidance of belief bias in reasoning after variation in cognitive ability has been controlled. The first regression analysis in Table 3 represents a replication of this finding. When entered first as a predictor of performance on the belief bias syllogisms, SAT total predicted 19.4% of the variance and, when entered second, the thinking dispositions composite predicted a statistically significant 1.9% of variance ($p < .001$). The simultaneous regression of these two variables on belief bias syllogism performance indicated that cognitive ability was a more potent unique predictor (15.1% unique variance). A comparison of the simultaneous beta weights indicated the same thing.

The next regression analysis in Table 3 indicated that the thinking dispositions composite was likewise a unique predictor of the heuristics and biases composite score. When entered first as a predictor, SAT total predicted 15.2% of the variance and, when entered second, the thinking dispositions composite predicted a statistically significant 3.2% of variance ($p < .001$). The simultaneous regression of these two variables on heuristics and biases performance indicated that cognitive ability was a more potent unique predictor (10.9% unique variance). A comparison of the simultaneous beta weights converged with this conclusion, although the variables were more balanced predictors than was the case in the analysis of syllogistic reasoning.

As shown in Table 2, the two components of critical thinking displayed a moderate correlation with each other (.437). The next analysis explored whether the link between the two was entirely mediated by cognitive ability. The criterion variable in this analysis was the heuristics and biases composite score. When entered first as a predictor, SAT total predicted 15.2% of the variance and, when entered second, performance on the belief bias syllogisms predicted a substantial 8.7% of variance ($p < .001$). The simultaneous regression of these two variables on the heuristics and biases score indicates that syllogistic reasoning performance was a more potent unique predictor than SAT total score (8.7% versus 4.8% unique variance). A comparison of the simultaneous beta weights provides a redundant way of comparing the potency of the two predictors.

The last regression in Table 3 examined whether the variance overlap between the susceptibility to various cognitive biases and the ability to reason independently of prior belief remained when the variance due to both cognitive ability and thinking dispositions was controlled. When entered first as a predictor of the heuristics and biases composite, SAT total predicted 15.2% of the variance. When entered second, the thinking dispositions composite predicted an additional 3.2% of variance ($p < .001$). Finally, when entered last, performance on the belief bias syllogisms accounted for a substantial 7.4% additional variance. The simultaneous regression of these three variables on the heuristics and biases score indicated that syllogistic reasoning performance was a more potent unique predictor than either SAT total score or the thinking dispositions composite score (7.4% vs. 3.8% and 1.9% unique variance, respectively). A

Table 3
Hierarchical and Simultaneous Regression Results

	Hierarchical		Simultaneous		
Criterion variable/order entered	ΔR^2	F to enter	β weight	F for β	Unique variance explained
Belief bias syllogisms					
1. SAT total	.194	190.67	.403	151.71	.151
2. Thinking disposition composite	.019	18.60	.141	18.60	.019
Overall regression					
$F = 106.76$					
Multiple $R = .461$					
Multiple $R^2 = .213$					
Heuristics and biases composite					
1. SAT total	.152	142.00	.341	104.80	.109
2. Thinking disposition composite	.032	30.53	.184	30.53	.032
Overall regression					
$F = 88.92$					
Multiple $R = .429$					
Multiple $R^2 = .184$					
Heuristics and biases composite					
1. SAT total	.152	142.00	.245	50.28	.048
2. Belief bias/syllogisms	.087	90.63	.329	90.63	.087
Overall regression					
$F = 124.36$					
Multiple $R = .489$					
Multiple $R^2 = .239$					
Heuristics and biases composite					
1. SAT total	.152	142.00	.218	39.32	.038
2. Thinking disposition composite	.032	30.53	.141	19.18	.019
3. Belief bias syllogisms	.074	78.38	.306	78.38	.074
Overall regression					
$F = 91.21$					
Multiple $R = .507$					
Multiple $R^2 = .258$					

Note. All F values in this analysis are significant at the $p < .001$ level.

comparison of the simultaneous beta weights provides a redundant way of comparing the potency of the three predictors.

As an additional way to reveal the overlap in the variables as predictors of susceptibility to the cognitive biases, we conducted a commonality analysis (see Pedhazur, 1997) in which the variance explained by each variable was partitioned into a portion unique to that variable and portions shared with every possible combination of variables. The results of the commonality analysis are presented in Table 4. The first row indicates the unique variance in the heuristics and biases composite score explained by each of the predictors. The next row displays the explained variance in heuristics and biases scores that was common to SAT and thinking dispositions (.010). The third row displays the explained variance in heuristics and biases scores that was common to SAT and the belief bias syllogisms (.071). The fourth row displays the explained variance in heuristics and biases scores that was common to thinking dispositions and belief bias syllogisms (.013). The fifth row indicates that the explained variance in heuristics and biases scores that was common to all three predictors is .033. All of the variance components added together (.038 + .019 + .074 + .010 + .071 + .013 + .033) summed to the total variance explained in heuristics and biases scores by the three predictors (.258).

The unique variances explained here recapitulates those derived from the regression analyses. The commonality analysis reveals that the primary reason that SAT was not a more potent unique predictor was because of its overlap with variance explained by syllogistic reasoning performance (.071). Secondarily, the unique variance explained by cognitive ability (.038) was roughly similar to that in common to all three variables (.033). In contrast, syllogistic reasoning was more separable from other predictors. The proportion of unique variance that it accounted for was 2–3 times as large as that of the other two predictors, and it was the only indicator in which the amount of unique variance explained (.074) was larger than the explained variance held in common by all three predictors (.033).[3]

Table 4
Results of a Commonality Analysis Using the Heuristics and Biases Composite Score as a Criterion Variable

	Variable		
Commonality	1. SAT total	2. Thinking dispositions[a]	3. Belief bias/ syllogisms
Unique variance	.038	.019	.074
Common 1 & 2[b]	.010	.010	—
Common 1 & 3[c]	.071	—	.071
Common 2 & 3[d]	—	.013	.013
Common 1 & 2 & 3[e]	.033	.033	.033
Total unique variance + common variance	.152	.075	.191

Note. Unique variance = explained variance in heuristics and biases composite scores that is unique to that variable.
[a] Thinking disposition composite. [b] Common 1 & 2 = explained variance that was common to Variables 1 and 2. [c] Common 1 & 3 = explained variance that was common to Variables 1 and 3. [d] Common 2 & 3 = explained variance that was common to Variables 2 and 3. [e] Common 1 & 2 & 3 = explained variance that was common to all three variables.

Discussion

The cognitive biases tapped by our heuristics and biases tasks are not typically assessed by measures of critical thinking. Yet there are theoretical reasons to believe that they should be. To perform well on these heuristics and biases tasks, one needs both declarative knowledge and the proper mental strategies and metastrategies (Stanovich, 2008a, 2008b). The heuristic response must be inhibited and replaced with a more normatively appropriate response. The inhibition process depends not only on cognitive capacity but also on the metacognitive sensitivity (i.e., reflectivity; see Stanovich, 2008b) to begin the process of response suppression. Once the heuristic response is inhibited, the knowledge to make the right substitute response must be present. Declarative knowledge of probability, causal reasoning, scientific thinking, and logic comes in to play at this step.

Several cognitive theorists have analyzed critical thinking in terms of rational thinking concepts and the philosophy of rational thought (e.g., Kuhn, 2005; Moshman, 2004, in press; Siegel, 1997). For example, discussions of critical thinking often concern the thought

[3] Structural equation modeling analyses provided additional evidence consistent with the pattern of findings reported in Table 3 and Table 4. Four different models were evaluated for relative fit. Model 1, which is most consistent with results displayed in the regression (Table 3) and commonality (Table 4) analyses, posited four separate latent constructs, with those of cognitive ability, thinking dispositions, and belief bias syllogistic reasoning simultaneously predicting the heuristics/biases endogenous construct. Three indicators were used for cognitive ability: total SAT score, high school grade point average (GPA), and college GPA (either high school or college GPA was missing for 78 participants). Two indicators were used for thinking dispositions: Need for Cognition Scale and AOT z scores. We created two item parcels each as separate indicators for the belief bias syllogism and heuristics/biases latent constructs by including odd-numbered and even-numbered items in each item parcel. This resulted in two indicators for both the belief bias syllogistic reasoning and the heuristics/biases composite.

Three additional models were tested to explore how well combining the latent constructs in various ways could account for the data. The following combinations were explored: when cognitive ability and thinking dispositions were treated as one of two exogenous constructs (Model 2); when cognitive ability and syllogistic reasoning were treated as one of two exogenous constructs (Model 3); and when cognitive ability, thinking dispositions, and syllogistic reasoning were treated as the one and only exogenous construct (Model 4). The heuristics and biases composite was the endogenous construct in each model. Model 1 generated $\chi^2(21, N = 715) = 99.2$, $p = .0001$; confidence fit index (CFI) = .95; Tucker-Lewis index (TLI) = .91; root-mean-square error of approximation (RMSEA) = .07. Model 2 generated $\chi^2(24, N = 715) = 176.2$, $p = .0001$; CFI = .90; TLI = .84; RMSEA = .09. Model 3 generated $\chi^2(32, N = 715) = 208.6$, $p = .0001$; CFI = .88; TLI = .83; RMSEA = .09. Model 4 generated $\chi^2(26, N = 715) = 303.2$, $p = .0001$; CFI = .81; TLI= .74; RMSEA = .12. While the chi-square statistic was significant for all of these models due to the large sample size in this study, Model 1 had the best fit of those tested on the basis of other indices used in this analysis. This suggests that considering these four constructs as separate latent variables is a good fit for this data.

The correlations and standardized weights from Model 1 also converged strongly with the results in Tables 3 and 4. Cognitive ability was significantly correlated with belief bias syllogisms ($r = .57$, $p < .001$) and thinking dispositions ($r = .44$, $p < .001$), and thinking dispositions were significantly correlated with syllogistic reasoning ($r = .33$, $p < .001$). Further, the path coefficient from cognitive ability to heuristics and biases was significant (.32, $p < .002$), the path from thinking dispositions to heuristics and biases was significant (.23, $p < .001$), and the path from syllogistic reasoning to heuristics and biases was significant (.38, $p < .0001$).

processes that we use to reason about what we should believe and how we should act given those beliefs. These thought processes correspond, respectively, to epistemic rationality (rationality of belief) and practical rationality (rationality of action). The distinction between epistemic and practical rationality has been much discussed in philosophy and cognitive science (Over, 2004; Samuels & Stich, 2004).

If these theorists are correct, then critical thought is rational thought—and the tasks used in our heuristics and biases battery tap some of the most important strictures of rational thinking that cognitive scientists have identified (Over, 2004; Samuels & Stich, 2004). For example, our battery examined violations of rational thinking such as conjunction errors in probability judgment, failures of descriptive invariance, ignoring sample size, failure to weight the denominator of the likelihood ratio, ignoring base rates, and failure to regress predictions (see Evans, 2007; Gilovich et al., 2002; Kahneman & Tversky, 1996, 2000; Stanovich, 1999, 2008b).

Our measure of the ability to avoid these thinking errors was moderately correlated ($r = .437$) with a more traditional laboratory measure of critical thinking—the ability to reason logically when logic conflicts with prior belief. The correlation between these two classes of critical thinking skills was not due to a joint connection with cognitive ability. It remained statistically significant after the variance due to cognitive ability was partialed out (partial $r = .321$, $p < .001$). As the third regression in Table 3 shows, performance on the syllogistic reasoning task with belief bias predicted 8.7% unique variance ($p < .001$) in performance on the heuristics and biases battery after cognitive ability was controlled, and as the fourth regression in Table 3 indicates, performance on the syllogistic reasoning task predicted 7.4% unique variance ($p < .001$) in performance on the heuristics and biases battery after cognitive ability and thinking dispositions both were partialed out. As the commonality analysis in Table 4 indicates, the syllogistic reasoning task had as high a unique connection to heuristics and biases performance (7.4% variance explained) as it held in common with SAT as a predictor (7.1% variance explained).

Another triangulating conclusion that can be drawn from the commonality analysis in Table 4 is that our new class of critical thinking skills (the heuristics and biases tasks) is more specifically related to a classic critical thinking skill (the ability to reason logically when logic conflicts with prior belief) than it is to cognitive ability. The basis of this conclusion is that the unique variance in heuristics and biases performance explained by syllogistic reasoning performance (7.4%) was nearly double the unique variance explained by SAT scores (3.8%).

A notable finding of our study was that the thinking dispositions measures were independent predictors (after cognitive ability was controlled) of both classes of critical thinking skill (see the first and second regressions in Table 3). That thinking dispositions were an independent predictor of the ability to avoid bias by prior belief when reasoning converges with several previous studies. For example, Schommer (1990) found that a measure of the disposition to believe in certain knowledge predicted the tendency to draw one-sided conclusions from ambiguous evidence even after verbal ability was controlled. Kardash and Scholes (1996) found that the tendency to properly draw inconclusive inferences from mixed evidence was related to the disposition to believe in certain knowledge and to a measure of need for cognition. Furthermore, these relationships were not mediated by verbal ability because a vocabulary measure was essentially unrelated to evidence evaluation. Likewise, Klaczynski (1997; see also Klaczynski & Gordon, 1996; Klaczynski & Robinson, 2000) found that the degree to which adolescents criticized belief-inconsistent evidence more than belief-consistent evidence was unrelated to cognitive ability (see also Perkins, Farady, & Bushey, 1991) but often was related to thinking dispositions associated with epistemic regulation. The evaluation of personally relevant arguments has been found to be related to actively open-minded thinking dispositions independent of cognitive ability (Sá et al., 1999; Stanovich & West, 1997).

The results of the first regression of Table 3—that our thinking dispositions measures predicted belief bias in syllogistic reasoning independent of cognitive ability—converges with these previous findings. What is even more notable is that the second regression in Table 3 indicates that the thinking dispositions were also unique predictors of a very different class of critical thinking skill—the susceptibility to the large class of cognitive biases discovered and reported in the heuristics and biases literature.

We wish here to introduce another class of thinking processes that needs to be considered when measuring the critical thinking abilities of individuals. Our goal has not been to develop a new critical thinking psychometric assessment instrument but, instead, to introduce theorists to a wide variety of critical thinking skills that are largely untapped by currently used critical thinking tests and to provide a preliminary indication of their likely empirical relationships to related constructs (cognitive ability and belief bias in reasoning). Typically, individual studies in the heuristics and biases literature have examined a single problem or only a few problems in an effort to understanding the processes that underlie responses that violate rational strictures (e.g., Gilovich et al., 2002; Kahneman, 2003; Kahneman & Tversky, 2000). Alternatively, the psychometric approach is characterized by the use of several similar items, which results in high internal consistency and reliability, as in measures of intelligence. In contrast, our approach has been to use a carefully selected set of different well-established heuristics and biases tasks that were expected to reflect relatively distinct cognitive skills, as opposed to a psychometric validation of a collection of problems thought to be characterized by similar reasoning tendencies. Thus, the rational thinking tendencies that we tapped with our measures reflect more than a single construct. Although some theoretical taxonomies have been proposed (Reyna et al., 2003; Stanovich 2008a, 2008b, in press; Stanovich, Toplak, & West, 2008), empirical work on construct differentiation is as yet virtually nonexistent—it is work we hope to provoke with our study. Given that we measured a multifarious concept with single tasks, the modest reliability of our heuristics and biases score was perhaps better than might have been expected. In the only studies (that we are aware of) in which multiple-item measures of different cognitive biases were used, Parker and Fischhoff (2005) and Bruine de Bruin et al. (2007) found mean interitem correlations of .12 and .16, respectively—higher than our mean of .066 but indicative that the biases are not reflective of a single underlying mechanism.[4] Further work will need to elaborate the structure of these biases so that they eventually might become more fully integrated with more classic notions of critical thinking which stress aspects of non-egocentric processing such as unbiased reasoning in the face of prior belief.

[4] An analysis of the correlation matrix in Table 2 of Klaczynski (2001) likewise yielded mean intertask correlations of roughly .15 among single-task measures of each bias when a single nonredundant variable was used for each task. Similarly, the mean correlation in a study including a battery of heuristics and biases tasks conducted by Slugoski, Shields, and Dawson (1993) was .03.

References

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology, 35,* 303–314.

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

Baron, J. (1991). Beliefs about thinking. In J. Voss, D. Perkins, & J. Segal (Eds.), *Informal reasoning and education* (pp. 169–186). Hillsdale, NJ: Erlbaum.

Baron, J. (2000). *Thinking and deciding* (3rd ed.). Cambridge, England: Cambridge University Press.

Baron, J., Bazerman, M. H., & Shonk, K. (2006). Enlarging the societal pie through wise legislation. A psychological perspective. *Perspectives on Psychological Science, 1,* 123–132.

Beyth-Marom, R., & Fischhoff, B. (1983). Diagnosticity and pseudodiagnositicity. *Journal of Personality and Social Psychology, 45,* 1185–1195.

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.

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.

Cassady, J. C. (2001). Self-reported GPA and SAT: A methodological note. *Practical Assessment, Research, and Evaluation, 7*(12).

Casscells, W., Schoenberger, A., & Graboys, T. (1978). Interpretation by physicians of clinical laboratory results. *New England Journal of Medicine, 299,* 999–1001.

College Board. (2006). *College Board SAT: 2006 college-bound seniors. Total group profile report*. Retrieved June 9, 2008, from http://www.collegeboard.com/prod_downloads/about/news_info/cbsenior/yr2006/national-report.pdf

Denes-Raj, V., & Epstein, S. (1994). Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology, 66,* 819–829.

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science, 17,* 428–433.

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities. In J. Baron, & R. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 9–26). New York: Freeman.

Ennis, R. H. (1996). *Critical thinking.* Upper Saddle River, NJ: Prentice-Hall.

Ennis, R. H., Millman, J., & Tomko, T. N. (1985). *Cornell Critical Thinking Tests*. Pacific Grove, CA: Midwest.

Evans, J. St. B. T. (1989). *Bias in human reasoning: Causes and consequences.* Hove, England: Erlbaum.

Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review, 13,* 378–395.

Evans, J. St. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgment.* New York: Psychology Press.

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.

Evans, J. St. B. T., & Curtis-Holmes, J. (2005). Rapid responding increases belief bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning, 11,* 382–389.

Evans, J. St. B. T., & Feeney, A. (2004). The role of prior belief in reasoning. In J. P. Leighton & R. J. Sternberg (Eds.), *The nature of reasoning* (pp. 78–102). New York: Cambridge University Press.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning: The psychology of deduction.* Hove, England: Erlbaum.

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning.* Hove, England: Psychology Press.

Facione, P. (1992). *California Critical Thinking Skills Test & California Critical Thinking Dispositions Inventory.* La Cruz, CA: California Academic Press.

Facione, P. (2007). *Critical thinking: What it is and why it counts.* Millbrae, CA: Insight Assessment, California Academic Press. Retrieved February 13, 2008, from http://www.insightassessment.com/pdf_files/DEXadobe.PDF

Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology, 18,* 253–292.

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or *g*? The relationship between the Scholastic Assessment Test and general cognitive ability. *Psychological Science, 15,* 373–378.

Garnham, A., & Oakhill, J. V. (2005). Accounting for belief bias in a mental model framework: Comment on Klauer, Musch, and Naumer (2000). *Psychological Review, 112,* 509–518.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment.* New York: Cambridge University Press.

Halpern, D. (2008). *Halpern critical thinking assessment: Background and scoring standards.* Unpublished manuscript, Claremont McKenna College, CA.

Harcourt Assessment. (2006). *Watson–Glaser Critical Thinking Appraisal: Sample questions* [Brochure]. Retrieved January 8, 2008, from http://harcourt.assessment.com/hai/Images/dotCom/HTC/WG/SampleQuestions.pdf

Hastie, R., & Dawes, R. M. (2001). *Rational choice in an uncertain world.* Thousand Oaks, CA: Sage.

Higgins, D. M., Peterson, J. B., Pihl, R. O., & Lee, A. G. M. (2007). Prefrontal cognitive ability, intelligence, Big Five personality, and the prediction of advanced academic and workplace performance. *Journal of Personality and Social Psychology, 93,* 298–319.

Hilton, D. J. (2003). Psychology and the financial markets: Applications to understanding and remedying irrational decision making. In I. Brocas & J. D. Carrillo (Eds.), *The psychology of economic decisions: Vol. 1. Rationality and well-being* (pp. 273–297). Oxford, England: Oxford University Press.

Johnson-Laird, P. N. (1999). Deductive reasoning. *Annual Review of Psychology, 50,* 109–135.

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

Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review, 103,* 582–591.

Kahneman, D., & Tversky, A. (Eds.) (2000). *Choices, value, and frames.* New York: Cambridge University Press.

Kahneman, D. A. (2003). Perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58,* 697–720.

Kardash, C. M., & Scholes, R. J. (1996). Effects of pre-existing beliefs, epistemological beliefs, and need for cognition on interpretation of controversial issues. *Journal of Educational Psychology, 88,* 260–271.

Kirkpatrick, L., & Epstein, S. (1992). Cognitive-experiential self-theory and subjective probability: Evidence for two conceptual systems. *Journal of Personality and Social Psychology, 63,* 534–544.

Klaczynski, P. A. (1997). Bias in adolescents' everyday reasoning and its relationship with intellectual ability, personal theories, and self-serving motivation. *Developmental Psychology, 33,* 273–283.

Klaczynski, P. A. (2001). Analytic and heuristic processing influences on adolescent reasoning and decision making. *Child Development, 72,* 844–861.

Klaczynski, P. A., & Gordon, D. H. (1996). Self-serving influences on adolescents' evaluations of belief-relevant evidence. *Journal of Experimental Child Psychology, 62,* 317–339.

Klaczynski, P. A., & Lavallee, K. L. (2005). Domain-specific identity,

epistemic regulation, and intellectual ability as predictors of belief-based reasoning: A dual-process perspective. *Journal of Experimental Child Psychology, 92,* 1–24.

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.

Klauer, K. C., Musch, J., & Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychological Review, 107,* 852–884.

Kokis, J., Macpherson, R., Toplak, M., 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.

Kuhn, D. (2005). *Education for thinking.* Cambridge, MA: Harvard University Press.

Kuncel, N. R., Crede, M., & Thomas, L. L. (2005). The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Educational Research, 75,* 63–82.

Lehman, D. R., Lempert, R. O., & Nisbett, R. E. (1988). The effect of graduate training on reasoning. *American Psychologist, 43,* 431–442.

Levesque, H. J. (1986). Making believers out of computers. *Artificial Intelligence, 30,* 81–108.

Levesque, H. J. (1989). Logic and the complexity of reasoning. In R. H. Thomason (Ed.), *Philosophical logic and artificial intelligence* (p. 73–107). Dordrecht, the Netherlands: Kluwer Academic.

Lichtenstein, S., & Slovic, P. (Eds.). (2006). *The construction of preference.* Cambridge, England: Cambridge University Press.

Macpherson, R., & Stanovich, K. E. (2007). Cognitive ability, thinking dispositions, and instructional set as predictors of critical thinking. *Learning and Individual Differences, 17,* 115–127.

Margolis, H. (1987). *Patterns, thinking, and cognition.* Chicago: University of Chicago Press.

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

Moshman, D. (2004). From inference to reasoning: The construction of rationality. *Thinking and Reasoning, 10,* 221–239.

Moshman, D. (2005). *Adolescent psychological development: Rationality, morality, and identity* (2nd ed.). Mahwah, NJ: Erlbaum.

Moshman, D. (in press). The development of rationality. In H. Siegel (Ed.), *Oxford handbook of philosophy of education.* Oxford, England: Oxford University Press.

Myers, D. G. (2002). *Intuition: Its powers and perils.* New Haven, CT: Yale University Press.

Nickerson, R. S. (2004). *Cognition and chance: The psychology of probabilistic reasoning.* Mahwah, NJ: Erlbaum.

Norris, S. P., & Ennis, R. H. (1989). *Evaluating critical thinking.* Pacific Grove, CA: Midwest.

Over, D. E. (2004). Rationality and the normative/descriptive distinction. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 3–18). Malden, MA: Blackwell Publishing.

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

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Fort Worth, TX: Harcourt Brace.

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

Piaget, J. (1972). Intellectual evolution from adolescence to adulthood. *Human Development, 15,* 1–12.

Reyna, V. F. (1991). Class inclusion, the conjunction fallacy, and other cognitive illusions. *Developmental Review, 11,* 317–336.

Reyna, V. F. (2004). How people make decisions that involve risk. *Current Directions in Psychological Science, 13,* 60–66.

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: Wiley.

Reyna, V. F., & Brainerd, C. J. (2007). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences, 17.*147–159.

Reyna, V. F., & Farley, F. (2006). Risk and rationality in adolescent decision making. *Psychological Science in the Public Interest, 7,* 1–44.

Reyna, V. F., & Lloyd, F. J. (2006). Physician decision making and cardiac risk: Effects of knowledge, risk perception, risk tolerance, and fuzzy processing. *Journal of Experimental Psychology: Applied, 12,* 179–195.

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). New York: Cambridge University Press.

Sá, W., West, R. F., & Stanovich, K. E. (1999). The domain specificity and generality of belief bias: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 91,* 497–510.

Samuels, R., & Stich, S. P. (2004). Rationality and psychology. In A. R. Mele & P. Rawling (Eds.), *The Oxford handbook of rationality* (pp. 279–300). Oxford, England: Oxford University Press.

Schommer, M. (1990). Effects of beliefs about the nature of knowledge on comprehension. *Journal of Educational Psychology, 82,* 498–504.

Siegel, H. (1988). *Educating reason: Rationality, critical thinking, and education.* London: Routledge.

Siegel, H. (1997). *Rationality redeemed? Further dialogues on an educational ideal.* London: Routledge.

Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational Behavior and Human Decision Processes, 91,* 296–309.

Slugoski, B. R., Shields, H. A., & Dawson, K. A. (1993). Relation of conditional reasoning to heuristic processing. *Personality and Social Psychology Bulletin, 19,* 158–166.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning.* Mahwah, NJ: Erlbaum.

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

Stanovich, K. E. (2008a). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond.* Oxford, England: Oxford University Press.

Stanovich, K. E. (2008b). *The psychology of rational thought: What intelligence tests miss.* New Haven, CT: Yale University Press.

Stanovich, K. E. (in press). *Rationality and the reflective mind: Toward a tri-process model of cognition.* New York: 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.

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.

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

Stanovich, K. E., & West, R. F. (1999). Discrepancies between normative and descriptive models of decision making and the understanding/acceptance principle. *Cognitive Psychology, 38,* 349–385.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences, 23,* 645–726.

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

Stanovich, K. E., & West, R. F. (2008). On the relative independence of

thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94,* 672–695.

Stein, E. (1996). *Without good reason: The rationality debate in philosophy and cognitive science.* Oxford, England: Oxford University Press.

Sternberg, R. J. (1997). *Thinking styles.* Cambridge, England: Cambridge University Press.

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

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity synthesized.* Cambridge, England: Cambridge University Press.

Sunstein, C. R. (2002). *Risk and reason: Safety, law, and the environment.* Cambridge, England: Cambridge University Press.

Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences, 28,* 531–573.

Toplak, M., Liu, E., Macpherson, R., Toneatto, T., & Stanovich, K. E. (2007). The reasoning skills and thinking dispositions of problem gamblers: A dual-process taxonomy. *Journal of Behavioral Decision Making, 20,* 103–124.

Toplak, M., & Stanovich, K. E. (2002). The domain specificity and generality of disjunctive reasoning: Searching for a generalizable critical thinking skill. *Journal of Educational Psychology, 94,* 197–209.

Tversky, A., & Kahneman, D. (1974, September 27). Judgment under uncertainty: Heuristics and biases. *Science, 185,* 1124–1131.

Tversky, A., & Kahneman, D. (1979). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology.* Hillsdale, NJ: Erlbaum.

Tversky, A., & Kahneman, D. (1981, January 30). The framing of decisions and the psychology of choice. *Science, 211,* 453–458.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review, 90,* 293–315.

Tversky, A., & Kahneman, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59,* 251–278.

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.

U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Evaluation. (2003). Trends in the well-being of America's children and youth, 2003. Retrieved June 10, 2006, from http://aspe.hhs.gov/HSP/03trends/index.htm

Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology* (pp. 135–151). Harmonsworth, England: Penguin.

Watson, G., & Glaser, E. M. (1980). *Watson-Glaser Critical Thinking Appraisal.* New York: Psychological Corp.

West, R. F., & Stanovich, K. E. (1991). The incidental acquisition of information from reading. *Psychological Science, 2,* 325–330.

Received September 25, 2007
Revision received May 6, 2008
Accepted May 8, 2008 ■

Low Publication Prices for APA Members and Affiliates

Keeping you up-to-date. All APA Fellows, Members, Associates, and Student Affiliates receive—as part of their annual dues—subscriptions to the *American Psychologist* and *APA Monitor.* High School Teacher and International Affiliates receive subscriptions to the *APA Monitor,* and they may subscribe to the *American Psychologist* at a significantly reduced rate. In addition, all Members and Student Affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

Essential resources. APA members and affiliates receive special rates for purchases of APA books, including the *Publication Manual of the American Psychological Association,* and on dozens of new topical books each year.

Other benefits of membership. Membership in APA also provides eligibility for competitive insurance plans, continuing education programs, reduced APA convention fees, and specialty divisions.

More information. Write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242.