

Cognitive ability, thinking dispositions, and instructional set as predictors of critical thinking[☆]

Robyn Macpherson, Keith E. Stanovich*

University of Toronto, Canada

Received 25 September 2006; received in revised form 12 May 2007; accepted 18 May 2007

Abstract

This study examined the predictors of belief bias in a formal reasoning paradigm (a syllogistic reasoning task) and myside bias in two informal reasoning paradigms (an argument generation task and an experiment evaluation task). Neither cognitive ability nor thinking dispositions predicted myside bias, but both cognitive ability and thinking dispositions were significant predictors of the ability to overcome belief bias in the syllogistic reasoning task. However, instructional set (either decontextualizing or non-directive instructions) had a significant effect on myside bias in the argument generation task, as well as a marginal effect on the syllogistic reasoning task. On the latter, and to some extent on the former task, instructional set interacted with cognitive ability. The debiasing effect of decontextualizing instructions was particularly large for those participants in the lowest quartile of cognitive ability.

© 2007 Elsevier Inc. All rights reserved.

Keywords: Individual differences; Belief bias; Instructional set; Myside bias; Critical thinking; Cognitive ability

1. Introduction

In the educational literature on critical thinking, the ability to decouple prior beliefs and opinions from the evaluation of evidence and arguments is deemed to be a skill of paramount importance (Baron, 1991, 2000; Ennis, 1987, 1995; Norris & Ennis, 1989; Paul, 1984, 1987; Perkins, 1995; Stanovich, 1999, 2004; Sternberg, 1997, 2001, 2003; Wade & Tavris, 1993). In the empirical literature, this skill is often operationalized as the ability to avoid myside bias and belief bias. Myside bias is the tendency to evaluate evidence, generate evidence, and test hypotheses in a manner biased toward one's own opinions. Strong myside bias effects have been demonstrated in numerous studies (Baron, 1991, 1995; Greenhoot, Semb, Colombo, & Schreiber, 2004; Kuhn, 1991; Nussbaum & Kardash, 2005; Perkins, 1985; Perkins, Farady, & Bushey, 1991; Stanovich & West, 2007; Toplak & Stanovich, 2003). The evidence is

[☆] This research was supported by grants from the Social Sciences and Humanities Research Council of Canada and the Canada Research Chairs program to Keith E. Stanovich.

* Corresponding author. Department of Human Development and Applied Psychology, University of Toronto, 252 Bloor St. West, Toronto, Ontario, Canada M5S 1V6. Tel.: +1 416 923 6641 2447.

E-mail address: kstanovich@oise.utoronto.ca (K.E. Stanovich).

similar with respect to belief bias—accepting or rejecting an argument's conclusion based on its believability rather than on its logical validity. Belief bias exists because people have difficulty evaluating conclusions that conflict with what they think they know about the world (De Neys, 2006; Evans, 2002b; Evans, Barston, & Pollard, 1983; Evans, Newstead, Allen, & Pollard, 1994; Goel & Dolan, 2003; Klauer, Musch, & Naumer, 2000). Belief bias arises from people's factual knowledge about the world (i.e., roses are flowers) in contrast to myside bias which is reasoning biased toward personal opinions or stances.

Nickerson (1998) has reviewed the literature on these and similar critical thinking skills and uses the term confirmation bias to refer to the demonstrated phenomenon of “inappropriate bolstering of hypotheses or beliefs whose truth is in question” (p. 175). This inappropriate bolstering can be driven by prior opinion (as in myside bias) or prior knowledge (as in the belief bias effect in syllogistic reasoning, see Klauer et al., 2000). In all cases, it involves an inability to decouple opinion or knowledge from evaluative information processing tasks.

Researchers have only recently begun to examine the individual difference variables that predict skill at decoupling and avoiding confirmation biases. Individual differences in myside bias have proven difficult to predict from other psychological variables. Correlations between the degree of myside bias and cognitive ability have proven to be extremely modest and are often non-significant (Klaczynski & Gordon, 1996; Klaczynski, Gordon, & Fauth, 1997; Klaczynski & Lavalley, 2005; Klaczynski & Robinson, 2000; Sá, Kelley, Ho, & Stanovich, 2005; Stanovich & West, 2007; Toplak & Stanovich, 2003). Some studies have, however, revealed that thinking dispositions associated with epistemic regulation (e.g., actively open-minded thinking, dogmatism) are related to individual differences in myside bias (Klaczynski et al., 1997; Klaczynski & Lavalley, 2005; Klaczynski & Robinson, 2000; Sá et al., 2005), although others have failed to replicate this relationship (Stanovich & West, 2007).

In contrast to the informal reasoning tasks that have been used to assess myside bias, belief bias has usually been measured using formal reasoning tasks. The most popular formal reasoning task has been the syllogistic reasoning paradigm where validity and the believability of the conclusion are put in conflict (Evans et al., 1983; Evans & Curtis-Holmes, 2005; Klauer et al., 2000). For example, the syllogism ‘All flowers have petals; roses have petals; therefore roses are flowers’ has a believable conclusion but is logically invalid. Significant correlations of moderate size have consistently been found between cognitive ability and belief bias in this task (De Neys, 2006; Gilinsky & Judd, 1994; Handley, Capon, Beveridge, Dennis, & Evans, 2004; Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Newstead, Handley, Harley, Wright, & Farrelly, 2004; Sá, West, & Stanovich, 1999; Simoneau & Markovits, 2003; Stanovich & West, 1998; however, see Torrens, Thompson, & Cramer, 1999). Thinking dispositions related to epistemic regulation have also displayed moderate correlations with belief bias in some studies (Kokis et al., 2002; Sá et al., 1999; Stanovich & West, 1998).

Thus, there appears to be a lack of convergence between individual difference studies of myside bias and belief bias regarding associations with cognitive ability in the critical thinking literature. This result is somewhat surprising as cognitive ability is the most widely studied psychological construct and decades of empirical research have shown correlations between cognitive ability and other measures of psychological performance (Brody, 1997; Gottfredson, 1997; Lubinski & Humphreys, 1997). If the suggestion that cognitive ability is correlated with belief bias (suggesting that cognitive ability facilitates overriding one's prior knowledge of the world in an assessment of logical validity), but not consistently correlated with myside bias (suggesting that cognitive ability does *not* facilitate overriding one's prior opinions when asked to assess the various sides of an argument) is confirmed, certain theoretical and practical implications follow. It would suggest that the type of content under consideration influences an individual's reasoning ability, and reasoning quality can vary within an individual depending on the subject he or she is reasoning about. The teaching of critical thinking skills will need to incorporate the very different ways in which myside bias and belief bias relate to individual differences. Furthermore, many of the contentious issues facing the world today (climate change, overpopulation, and biomedical advances) implicate a complex mix of myside bias and belief bias. Nevertheless, the apparent lack of convergence between myside bias studies and belief bias studies must be inferred across investigations because researchers tend not to include assessments of both myside bias and belief bias in the same experiment. In the present experiment, we examined both biases; belief bias was assessed in the traditional way, using a syllogistic reasoning task, and myside bias was assessed using two informal reasoning tasks.

A review of the critical thinking literature reveals that there is another aspect of the experimental paradigms employed that varies unsystematically and that might affect the degree of bias displayed. That factor is the explicitness of the instructions provided to the participants in studies of bias. In belief bias studies, participants are usually

explicitly and specifically instructed to ignore prior knowledge. For example, in a typical syllogistic reasoning task participants might be told “You must *suppose that the premises are all true* and limit yourself only to the information contained in the premises” (e.g., Sá et al., 1999). Such instructions strongly emphasize the necessity of decoupling prior knowledge and prior beliefs from the validity evaluation (decontextualizing instructions). In contrast, in the informal reasoning tasks used to assess myside bias, there is usually no explicit instruction to set aside prior opinions about the issue that is the focus of the experiment (non-directive instructions).

Evans (2002a) notes that compared to the large amount of research that has been done on issues such as figure and form in deductive reasoning paradigms, little attention has been paid to the effect of instructions on the level of belief bias in syllogistic reasoning. Evans et al. (1994) conducted one of the few such investigations (see also, Dias, Roazzi, & Harris, 2005). They examined several ways of manipulating the instructional set so as to emphasize logical validity over believability (and thus potentially reduce belief bias effects).

For example, in addition to standard instructions in one condition, one group was given augmented instructions that read as follows: “Please note that according to the rules of deductive reasoning, you can only endorse a conclusion if it definitely follows from the information given. A conclusion that is merely possible, but not necessitated by the premises is not acceptable. Thus, if you judge that the information given is insufficient and you are not absolutely sure that the conclusion follows you must reject it and answer “NO; [invalid]” (p. 272). A second phrase, typed in capital letters, reminded participants about the need to respond if and only if the conclusion followed logically.

In another condition, Evans et al. (1994) employed quite complex instructions that consisted of eight paragraphs which contained a very detailed description of the logical properties of syllogisms, including information about form, the relationship between premises and conclusion, and such details as the need to interpret the word ‘some’ strictly by its logical meaning of ‘at least one and possibly all.’ The instructions also guided participants through the steps required to complete the task, with multiple reminders of the need to rely solely on the logical form of the syllogism and to accept the content as true. After studying several such instructional manipulations, Evans et al. (1994) did find some evidence that the magnitude of the belief bias effect was related to the strength of the decontextualizing instructions.

There has been even less research on the effects of instructional set on myside bias. In one of the few studies, Nussbaum and Kardash (2005) manipulated instructional set in an essay writing task. Participants who were given instructions to “discuss two or three reasons why others might disagree with you, and why those reasons are wrong” generated significantly more counterclaims and rebuttals than did a control group. This result revealed that mysided thinking can be attenuated by instructions that prompt consideration of counterarguments. However, this paradigm (argument generation) allows procedures necessary to avoid myside bias to be quite explicitly presented to the participants. Participants are simply told to generate counterarguments.

In the experiment reported here, we manipulated the instructional set. One group of participants was given decontextualizing instructions and the other group was given non-directive instructions. Decontextualizing instructions direct participants to decouple prior opinion and knowledge when reasoning. Non-directive instructions simply let participants interpret the task as they wish, without any admonition to ignore prior opinion or knowledge. They contain no direction to focus on logic and/or evidence when reasoning.

Note that none of the previous studies examined the influence of individual difference variables and instructional set within a single experiment. It remains an open question whether different types of instructions interact with different levels of the individual difference variables in belief and myside bias tasks. If interactions do exist, it would suggest that different pedagogical approaches might be warranted for different types of students.

In summary, in the present study we investigated both belief bias and myside bias — the former with the traditionally used syllogistic reasoning task and the latter with two informal reasoning tasks (argument generation and experiment evaluation). Examining both types of bias within the same experiment allowed us to test the suggestion in the literature (based on between-experiment comparisons) that cognitive ability is associated with belief bias but not with myside bias. Additionally, we examined whether thinking dispositions such as actively open-minded thinking and need for cognition are predictors of the two biases. In addition to individual difference variables, we examined the influence of instructional set on belief bias and myside bias. Thus, we were able to investigate whether the effects of instructional set were moderated by individual difference variables. Our investigation was distinctive in studying belief bias and myside bias with the same set of participants and in conjoining a manipulation designed to influence the degree of bias with individual difference variables.

2. Methods

2.1. Participants

Participants were 195 students (98 males and 97 females) recruited through poster advertisements from the downtown campus of a large Canadian university. The mean age was 21.1 years ($SD=1.8$). Participants were paid \$20 for their cooperation. Participants were randomly assigned to one of two groups. Half of the sample ($n=98$) received decontextualizing instructions and the other half ($n=97$) received non-directive instructions for all three critical thinking tasks.

2.2. Critical thinking tasks

Participants completed three critical thinking tasks: a syllogistic reasoning task (to measure belief bias), an argument generation task (to measure myside bias), and an experiment evaluation task (to measure myside bias). All participants completed the tasks in the same order.

2.2.1. Belief bias: syllogistic reasoning task

Sixteen syllogistic reasoning problems, drawn from Markovits and Nantel (1989), were completed by the participants. Eight of the problems were worded such that the validity judgment was congruent with the believability of the conclusion (e.g., All fish can swim; tuna are fish; therefore, tuna can swim—which is valid). These were termed the consistent items. Eight of the problems were worded such that the validity judgment was in conflict with the believability of the conclusion (e.g., All flowers have petals; roses have petals; therefore roses are flowers—which is invalid). These were termed the inconsistent items. Note that there were two types of consistent and two types of inconsistent syllogisms. One type of consistent syllogisms had believable premises and conclusions in a logically valid format, while the other type had unbelievable premises and conclusions in a logically invalid format. Therefore, the believability of the content was consistent with the logical format of the syllogism in both types. Correspondingly, one type of inconsistent syllogism had believable premises and conclusions in a logically invalid format, while the other type had unbelievable premises and conclusions in a logically valid format. Therefore, the believability of the content was inconsistent with the logical format of the syllogism in both types.

Definitionally, belief bias is the difference between the number of consistent problems answered correctly and the number of inconsistent items answered correctly. With adult participants, however, performance on the consistent items is constricted because of ceiling effects and the difference score is almost always less reliable than the raw number of inconsistent items answered correctly. For these statistical reasons – and because of theoretical arguments and empirical evidence that performance on the inconsistent items is a better indicator of the ability to override automatically activated prior knowledge (De Neys, 2006; Goel & Dolan, 2003; Kokis et al., 2002; Newstead et al., 2004) – the number of inconsistent items answered correctly will be the primary variable analyzed from this task.

The investigator read the following instructions out loud to participants in the decontextualizing group while they followed along on the printed page:

In the following problems, you will be given two premises, which *you must assume are true*. A conclusion from the premises then follows. You must decide whether the conclusion *follows logically* from the premises or not. You must *suppose that the premises are all true*—even if these statements appear to be false based on your prior knowledge of the world. This is very important. Decide if the conclusion follows logically from the premises, assuming the premises are true, and circle your response.

Verbal emphasis was given to phrases that discussed the need to accept the content of each task as true—even if the content was false based on the participants' knowledge of the world and/or personal beliefs.

The following instructions were provided to the non-directive group; the investigator did not read them aloud nor did the instructions mention the need to put aside knowledge of the world and/or personal beliefs. This protocol was

followed in order to replicate previous research where only written instructions were provided, and this protocol was used for all three critical thinking tasks:

In the following problems, you will be given two premises. A conclusion drawn from the premises is then presented. You must decide whether the conclusion follows from the premises or not.

After each syllogism, participants circled one of the two alternatives: a) Conclusion follows logically from premises; or b) Conclusion does not follow logically from premises.

2.2.2. *Myside bias: argument generation task*

This was an informal reasoning task in which participants were asked to generate written arguments about two different topic statements (cost of tuition and file sharing over the Internet). The cost of tuition statement was worded as follows:

The difference between the actual cost of an undergraduate education at the University of Toronto and what a student pays in tuition is approximately \$8000 per year. The taxpayer pays this \$8000 difference. All U of T undergrads should pay the actual cost of their university education.

The file sharing over the Internet statement was worded as follows:

Music file sharing over the Internet is a growing phenomenon. People should be able to share music over the Internet without paying for the music.

The following instructions were read out loud to participants in the decontextualizing group as participants followed the text on the printed page:

Please carefully read the following and take your time to think about the issue from opposite sides. [Cost of tuition or file-sharing statement was inserted here]. We would like you to put aside your personal beliefs on this issue. Write down arguments both for and against. Write as much as you can, and try to *give both reasons for and reasons against*.

Verbal emphasis was given to the phrases that discussed the need to think about the issue from opposite sides, and to put aside personal beliefs. The following instructions were provided to participants in the non-directive group:

Please read the following and take your time to think about it. [Cost of tuition or file-sharing statement was inserted here]. Write down arguments about this issue. Try to write as much as you can and please feel free to take your time.

For both groups, the statements were followed by eight sections labelled Argument # 1, Argument # 2, ... Argument # 8. Participants were asked to write their arguments on the blank lines.

Participants' prior opinion for both topics was captured in their responses to two items embedded in a thinking dispositions questionnaire (described in a following section). The two items were: "Universities should require students to pay the true cost of a university education (which is considerably higher than what students currently pay in tuition)" and "I think file-sharing music over the Internet should be free." Participants responded on a scale ranging from disagree strongly (1) to agree strongly (6). Scores of 1 to 3 were classified as opposing the tuition and file-sharing topic statements, and scores of 4 to 6 were classified as supporting the tuition and file-sharing topic statements.

Participants were quite adamant in their opposition to the tuition proposition ($M=1.62$; $SD=1.02$), halfway between strongly and moderately disagree, and there was little difference between the means of the decontextualizing group ($M=1.73$; $SD=1.06$) and non-directive group ($M=1.51$; $SD=0.98$), $t(187)=1.49$, ns. Participants supported the file-sharing proposition, but their support was quite moderate ($M=4.55$; $SD=1.29$), halfway between slightly and moderately agree. There was little difference between the means of the decontextualizing group ($M=4.63$; $SD=1.26$) and non-directive group ($M=4.47$; $SD=1.32$), $t(187)=0.86$, ns.

Toplak and Stanovich's (2003) coding scheme was utilized in this study. Participant-generated arguments were scored as either supporting or refuting the two topic statements. For example, the argument that 'taxpayers should not have to foot the bill for those who they don't even know' was classified as supporting the tuition topic statement, whereas the argument 'many students would simply not be able to attend college' was classified as refuting it.

Arguments were then classified as either *myside* or *otherside* depending on the participant's prior belief. For example, 'taxpayers should not have to foot the bill for those who they don't even know' was classified as a *myside* argument if the participant supported the tuition proposition, but as an *otherside* argument if the participant did not support the statement. Arguments that did not support or refute the statements were classified as *miscellaneous*.

Participants generated almost equal numbers of arguments by topic: 1140 for tuition and 1165 for file sharing (2305 arguments in total). There were very similar numbers of arguments generated by type of argument: *myside* arguments (666 for tuition and 678 for file sharing), *otherside* arguments (378 for tuition and 401 for file sharing), and *miscellaneous* arguments (96 for tuition and 86 for file sharing). Just over 92% of the arguments were coded as either *myside* or *otherside*, while just under 8% were coded as *miscellaneous*. Arguments were coded by two independent raters who used the same scoring key and were blind to the participants' prior opinions. There was 95% inter-rater agreement between the coders, and all discrepancies were resolved by a discussion of the differences and agreement on the appropriate coding category.

The *myside* bias index for each topic was calculated by subtracting the number of *otherside* arguments from the number of *myside* arguments.

The order of presentation of the two topics (tuition and file sharing) was counterbalanced. Half of the participants received the tuition topic first followed by the file-sharing topic (order 1), while the other half received the reverse order (order 2). Six participants did not complete the task consistent with the instructions provided¹ and were excluded from the analyses, resulting in a sample size of 189 for this task.

2.2.3. *Myside bias: experiment evaluation task*

This task was an informal reasoning task modeled on a paradigm used by Klaczynski (1997, 2000; Klaczynski & Gordon, 1996; Klaczynski et al., 1997; Klaczynski & Lavalley, 2005). Participants evaluated two experiments about the relationship between a parental characteristic (stay-at-home or working mother) and the behavior of offspring (the social responsibility of the child).

Participants' prior opinion on this topic was captured in the response to an item embedded in a thinking dispositions questionnaire (described in a following section). The item was: "I think that children who were raised by mothers who stayed at home grow up to be more socially responsible than children who were raised by mothers who worked outside the home." Participants responded on a scale ranging from disagree strongly (1) to agree strongly (6). Scores of 1 to 3 were classified as 'pro working mother,' and scores of 4 to 6 as 'pro stay-at-home mother.'

The two experiments had opposing conclusions; one claimed that children raised by stay-at-home mothers were more socially responsible than children raised by working mothers, while the other claimed the opposite. Subsequent to reading the description of each experiment, participants responded to three evaluative questions using a six-point Likert rating scale ranging from unfavorable assessments (1) to favorable assessments (6). The evaluative questions concerned: the strength of the conclusion (How strongly is the conclusion supported by the results of the experiment?); the quality of the experiment (What is your overall evaluation of the quality of this experiment?); and the persuasiveness of the experiment (How persuasive is this experiment?). The ratings for the three questions were summed to yield a total experiment evaluation score for each experiment.

The following instructions were read out loud to the decontextualizing group as participants followed the text on the printed page. Verbal emphasis was provided to the phrases that discussed the need to put aside personal beliefs and base the evaluation only on the information provided in the summary:

We are interested in what you think about the quality and persuasiveness of some of these interesting experiments. In the following pages we have summarized two experiments. For each experiment, you will first read the conclusion of that experiment followed by a very brief summary of what was done in that experiment. We would like you to put aside your personal beliefs on this issue, and then read the summaries and evaluate each experiment on the subsequent scales. It is important to base your evaluation *only* on the information provided in the summary, and not on your personal beliefs about this matter.

¹ Five participants in the non-directive condition only generated arguments that agreed with the conclusion of the topic statements—despite holding prior beliefs that disagreed with the conclusion of the topic statements. As such, the direction and magnitude of their *myside* bias scores were affected, and the participants were excluded from the analyses of this task. One participant failed to complete the task.

The same instructions were provided to participants in the non-directive group *except* for the final sentence. The order of the two experiments (pro stay-at-home or pro working mother first) was counterbalanced (order 1: pro working mother conclusion first followed by the pro stay-at-home mother conclusion; order 2 was the reverse).

Both experiments shared a confound — testers of a different gender in the experimental and control groups. The experiment with a pro stay-at-home conclusion contained an additional confound (different time of testing for experimental and control groups), which was overlooked during the design of this study. Therefore, the pro stay-at-home mother experiment was objectively worse than the pro working mother experiment. This imbalance would inflate the myside bias score of participants with pro working mother prior opinions and attenuate the myside bias score of the pro stay-at-home mother group. This is because the pro stay-at-home mother experiment would likely be given lower ratings than the pro working mother experiment—regardless of the potential influence of prior opinions. Thus, it was decided that all the analyses for this task would be limited to participants with pro working mother prior beliefs. Given that almost 75% of participants ($n=144$) were pro working mother, sufficient statistical power was retained.

The mean for the prior beliefs of the restricted sample was 1.81 ($SD=0.78$), close to the value of 2, which represented moderately pro working mother. The decontextualizing group had a mean of 1.73 ($SD=0.76$) and the non-directive group had a mean of 1.89 ($SD=0.79$). There was no between-group difference: $t(142)=-1.27, p>.05$.

2.3. Cognitive ability measure

Participants completed a short form of the Wechsler Abbreviated Scale of Intelligence (WASI) consisting of the vocabulary (verbal measure) and matrix reasoning (nonverbal measure) subtests. Previous research has shown that these two subtests correlate strongly (vocabulary, $r=.88$) and moderately (matrix reasoning, $r=.66$) with the corresponding subtests on the Wechsler Adult Intelligence Scale (WAIS-III), achieve consistently high reliabilities, and are good measures of psychometric g (Wechsler, 1999). Scores on these two subtests can be pro-rated to obtain a short form IQ score that correlates strongly with WAIS-III Full Scale IQ ($r=.87$). Accordingly, these two subtests were combined to provide a full-scale equivalent IQ score (pro-rated IQ) following the procedures provided in the WASI manual.

2.4. Thinking dispositions questionnaire

Participants completed a thinking dispositions questionnaire in which they responded using a six-point Likert rating scale (1 = strongly disagree; 2 = moderately disagree; 3 = slightly disagree; 4 = slightly agree; 5 = moderately agree; 6 = strongly agree). Sixty-eight individual items from the three different scales discussed below were randomly intermixed.

2.4.1. Actively open-minded thinking scale (AOT)

The actively open-minded thinking scale was composed of 41 items drawn from a variety of sources: 10 items from a flexible thinking scale developed by Stanovich and West (1997); 8 items from the Openness-Values facet of the Revised NEO Personality Inventory (Costa & McCrae, 1992); 9 items measuring dogmatism (Paulhus & Reid, 1991; Robinson, Shaver, & Wrightsman, 1991; Troidahl & Powell, 1965); 3 items from the categorical thinking subscale of Epstein and Meier's (1989) constructive thinking inventory; 9 items from the belief identification scale developed by Sá et al. (1999); 2 items from a counterfactual thinking scale developed by Stanovich and West (1997). All items were scored in the direction 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," "No one can talk me out of something I know is right" (reverse scored), "I tend to classify people as either for me or against me" (reverse scored), "Certain beliefs are just too important to abandon no matter how good a case can be made against them" (reverse scored), and "My beliefs would not have been very different if I had been raised by a different set of parents" (reverse scored). The score on the scale was obtained by summing the responses to the 41 items (mean=163.1; $SD=19.3$). Cronbach's alpha was .84.

2.4.2. Superstitious thinking scale

This scale consisted of nine items from three related subscales. Four items that reflected conceptions of luck were used, e.g., "The number 13 is unlucky" (Stanovich & West, 1998). Two items from the Paranormal Belief Scale (Jones, Russell, & Nickel, 1977) were included ("It is advisable to consult your horoscope daily"); and three items from the

Table 1
Intercorrelations between the bias indices and the individual difference variables

	Inconsistent syllogisms correct ($n=195$)	Myside bias: argument generation task ($n=189$)	Myside bias: experiment evaluation task ($n=144$)
Cognitive ability			
Pro-rated IQ	.41***	-.04	.08
Thinking dispositions			
Need for cognition	.19**	-.01	.18*
AOT	.21**	-.03	.16
Superstitious thinking	-.23**	-.01	-.04

* $p < .05$; ** $p < .01$; *** $p < .001$ (two-tailed).

Note: AOT = actively open-minded thinking scale.

Superstitious Thinking scale (Epstein & Meier, 1989) (“I do not believe in any superstitions,” reverse scored). Higher scores on the scale indicated greater belief in superstitious thinking. The score on the scale was obtained by summing the responses to the nine items (mean = 19.0; SD = 7.1). Cronbach’s alpha was .79.

2.4.3. Need for cognition

The 18 item need for cognition scale published by Cacioppo, Petty, Feinstein, and Jarvis (1996) was employed in this study. Sample items included: “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 score on the scale was obtained by summing the responses to the 18 items (mean = 77.6; SD = 12.6). Cronbach’s alpha was .88.

2.5. Procedure

Participants completed the tasks in the following order: informed consent, demographic questionnaire, syllogistic reasoning task, WASI, argument generation task, thinking dispositions questionnaire, experiment evaluation task.²

3. Results

When the results were collapsed across the two groups (decontextualizing and non-directive), there were significant indications of belief bias or myside bias on each of the three tasks examined in this investigation.

Significant belief bias was displayed on the syllogistic reasoning task. The mean number of consistent syllogisms answered correctly (7.19, SD = 1.10) was significantly higher than the mean number of inconsistent syllogisms answered correctly (6.05, SD = 1.86), $t(194) = 9.42$, $p < .001$.

Significant myside bias was displayed on the argument generation task. For the tuition topic, the mean number of myside arguments (3.52, SD = 1.62) was significantly higher than the mean number of otherside arguments (1.85, SD = 1.29), $t(188) = 10.29$, $p < .001$. For the file-sharing topic, the mean number of myside arguments (3.44, SD = 1.51) was also significantly higher than the mean number of otherside arguments (2.12, SD = 1.31), $t(188) = 8.71$, $p < .001$.

Finally, there was significant myside bias displayed on the experiment evaluation task (although qualified by the confound mentioned in the Methods section, which impeaches the mean comparison reported here but not the correlational analyses to follow). The mean rating given to the pro working mother experiment (9.56, SD = 3.54) was significantly higher than the mean rating given to the pro stay-at-home mother experiment (7.28, SD = 3.32), $t(143) = 6.49$, $p < .001$.

Table 1 presents the correlations between the individual differences variables (cognitive ability and thinking dispositions) and the bias indices for each task. As mentioned in the Methods section, for the syllogistic reasoning task, the number of inconsistent syllogisms answered correctly is the best correlational measure of the ability to overcome

² Thinking dispositions questionnaire was placed between the argument generation and experiment evaluation tasks to avoid participants completing two myside tasks in a row.

Table 2

Mean scores (standard deviations in parentheses) on the three critical thinking tasks as a function of instructional set

	Decontextualizing	Non-directive	<i>t</i> -test	Effect size Cohen's <i>d</i>
Inconsistent syllogisms correct ^a	6.27 (1.73)	5.83 (1.97)	$t(193)=1.66^*$.24
Myside bias: argument generation task ^b	1.33 (2.22)	4.74 (3.82)	$t(187)=-7.55^{**}$	1.09
Myside bias: experiment evaluation task ^c	2.61 (4.20)	1.96 (4.23)	$t(142)=.93$.15

* $p < .05$ (one-tailed); ** $p < .001$.^a $N=195$, decontextualizing=98, non-directive=97.^b $N=189$, decontextualizing=97, non-directive=92.^c $N=144$, decontextualizing=70, non-directive=74.

the biasing effects of prior knowledge. This score should be positively correlated with all of the individual difference variables except superstitious thinking (with which it would be expected to show a negative correlation). The myside bias indicator for the argument generation task was the total number of myside arguments produced for the tuition and file-sharing topics combined minus the total number of otherside arguments produced for the tuition and file-sharing topics combined. This index should be negatively correlated with the individual difference variables with the exception of superstitious thinking. The myside bias index for the experiment evaluation task was the total experiment evaluation score (combined rating given to the three evaluative questions) on the pro working mother experiment minus the total experiment evaluation score for the pro stay-at-home mother experiment. This index should show correlational patterns similar to those displayed for the argument generation task.

As indicated in Table 1, pro-rated IQ and all three thinking dispositions displayed significant correlations with performance on the syllogistic reasoning task and all were in the expected direction. This outcome was in complete contrast to the results for the argument generation task. None of the individual difference variables displayed significant correlations with myside bias on this task. In fact, all the correlations were close to zero. Likewise, myside bias in the experiment evaluation task failed to correlate with pro-rated IQ. One of the three correlations with thinking dispositions was significant (need for cognition) but it was in the opposite direction than that expected. Thus, the individual differences variables largely failed to predict the degree of myside bias on the two informal reasoning tasks. Importantly though, the pattern of correlations displayed in this study was consistent with the trend discussed in the Introduction. The ability to overcome the effects of belief bias (i.e., knowledge bias) was significantly related to cognitive ability in a formal reasoning task, but the ability to overcome myside bias in informal reasoning tasks was independent of cognitive ability.³

The next series of analyses examined whether the manipulation of instructional set was a predictor of the bias indices. Table 2 presents the mean bias scores for each task as a function of instructional set (decontextualizing versus non-directive instructions). The means were in the expected direction for two of the three tasks (syllogistic reasoning and argument generation tasks). The effect for the experiment evaluation task was not significant and, in any case, it was in the wrong direction. The effect of instructional set was significant for the argument generation task, and its effect size was quite large (Cohen's $d=1.09$). Instructional set had a small effect size in the syllogistic reasoning task (Cohen's $d=0.24$), and the difference on this task was only significant on a one-tailed test.

The analyses presented in Tables 1 and 2 have indicated that both the manipulated variable of instructional set and the individual difference variable of cognitive ability were associated with performance on the inconsistent syllogisms. The results displayed in Table 3 indicate that the effects of these two factors might be in an interactive relationship. The table presents the means for the inconsistent syllogisms as a function of instructional set and cognitive ability after the latter variable had been partitioned into four quartiles. Participants were assigned to one of four cognitive ability groups based on their pro-rated IQ score: Quartile 1 (scores from 93 to 110), Quartile 2 (scores from 111 to 117), Quartile 3 (scores from 118 to 124), and Quartile 4 (scores of 125 and above). A perusal of Table 3 indicates the locus of the interactive effects of instructional set and cognitive ability. Specifically, the nature of the instructions appears to make

³ Analyses were also run investigating the influence of the most extreme prior belief scores on the relationships between the bias scores and the individual difference variables for the argument generation task. For the argument generation task, analyses using participants with a prior belief score of 1 ($n=120$) on the tuition proposition (strongly disagree that students should pay the true cost of tuition) revealed similar relationships to those found for all participants. Analyses using participants with a prior belief score of 6 ($n=56$) for the file-sharing proposition (strongly agree that file sharing should be free) revealed similar relationships to those found for all participants.

Table 3

Mean performance (standard deviations in parentheses) on the number of inconsistent syllogisms correct as a function of cognitive ability and instructional set ($N=195$)

	Decontextualizing instructions	Non-directive instructions
<i>Cognitive ability quartile</i>		
Quartile 1 ($n=51$)	5.64 (1.57)	4.09 (2.02)
Quartile 2 ($n=46$)	5.41 (2.40)	5.90 (1.66)
Quartile 3 ($n=49$)	6.80 (1.38)	6.38 (1.77)
Quartile 4 ($n=49$)	6.93 (1.27)	7.00 (1.18)

little difference for the groups in the highest three quartiles. This is why the overall effect of instructional set was weak. However, participants in the lowest quartile of cognitive ability were markedly affected by the instructional set—answering over one and a half more items correctly when given the decontextualizing instructions (an improvement of 38.1%). A 2 (instructional set) \times 4 (cognitive ability) ANOVA confirmed this impression, as the interaction between instructional set and cognitive ability was statistically significant, $F(3, 187)=3.38, p<.05$.

The same trend was apparent in the data of the argument generation task but it did not quite attain statistical significance. Table 4 presents the means of the number of myside and otherside arguments generated as a function of instructional set and cognitive ability after the latter variable had been partitioned into the same four quartiles as in the previous analysis. A perusal of Table 4 indicates a trend for the decontextualizing instructions to attenuate the myside bias effect (myside arguments minus otherside arguments) more for the lowest quartile than for the other three. A 2 (instructional set) \times 4 (cognitive ability) \times 2 (argument type) ANOVA was run to test this trend. However, the three-way interaction between instructional set, cognitive ability, and argument type did not quite reach conventional levels of statistical significance, $F(3, 181)=1.82, p=.145$.

4. Discussion

The results of the present study indicated that belief bias was more strongly related to the individual difference variables measured here than was myside bias. Cognitive ability was a moderate predictor of the ability to overcome belief bias and several thinking dispositions (need for cognition, actively open-minded thinking, and superstitious thinking) were significant (albeit weak) predictors. In contrast, cognitive ability displayed near zero correlations with myside bias as measured in two different paradigms. Five of six correlations involving myside bias and a thinking disposition were non-significant and the sign of sixth was in the unexpected direction.

On the face of it, the failure to find an association between cognitive ability and myside bias seems somewhat surprising. Ever since Spearman (1904) first discovered positive manifold, intelligence indicators have correlated with a plethora of cognitive/personality traits and thinking abilities (e.g., Ackerman, Kyllonen, & Richards, 1999; Deary, 2001; Deary, Whiteman, Starr, Whalley, & Fox, 2004; Geary, 2005; Lubinski, 2000, 2004; Lubinski & Humphreys, 1997). The fact that the two myside bias tasks were informal reasoning paradigms might be critical here (see Neuman, Weinstock, & Glasner, 2006). In such informal reasoning paradigms, participants of high cognitive ability may be no more likely to *recognize* the need for unbiased processing than are participants of low cognitive ability (Stanovich & West, 2007). However, if this were the case, it might be thought that thinking dispositions (toward more open-minded

Table 4

Mean performance (standard deviations in parentheses) on the argument generation task as a function of cognitive ability and instructional set ($N=189$)

	Decontextualizing		Non-directive	
	Myside arguments generated	Otherside arguments generated	Myside arguments generated	Otherside arguments generated
<i>Cognitive ability quartile</i>				
Quartile 1 ($n=51$)	5.86 (1.98)	5.04 (1.64)	7.65 (3.58)	1.65 (1.70)
Quartile 2 ($n=43$)	6.94 (1.56)	5.35 (1.62)	7.46 (2.82)	3.27 (2.15)
Quartile 3 ($n=48$)	6.64 (1.98)	5.44 (1.83)	6.74 (2.73)	2.52 (2.21)
Quartile 4 ($n=47$)	6.81 (2.54)	5.00 (1.75)	7.90 (3.19)	3.30 (1.75)

thinking or toward more thinking in general) would have been more potent predictors in our study, and this was not the case.⁴

Alternatively, the content used in the tasks might have accounted for the difference. The formal reasoning task required decoupling from prior personal knowledge of the world, whereas the informal reasoning tasks required decoupling from prior opinions about controversial issues (tuition levels, file sharing, and working mothers). As noted in the Introduction, these findings suggest that reasoning ability seems to be influenced by the subject an individual is reasoning about. When the subject is related to personal opinions rather than factual knowledge of the world, even individuals with greater cognitive ability had difficulty overriding their personal opinions.

Despite the failure of individual difference variables to predict the degree of myside bias, instructional set was found to be a significant predictor, at least in one task. The differential effectiveness of the instructional manipulation across the three tasks is perhaps explainable in terms of how well it can be specified just what unbiased processing is in each of the three tasks. In the argument generation task, it is possible to give the participant quite explicit instructions regarding what would be considered unbiased responding (e.g., “give both reasons for *and* reasons against”). Correspondingly, this task displayed the strongest effect of instructions. In contrast, the task that displayed the weakest effect of instructions – the experiment evaluation task – is the most obscure in indicating just what the participant should do to avoid bias (e.g., “base your evaluation *only* on the information provided in the summary, and not on your personal beliefs about this matter”). In between these two tasks is the syllogistic reasoning task where the decontextualizing instructions are more explicit than in the experiment evaluation task, but not as explicit as those in the argument generation task. This task did not display a significant effect of instructions overall, but there were indications that the instructions were effective for the lowest ability group.

Results consistent with this explanation have been reported by Neuman et al. (2006). In their research, participants were provided with dialogues which contained some informal reasoning fallacies. Neuman et al. (2006) manipulated the context of the dialogues—in one condition the dialogue was presented as a reasoned dialogue and in the other it was presented as a quarrel or debate. The former group was more likely to identify the informal reasoning fallacies, thus indicating that the context of the argument effected argument evaluation. The ability of our training instructions to induce different contexts might have varied from task to task, resulting in variation in the size of the reasoning biases. Additionally, the different tasks themselves might have represented contexts of varying demand for unbiased reasoning. Some of these factors might have mitigated the influence of individual difference variables.

The trend in the syllogistic reasoning task for the group lowest in cognitive ability to be more affected by the manipulation of instructions than the higher ability group was also apparent in the data from the argument generation task, although the latter did not attain statistical significance. For these two tasks perhaps, instructions and cognitive ability worked together to facilitate reasoning in the presence of potential bias. This pattern supports the conjecture that increased cognitive ability serves an override function when a cognitive tool such as decontextualizing instructions is absent. Nonetheless, the low quartile participants who were advised to concentrate on the logical validity (or on evensided argumentation) and not the believability of the content, were able to override their biases related to prior knowledge and opinion.

It remains an open question as to why no significant correlations were found between the thinking dispositions and myside bias scores. As mentioned in the Introduction, some previous research has found a relationship between thinking dispositions associated with epistemic regulation and individual differences in myside bias (Klaczynski et al., 1997; Klaczynski & Lavalley, 2005; Klaczynski & Robinson, 2000; Sá et al., 2005). Additionally, Mason and Scirica (2006) used an informal reasoning task and found that an individual difference variable measuring epistemological understanding predicted the ability to generate arguments, counterarguments and rebuttals about controversial topics, as well as the quality of the three types of arguments. Weinstock, Neuman, and Tabak (2004) administered a similar measure of epistemology (familiarity with argumentation norms) and an informal reasoning task (ability to identify fallacious arguments). Participants who were familiar with the norms outperformed those who were unfamiliar on the informal reasoning task. Kuhn and Weinstock (2002) also found that individual differences in the quality of argument

⁴ It was not the case that cognitive ability and thinking dispositions were unable to account for variance in other variables. Both predicted performance on the inconsistent syllogisms, as indicated in the Results section. To take another example, pro-rated IQ and actively open-minded thinking were significant predictors of superstitious thinking (multiple $R = .436$, $p < .001$) and both were significant independent predictors in the simultaneous regression equation: pro-rated IQ standardized beta = $-.163$, $F(1, 192) = 5.59$, $p < .05$ (unique variance explained = $.024$); actively open-minded thinking standardized beta = $-.353$, $F(1, 192) = 26.27$, $p < .001$ (unique variance explained = $.111$).

generation and in epistemological understanding predicted outcome measures in a juror reasoning task (see also, Kuhn, Weinstock, & Flaton, 1994). These findings suggest that the inclusion of a measure of epistemological understanding – in addition to the AOT – would provide further insight into the relationship between myside bias and epistemic regulation.

References

- Ackerman, P., Kyllonen, P. & Richards, R. (Eds.). (1999). *Learning and individual differences: Process, trait, and content determinants*. Washington, DC: American Psychological Association.
- 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. (1995). Myside bias in thinking about abortion. *Thinking and Reasoning*, 1, 221–235.
- Baron, J. (2000). *Thinking and deciding*, (Third Edition) Cambridge, MA: Cambridge University Press.
- Brody, N. (1997). Intelligence, schooling, and society. *American Psychologist*, 52, 1046–1050.
- 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.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory*. Odessa, FL: Psychological Assessment Resources.
- Deary, I. J. (2001). *Intelligence: A very short introduction*. Oxford: Oxford University Press.
- Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology*, 86, 130–147.
- De Neys, W. (2006). Dual processing in reasoning—Two systems but one reasoner. *Psychological Science*, 17, 428–433.
- Dias, M., Roazzi, A., & Harris, P. L. (2005). Reasoning from unfamiliar premises: A study with unschooled adults. *Psychological Science*, 16, 550–554.
- 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: W.H. Freeman.
- Ennis, R. H. (1995). *Critical thinking*. New York: Prentice Hall.
- Epstein, S., & Meier, P. (1989). Constructive thinking: A broad coping variable with specific components. *Journal of Personality and Social Psychology*, 57, 332–350.
- Evans, J. St. B. T. (2002a). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin*, 128, 978–996.
- Evans, J. St. B. T. (2002b). The influence of prior belief on scientific thinking. In P. Carruthers, S. Stich, & M. Siegal (Eds.), *The cognitive basis of science* (pp. 193–210). Cambridge: Cambridge University Press.
- Evans, J. St. B. T., Barston, J., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and 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 and Reasoning*, 11, 382–389.
- Evans, J. St. B. T., Newstead, S., Allen, J., & Pollard, P. (1994). Debiasing by instruction: The case of belief bias. *European Journal of Cognitive Psychology*, 6, 263–285.
- Geary, D. C. (2005). *The origin of the mind: Evolution of brain, cognition, and general intelligence*. Washington, DC: American Psychological Association.
- Gilinsky, A., & Judd, B. B. (1994). Working memory and bias in reasoning across the life span. *Psychology and Aging*, 9, 356–371.
- Goel, V., & Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition*, 87, B11–B22.
- Gottfredson, L. S. (1997). Why *g* matters: The complexity of everyday life. *Intelligence*, 24, 79–132.
- Greenhoot, A. F., Semb, G., Colombo, J., & Schreiber, T. (2004). Prior beliefs and methodological concepts in scientific reasoning. *Applied Cognitive Psychology*, 18, 203–221.
- Handley, S. J., Capon, A., Beveridge, M., Dennis, I., & Evans, J. St. B. T. (2004). Working memory, inhibitory control and the development of children's reasoning. *Thinking and Reasoning*, 10, 175–195.
- Jones, W., Russell, D., & Nickel, T. (1977). Belief in the paranormal scale: An objective instrument to measure belief in magical phenomena and causes. *JSAS Catalog of Selected Documents in Psychology*, 7 (100 (Ms. No. 1577)).
- 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. (2000). Motivated scientific reasoning biases, epistemological beliefs, and theory polarization: A two-process approach to adolescent cognition. *Child Development*, 71, 1347–1366.
- 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., Gordon, D. H., & Fauth, J. (1997). Goal-oriented critical reasoning and individual differences in critical reasoning biases. *Journal of Educational Psychology*, 89, 470–485.
- Klaczynski, P. A., & Lavalley, 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. (1991). *The skills of argument*. Cambridge: Cambridge University Press.
- Kuhn, D., & Weinstock, M. (2002). What is epistemological thinking and why does it matter? In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 121–144). Mahwah, NJ: Lawrence Erlbaum.
- Kuhn, D., Weinstock, M., & Flaton, R. (1994). How well do jurors reason? Competence dimensions of individual variation in a juror reasoning task. *Psychological Science*, *5*, 289–296.
- Lubinski, D. (2000). Scientific and social significance of assessing individual differences: “Sinking shafts at a few critical points”. *Annual Review of Psychology*, *51*, 405–444.
- Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after Spearman’s (1904) “General Intelligence, Objectively Determined and Measured”. *Journal of Personality and Social Psychology*, *86*, 96–111.
- Lubinski, D., & Humphreys, L. G. (1997). Incorporating general intelligence into epidemiology and the social sciences. *Intelligence*, *24*, 159–201.
- Markovits, H., & Nantel, G. (1989). The belief-bias effect in the production and evaluation of logical conclusions. *Memory and Cognition*, *17*, 11–17.
- Mason, L., & Scirica, F. (2006). Prediction of students’ argumentation skills about controversial topics by epistemological understanding. *Learning and Instruction*, *16*, 492–509.
- Neuman, Y., Weinstock, M., & Glasner, A. (2006). The effect of contextual factors on the judgment of informal reasoning fallacies. *Quarterly Journal of Experimental Psychology*, *59*, 411–425.
- Newstead, S. E., Handley, S. J., Harley, C., Wright, H., & Farrelly, D. (2004). Individual differences in deductive reasoning. *Quarterly Journal of Experimental Psychology*, *57A*, 33–60.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220.
- Norris, S. P., & Ennis, R. H. (1989). *Evaluating critical thinking*. Pacific Grove, CA: Midwest Publications.
- Nussbaum, E. M., & Kardash, C. M. (2005). The effects of goal instructions and text on the generation of counterarguments during writing. *Journal of Educational Psychology*, *97*, 157–169.
- 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*(1), 4–14.
- 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). Hillsdale, NJ: Erlbaum.
- Paulhus, D. L., & Reid, D. B. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, *60*, 307–317.
- Perkins, D. N. (1985). Postprimary education has little impact on informal reasoning. *Journal of Educational Psychology*, *77*, 562–571.
- 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.
- Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (1991). *Measures of personality and social psychological attitudes, vol. 1* (pp. 560–564). San Diego, CA: Academic Press.
- 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.
- Sá, W. C., 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.
- 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.
- Spearman, C. (1904). General intelligence, objectively determined and measured. *American Journal of Psychology*, *15*, 201–293.
- Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Stanovich, K. E. (2004). *The robot’s rebellion: Finding meaning in the age of Darwin*. Chicago: University of Chicago Press.
- 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. (2007). Natural myside bias is independent of cognitive ability. *Thinking and Reasoning*, *13*.
- Sternberg, R. J. (1997). *Thinking styles*. Cambridge: 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: Cambridge University Press.
- 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.
- Torrens, D., Thompson, V. A., & Cramer, K. M. (1999). Individual differences and the belief bias effect: Mental models, logical necessity, and abstract reasoning. *Thinking and Reasoning*, *5*, 1–28.
- Troldahl, V., & Powell, F. (1965). A short-form dogmatism scale for use in field studies. *Social Forces*, *44*, 211–215.
- Wade, C., & Tavis, C. (1993). *Critical and creative thinking*. New York: Harper Collins.
- Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. Toronto, ON: The Psychological Corporation, Harcourt Brace Jovanovich, Inc.
- Weinstock, M., Neuman, Y., & Tabak, I. (2004). Missing the point or missing the norms? Epistemological norms as predictors of students’ ability to identify fallacious arguments. *Contemporary Educational Psychology*, *29*, 77–94.