Research Article

Dual Processing in Reasoning

Two Systems but One Reasoner

Wim De Neys

University of Leuven, Leuven, Belgium

ABSTRACT—*Human reasoning has been characterized as an interplay between an automatic belief-based system and a demanding logic-based reasoning system. The present study tested a fundamental claim about the nature of individual differences in reasoning and the processing demands of both systems. Participants varying in working memory capacity performed a reasoning task while their executive resources were burdened with a secondary task. Results were consistent with the dual-process claim: The executive burden hampered correct reasoning when the believability of a conclusion conflicted with its logical validity, but not when beliefs cued the correct response. However, although participants with high working memory spans performed better than those with lower spans in cases of a conflict, all reasoners showed similar effects of load. The findings support the idea that there are two reasoning systems with differential processing demands, but constitute evidence against qualitative individual differences in the human reasoning machinery.*

Human reasoners are prone to various errors in logical reasoning and judgment. No matter how small the probability of a new terrorist attack on the 11th day of the ninth month of the year may be, for example, many people will nevertheless refrain from flying on that specific day. They believe that taking the car is the safer option. However, the resulting increased risk of dying in a fatal traffic accident may actually outweigh the risks associated with flying (e.g., Gigerenzer, 2004). People's objective risk judgment is thus biased by the vividness of their recollection of the collapsing twin towers. This common human tendency to base judgments on prior beliefs and intuition rather than on a logical reasoning process biases performance in many classic reasoning tasks (Evans, 2002; Kahneman, Slovic, & Tversky, 1982).

Address correspondence to Wim De Neys, Post Doctoraal Onderzoeker FWO, Lab Experimentele Psychologie, K.U. Leuven, Tiensestraat 102, 3000 Leuven, Belgium, e-mail: wim.deneys@psy.kuleuven.be.

Influential dual-process theories of thinking have explained people's "rational thinking failure" by positing two different human reasoning systems (e.g., Epstein, 1994; Evans, 2003; Evans & Over, 1996; Sloman, 1996; Stanovich & West, 2000). Dual-process theories come in many flavors, but generally they assume that a first system (often called the heuristic system) will tend to solve a problem by relying on prior knowledge and beliefs, whereas a second system (often called the analytic system) allows reasoning according to logical standards. The heuristic default system is assumed to operate rapidly and automatically, whereas the operations of the analytic system are believed to be slow and heavily demanding of people's computational resources.

Dual-process theories state that the heuristic and analytic systems will often act in concert. Hence, on these occasions, the heuristic default system will provide fast, frugal, and correct conclusions. However, prepotent heuristics can also bias reasoning in situations that require more elaborate, analytic processing. That is, the two systems will sometimes conflict and cue different responses. In these cases, the analytic system will need to override the belief-based response generated by the heuristic system (Stanovich & West, 2000). The inhibition of the heuristic system and the computations of the analytic system are assumed to draw on limited executive working memory (WM) resources. Therefore, correct analytic reasoning in the case of a conflict between belief and logic would be characteristic of individuals who are highest in WM span: The more resources that are available, the more likely that the analytic system will be successfully engaged and the correct response will be calculated.

Paradigmatic support for the dual-process framework has come from individual differences studies on belief bias in syllogistic reasoning. Belief bias refers to the intuitive tendency to judge the validity of a syllogism by evaluating the believability of the conclusion (Oakhill, Johnson-Laird, & Garnham, 1989). For some problems, referred to as *conflict items*, the logical status of the conclusion conflicts with background beliefs. Consider, for example, the following syllogism: "All mammals can walk. Whales are mammals. Therefore, whales can walk." In this case, the conclusion is valid but unbelievable. The heuristic, belief-based system thus triggers a logically erroneous

response, and providing the correct response requires demanding analytic computations. Consequently, many people fail to solve conflict items correctly. For other syllogisms, referred to as *no-conflict items*, the logical status of the conclusion is consistent with the believability of the conclusion. Consider, for example, the following syllogism: "All fruits can be eaten. Hamburgers can be eaten. Therefore, hamburgers are fruit." In this case, the conclusion is both invalid and unbelievable, and the heuristic system cues the correct response. Rates of correct solution are uniformly high for these no-conflict items.

Studies consistent with the dual-process framework (e.g., Newstead, Handley, Harley, Wright, & Farrelly, 2004; Stanovich & West, 2000) have shown that individual differences in cognitive (WM) capacity predict performance on conflict items, but not on no-conflict items. Indeed, the heuristic system is assumed to operate automatically; that is, it should not burden limited executive resources. Hence, when the heuristic, belief-based response is consistent with the logical response, even people with low WM spans will get the right answer, because they can rely on heuristic reasoning. However, on conflict items, only people with higher spans will manage to block the heuristic system and reason analytically to arrive at the logically correct answer.

Older studies had already indicated that solving logical reasoning tasks requires cognitive resources (e.g., Kyllonen & Christal, 1990; Stanovich & West, 1998). The individual differences studies on belief bias made it clear that correct reasoning can also be nondemanding. The belief-bias studies suggest that a reasoning process can have different cognitive demands depending on whether or not beliefs and logic conflict. This pattern fits the dual-process assumptions (Stanovich & West, 2000).

Despite these individual differences studies, the dual-process framework has been criticized severely (e.g., De Neys, in press; Gigerenzer & Regier, 1996; Reyna, Lloyd, & Brainerd, 2003; see also commentaries in Stanovich & West, 2000). A crucial critique concerns the fact that the framework has focused exclusively on people's response output and not on the underlying cognitive processes. This approach may result in dramatic confounds. For example, the individual differences studies were purely correlational, and the reported correlations do not establish the assumed causality: The findings indicate that selecting the correct, analytic conclusion in case of a conflict is associated with having a larger executive resource pool, but this does not imply that the executive resources are necessary for the analytic operations. Some other factor might account for the positive association (e.g., Klaczynski, 2000; Newton & Roberts, 2003; Sternberg, 2000). In other words, the framework's basic processing assumption, the differential involvement of executive WM resources in heuristic and analytic reasoning, lacks experimental support.

Furthermore, the individual differences studies are not clear about the nature of the established diversity in reasoning performance. The observed differences between groups varying in cognitive capacity are open to a quantitative or a qualitative interpretation (e.g., Moshman, 2004). Bluntly put, the issue boils down to whether the performance differences are due to differences in the efficiency with which a universal reasoning machinery operates or due to individual differences in the machinery itself. A weak, quantitative interpretation entails that all people have the two reasoning systems at their disposal and that all reasoners engage in analytic reasoning in the case of a belief-logic conflict. According to this interpretation, the machinery itself does not vary. Because of the demanding nature of the analytic operations, people with higher spans would simply be more likely to complete the analytic process successfully. A strong, qualitative interpretation, however, entails that people with low spans, unlike those with high spans, have no access to an analytic system, so that their reasoning is completely mediated by the heuristic system. According to this interpretation, there are two types of human reasoners with different cognitive reasoning architectures.[1]

The present study introduces the use of a secondary task to investigate the processing shortcomings uncovered in previous dual-process research. Participants solved a standard syllogistic reasoning task with conflict and no-conflict problems while their executive resources were burdened by the memorization of a dot pattern. Using a secondary task allows a direct test of the dual-process framework. On no-conflict problems, the heuristic belief-based system is assumed to trigger the correct response. If this system operates automatically, it should not be hindered by an executive burden. On conflict problems, in contrast, deriving the correct conclusion is assumed to require analytic reasoning. If this system draws on executive resources, it should be hindered by an executive burden. Consequently, the dual-process framework predicts that when participants are under load, reasoning performance on conflict problems should decrease, whereas performance on no-conflict problems should be unaffected.

All participants also completed a measure of executive WM capacity. Both the weak and the strong interpretations of individual differences in performance predict overall high performance on no-conflict problems and better performance for people with high spans than for people with low spans on conflict problems. The interpretations can be disentangled, however, by examining how load affects performance on conflict items. If individuals with high and low spans rely on qualitatively different reasoning systems, then the effects of load should differ between the two groups. Indeed, if reasoning among people with low spans is already based on purely automatic heuristic processing in the absence of a load, the additional secondary-task load should not affect their performance. If the difference between the groups is instead quantitative, and people with low spans do engage in executive analytic reasoning under standard conditions, burdening their resources should further decrease

[1] In the reasoning literature, there have been some attempts to unify the two systems identified by dual-process theorists under a single system (e.g., Osman, 2004). Note that the present study does not speak to this issue. Both the weak and the strong positions assume that there are two different reasoning systems.

their performance. Hence, the strong interpretation predicts an effect of load only among individuals with high spans, whereas the weak interpretation predicts an effect for both groups.

As do many cognitive scientists, most dual-process theorists assume that executive WM resources are the quintessential component of computational cognitive capacity (e.g., Evans, 2003; Stanovich & West, 2000). Therefore, the present study adopted a capacity measure and secondary task specifically designed to measure and burden executive WM resources.

In sum, in the present study, reasoners varying in WM span solved syllogisms while executive burden was manipulated. Three levels of load were adopted (no load, low load, and high load). Examining how load affected performance on conflict and no-conflict problems tested whether there are two kinds of reasoning systems with different processing characteristics. Examining the effect of load on different span groups tested whether there is qualitative variation in the reasoning mechanism itself.

EXPERIMENT

Method

A total of 308 first-year psychology students from the University of Leuven, Belgium, participated in return for credit in a psychology course.

Measure of WM

Participants' WM capacity was measured using a version of the Operation Span task (La Pointe & Engle, 1990) adapted for group testing (Gospan; for details, see De Neys, d'Ydewalle, Schaeken, & Vos, 2002). This task involves performing a series of simple mathematical operations while attempting to remember a list of unrelated words. The measure of WM capacity is based on the number of words remembered.

Reasoning Task

The syllogistic reasoning task was based on Sá, West, and Stanovich (1999). Participants evaluated eight syllogisms taken from the work of Markovits and Nantel (1989). Four of the problems had conclusions in which logic was in conflict with believability (i.e., conflict items). For the other four problems, the believability of the conclusion was consistent with its logical status (i.e., no-conflict items).

The following item format was adopted:

Premises: All fruits can be eaten.

Hamburgers can be eaten.

Conclusion: Hamburgers are fruits.

1. The conclusion follows logically from the premises.

2. The conclusion does not follow logically from the premises.

Type down the number that reflects your decision: _

Instructions, which showed an example item, emphasized that the premises should be assumed to be true and that a conclusion should be accepted only if it followed logically from the premises.

Dot Memory Task

The dot memory task is a classic spatial storage task (e.g., Bethell-Fox & Shepard, 1988; Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001). For the present study, a 3 × 3 matrix filled with three or four dots was presented for 850 ms (see Fig. 1). Participants memorized the pattern and were asked to reproduce it afterward.

In the high-load condition, the matrix was filled with a complex four-dot pattern (i.e., a "two- or three-piece" pattern based on Bethell-Fox & Shepard, 1988, and the work of Verschueren, Schaeken, & d'Ydewalle, 2004). Miyake et al. (2001) established that storage of similar complex dot patterns taps executive resources. In the low-load condition, the pattern consisted of three dots on a horizontal line (i.e., a "one-piece" pattern in Bethell-Fox & Shepard's terms). This pattern should place only a minimal burden on executive resources.

Procedure

Participants were tested in groups of 6 to 20 and were randomly assigned to one of three (no, low, or high) load conditions. They performed the Gospan task first and then completed the syllogistic reasoning task after a short break. In the low- and high-load conditions, a dot pattern was presented for 850 ms before each syllogism was presented. After participants had typed their response ("1" or "2"), an empty matrix was presented, and they had to reproduce the dot pattern. Instructions emphasized that it was crucial for the dot patterns to be reproduced correctly.

Results

Reasoning Task

Participants were split in three span groups (top, middle, and bottom third) based on the distribution of the Gospan-capacity scores.[2] This classification resulted in approximately equal numbers of individuals with low, medium, and high spans in the three load conditions. Reasoning scores (i.e., the number of logically correct responses) on the four conflict and four no-conflict items were entered in a 3 (load, between subjects) × 3 (span, between subjects) × 2 (conflict, within subjects) analysis of variance.

Figure 2 presents an overview of the findings. There were main effects of load, $F(2, 299) = 4.13$, $p_{rep} = .93$, $\eta_p^2 = .03$, and conflict, $F(1, 299) = 194.53$, $p_{rep} > .99$, $\eta_p^2 = .39$, and these two factors also interacted, $F(1, 299) = 7.86$, $p_{rep} = .99$, $\eta_p^2 = .05$. As predicted by the dual-process framework, the executive load

[2]Analyses based on median, quartile, and quintile splits for the span factor yielded similar results.

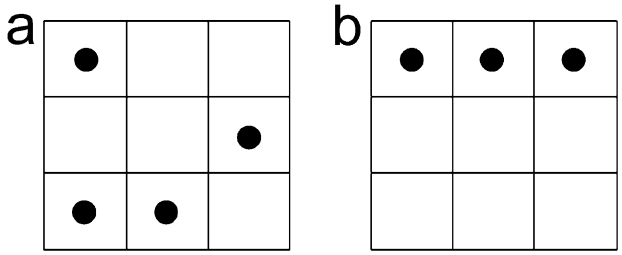

Fig. 1. Examples of the dot patterns in the high-load (a) and low-load (b) conditions.

did not affect reasoning performance on the no-conflict problems, $F(2, 299) < 1$, $p_{rep} = .31$. This result supports the claim that reasoning in the absence of a belief-logic conflict is mediated by an automatically operating system. Also as predicted, the executive burden had a clear impact on performance on the conflict items, $F(2, 299) = 7.64$, $p_{rep} = .99$, $\eta_p^2 = .05$. Performance decreased linearly with increasing secondary-task load, $F(1, 299) = 15.28$, $p_{rep} > .99$, $\eta_p^2 = .05$. This result supports the claim that reasoning in the case of a belief-logic conflict is mediated by a system that requires executive WM resources for its proper functioning.

As in previous individual differences studies, the span and conflict factors interacted, $F(2, 299) = 4.63$, $p_{rep} = .95$, $\eta_p^2 = .03$. Performance on the no-conflict items was uniformly high irrespective of span size, $F < 1$. On the conflict items, a larger executive resource pool did result in better performance, $F(1, 299) = 9.41$, $p_{rep} = .97$, $\eta_p^2 = .03$.

The span factor, however, did not interact with the load factor, $F(4, 299) < 1$, $p_{rep} = .09$ (the Load × Span × Conflict interaction was also nonsignificant, $F < 1$). As Figure 2 indicates, reasoning performance on the conflict items decreased for all span groups under load. This finding is consistent with the weak interpretation of individual differences in reasoning.

One might argue that the performance decrease observed in the present study does not necessarily follow from an impact on the reasoning process per se. Indeed, before participants can start reasoning, they have to read and mentally represent the premises. Such reading and comprehension processes can also demand some executive resources (e.g., Just, Carpenter, & Keller, 1996). However, it is crucial to note that the present executive load did not affect reasoning performance on the no-conflict items. If the load had merely interfered with the comprehension process, it should have affected performance on the no-conflict items, too. The fact that reasoning was affected only on the conflict items even among individuals with low spans indicates that the load specifically interfered with processing the conflict between belief and logic.

Figure 2 also suggests that the impact of the executive load tended to increase with decreasing span size. For the high-span group, performance on the conflict items was 13% lower in the high-load condition compared with the no-load condition, $F(1, 299) = 2.92$, $p_{rep} = .83$, $\eta_p^2 = .01$. The medium-span group showed a 16% decrease, $F(1, 299) = 4.78$, $p_{rep} = .91$, $\eta_p^2 = .02$, and the low-span group showed a 22% decrease, $F(1, 299) = 8.08$, $p_{rep} = .97$, $\eta_p^2 = .03$. Although the differences among the span groups did not reach traditional significance, $F < 1$, the direction of the trend cannot be reconciled with the strong interpretation of the groups' differential performance.

One might still claim, however, that the exclusively heuristic processing alluded to by the strong interpretation is found only among the most ungifted reasoners. In this sense, the "bottom third" selection criterion for the low-span group might have been too liberal. Therefore, the analysis was repeated with more extreme span groups—only participants in the top and bottom quintiles (20%) of the capacity distribution. However, in this analysis, the low-span group showed an even more pronounced performance decrease with a high load (23%), $F(1, 129) = 5.71$, $p_{rep} = .93$, $\eta_p^2 = .04$. Hence, contrary to the strong claim, selection of a more extreme low-span group tended to boost, rather than decrease, the effect of executive load.

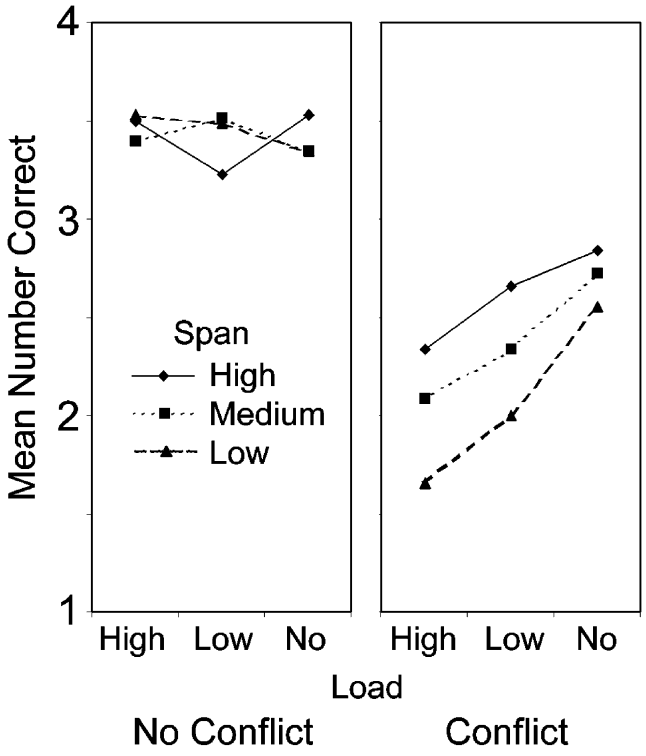

Fig. 2. Reasoning performance of the high-, medium-, and low-span groups as a function of executive load. Results are shown separately for conflict problems, in which the logical validity of the conclusion conflicted with its believability, and no-conflict problems, in which the logical validity and believability of the conclusion were consistent.

Dot Memory Task

Results for the dot memory task indicated that the task was performed properly. The mean number of correctly localized dots was 3.54 ($SD = 0.51$) for the complex four-dot patterns in the high-load condition and 2.84 ($SD = 0.34$) for the simple three-dot patterns in the low-load condition. Thus, overall, about

89% of the four-dot and 94% of the three-dot patterns were reproduced correctly. The mean proportion of correctly localized dots was entered in a 2 (dot complexity, between subjects) × 3 (span, between subjects) × 2 (conflict, within subjects) analysis of variance. Not surprisingly, the simple three-dot patterns were remembered better than the four-dot patterns, $F(1, 212) = 22.56$, $p_{rep} = .99$, $\eta_p^2 = .1$. The dot complexity and span factors also interacted, $F(2, 212) = 4.78$, $p_{rep} = .95$, $\eta_p^2 = .04$. The three span groups remembered the three-dot patterns equally well, $F < 1$, but the complex patterns were somewhat better remembered by the high-span group (93%) than the middle-span (87%) and low-span (87%) groups, $F(2, 111) = 5.99$, $p_{rep} = .97$, $\eta_p^2 = .1$. Other factors did not reach conventional significance, all Fs < 1. Although the somewhat lower performance on the four-dot patterns by the low-span group might have resulted in a lower executive burden for this group than for the other span groups, the low-span group's high proportion of correctly localized dots (87%) indicates that there was no systematic trade-off between reasoning and recall.

DISCUSSION

The present study introduced the use of a secondary task to validate basic processing assumptions in dual-process theories of reasoning. The dual-process framework postulates the existence of a heuristic belief-based reasoning system and an analytic logic-based reasoning system that are characterized by differential involvement of executive WM resources. Consistent with the basic claim, the present findings showed that when the heuristic system triggered the correct response to syllogistic problems, burdening executive resources with a secondary storage task did not affect reasoning performance. This finding supports the claim that the heuristic system operates automatically. Reasoning performance did decrease under experimental load, however, when the belief-based response conflicted with the logically correct response. This finding indicates that the analytic operations alleged to be necessary to override the heuristic response and compute the correct solution draw on executive resources. This study thereby completes previous correlational studies in the dual-process field by showing that erroneous reasoning in the case of a belief-logic conflict is not only associated with, but also directly caused by, limitations in executive resources.

Correlational individual differences studies are not clear about the reason for the established diversity in reasoning performance. The performance differences between reasoners varying in cognitive capacity can be interpreted quantitatively or qualitatively. A weak, quantitative interpretation entails that all capacity groups have the two reasoning systems at their disposal and that all reasoners engage in analytic reasoning in the case of a belief-logic conflict; however, because of the demanding nature of the analytic operations, people with higher spans are simply more likely than those with lower spans to complete the analytic process. A strong, qualitative interpretation entails that individuals with low spans have no access to an analytic system; their reasoning, in contrast with that of people with high spans, is completely mediated by the heuristic system.

The present findings regarding the effect of a secondary task directly contradict the strong interpretation. If the reasoning of individuals with low spans is mediated solely by the heuristic system, they should not be bothered by an additional executive load. Results clearly showed, however, that even for the lowest-capacity groups, performance on the conflict problems decreased under executive load, whereas performance in the absence of a conflict was unaffected. The only difference between the conflict and no-conflict problems was in the logical status of their belief-based conclusions. If a reasoner did not take logical validity into account, there would be nothing special about the conflict problems, and processing them would be as undemanding as processing the no-conflict problems. The effect of load on performance on the conflict problems indicates that the lowest-span groups detected the belief-logic conflict and allocated executive resources to resolve it. This is precisely the hallmark of analytic reasoning. Hence, there is no evidence for strong individual differences in the human reasoning machinery.

The present findings have interesting implications for the debate on human rationality (e.g., Stanovich & West, 2000; Stein, 1996). This rife debate centers around whether the traditional norms (such as standard logic) against which the rationality of people's decisions is measured are valid and whether there are individual differences in the norms people adhere to. The present evidence against qualitative individual differences in processing supports the view that all reasoners adhere to the same normative standard. If people with low spans did not know standard logic or did not attempt to adhere to it, solving the conflict and no-conflict items would not pose differential processing demands. One might claim that individuals with higher spans are quantitatively more rational in the sense that they manage better to adhere to the norm. The present study demonstrates, however, that although individuals with lower spans might be less successful, they at least attempt to adhere to the same norm.

The fact that both individuals with high spans and those with low spans take the logical status of the conclusions into account can also support the validity of standard logic as a normative standard itself. This validity is a moot issue in the rationality debate: It has been questioned why reasoning based on logic would be more rational or "correct" than pure belief-based reasoning (e.g., Todd & Gigerenzer, 2000; Oaksford & Chater, 1998). Stanovich and West (2000) defended the validity of traditional norms on the basis of their individual differences studies. Their rationale rested on the assumption that in a particular situation, more cognitively gifted people will be better at selecting the appropriate norm than less gifted people are. Bluntly put, the argument was that whatever the "smart" people do can be assumed to be right. Therefore, the finding that people with high spans prefer responses that are in line with traditional norms indicates that the norms are valid. This argument has

been criticized (e.g., Sternberg, 2000). Focusing on the processing demands of reasoning allows one to sidestep Stanovich and West's individual differences rationale. Indeed, the present findings stipulate that a different outcome does not necessarily imply a different kind of processing. Likewise, a belief-based response does not necessarily imply mere belief-based reasoning. The best support for the validity of a specific norm is not that people with high spans manage to adhere to it, but rather that all reasoners at least attempt to do so.

Finally, one should bear in mind that as did previous individual differences studies, the present study tested the reasoning performance of adult college students only. The executive capacities of the low-span group clearly fell within the normal range. It can be argued that in more special populations (e.g., patients with cognitive dysfunction or young children), one would find evidence for strong, qualitative differences in reasoning processes. Of course, the present claims apply only to the group of young, educated adults whose reasoning performance has been the subject of dual-process theorizing and the rationality debate. With this stipulation in mind, this study provides clear evidence that all reasoners have access to both an automatic, heuristic reasoning system and an executive, analytic reasoning system. In sum, the reasoning systems differ, but the reasoners do not.

REFERENCES

Bethell-Fox, C.E., & Shepard, R.N. (1988). Mental rotation: Effects of stimulus complexity and familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 12–23.

De Neys, W. (in press). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual task considerations. *Quarterly Journal of Experimental Psychology*.

De Neys, W., d'Ydewalle, G., Schaeken, W., & Vos, G. (2002). A Dutch, computerized, and group administrable adaptation of the operation span test. *Psychologica Belgica*, *42*, 177–190.

Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist*, *49*, 709–724.

Evans, J.St.B.T. (2002). Logic and human reasoning: An assessment of the deduction paradigm. *Psychological Bulletin*, *128*, 978–996.

Evans, J.St.B.T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences*, *7*, 454–459.

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

Gigerenzer, G. (2004). Dread risk, September 11, and fatal traffic accidents. *Psychological Science*, *15*, 286–287.

Gigerenzer, G., & Regier, T. (1996). How do we tell an association from a rule?: Comment on Sloman (1996). *Psychological Bulletin*, *119*, 23–26.

Just, M.A., Carpenter, P.A., & Keller, T.A. (1996). The capacity theory of comprehension: New frontiers of evidence and arguments. *Psychological Review*, *103*, 773–780.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.

Klaczynski, P.A. (2000). Is rationality really "bounded" by information processing constraints? *Behavioral and Brain Sciences*, *23*, 683–684.

Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working memory capacity. *Intelligence*, *14*, 389–433.

La Pointe, L.B., & Engle, R.W. (1990). Simple and complex word spans as measures of working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 1118–1133.

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

Miyake, A., Friedman, N.P., Rettinger, D.A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *Journal of Experimental Psychology: General*, *130*, 621–640.

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

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.

Newton, E.J., & Roberts, M.J. (2003). Individual differences transcend the rationality debate. *Behavioral and Brain Sciences*, *26*, 530–531.

Oakhill, J., Johnson-Laird, P.N., & Garnham, A. (1989). Believability and syllogistic reasoning. *Cognition*, *31*, 117–140.

Oaksford, M., & Chater, N. (1998). *Rationality in an uncertain world: Essays on the cognitive science of human reasoning*. Hove, England: Psychology Press.

Osman, M. (2004). An evaluation of dual-process theories of reasoning. *Psychonomic Bulletin & Review*, *11*, 988–1010.

Reyna, V.F., Lloyd, F.J., & Brainerd, C.J. (2003). Memory, development, and rationality: An integrative theory of judgement and decision-making. In S. 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.

Sloman, S.A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, *119*, 3–22.

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. (2000). Individual differences in reasoning: Implications for the rationality debate [Target article and commentaries]. *Behavioral and Brain Sciences*, *23*, 645–726.

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

Sternberg, R.J. (2000). The ability is not general, and neither are the conclusions. *Behavioral and Brain Sciences*, *23*, 697–698.

Todd, P.M., & Gigerenzer, G. (2000). Precis of simple heuristics that make us smart. *Behavioral and Brain Sciences*, *23*, 727–780.

Verschueren, N., Schaeken, W., & d'Ydewalle, G. (2004). Everyday conditional reasoning with working memory preload. In *Proceedings of the Twenty-Sixth Annual Meeting of the Cognitive Science Society* (pp. 1399–1404). Mahwah, NJ: Erlbaum.

(RECEIVED 4/6/05; REVISION ACCEPTED 8/19/05;
FINAL MATERIALS RECEIVED 9/8/05)