Title
Individual Differences in Rational Thinking Time

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN
1069-7977

Authors
De Neys, Wim
Dieussaert, Kristien

Publication Date
2005

Peer reviewed

Individual Differences in Rational Thinking Time

Wim De Neys (Wim.Deneys@psy.kuleuven.be)
Kristien Dieussaert (Kristien.Dieussaert@psy.kuleuven.be)
Department of Psychology, K.U.Leuven, Tiensestraat 102
B-3000 Leuven, Belgium

Abstract

Individual difference studies suggest that reasoners highest in cognitive capacity favor analytic, normative responses over fallacious, heuristic responses. The present study complemented reasoning accuracy with timing data to obtain an indication of the nature of the reasoning process underlying the response selection. A total of 199 participants were presented with a measure of working memory capacity and a syllogistic reasoning task. As predicted, higher spans were not only more likely to draw a correct conclusion but also reasoned longer when believability and logical status of the conclusion conflicted. Working memory capacity did not predict reasoning accuracy or time when believability and logic were consistent. Findings validate basic processing assumptions of a dual process framework of thinking.

Introduction

In recent years, studies on individual differences in reasoning performance moved to the center stage of the cognitive research on human thinking. Over the last decades an impressive body of research established that human reasoning frequently violates normative standards: In a wide range of reasoning tasks the majority of educated adults fails to give the response that is correct according to logic or probability theory (Evans, 2002; Kahneman, Slovic, & Tversky, 1982). The exact nature and consequences of this "rational thinking failure" instigated a rife debate (e.g., Stein, 1996). Systematic research on individual differences in cognitive capacity (e.g., Kokis, Macpherson, Toplak, West, & Stanovich, 2002; Klaczynski, 2001; Newstead, Handley, Harley, Wright, Farrelly, 2004; Stanovich & West, 1999, 2000; Torrens, Thompson, & Cramer, 1999) pointed to the crucial stipulation that participants highest in cognitive (working memory) capacity do manage to solve the problems correctly.

According to influential dual process theories of thinking, correct normative responding in classic reasoning tasks requires that an analytic, controlled reasoning process overrides prepotent heuristics (e.g., Epstein, 1994; Evans, 2003; Evans & Over, 1996; Sloman, 1996; Stanovich & West, 2000). Advocates of the dual process framework have distinguished two types of reasoning systems. In general, the first, so-called heuristic system is characterized by a tendency towards an automatic contextualization of a problem with prior knowledge and beliefs whereas the second, so-called analytic system tends to decontextualize a problem and allows reasoning according to normative standards. The heuristic default system is assumed to operate fast and automatically whereas the operations of the analytic system would be slow and heavily demanding of our limited working memory resources. Both systems have been shown to reside in different brain regions (Goel & Dolan, 2003).

It is assumed that in most situations the heuristic and analytic system interact in concert. Hence, most of the time the heuristic default system will provide us with fast, frugal, and correct conclusions. However, the prepotent heuristics may also bias reasoning in situations that require more elaborate, analytic processing. That is, both 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 would draw on limited, executive working memory resources. Therefore, correct analytic reasoning in case of a belief-logic conflict would be characteristic of those highest in working memory span: The more resources that are available, the more likely that the analytic system will be successfully engaged and the correct response calculated.

Individual difference studies on the belief bias effect in syllogistic reasoning provide paradigmatic support for the framework. Belief bias refers to the heuristic tendency to judge the validity of a syllogism by evaluating the believability of the conclusion (e.g., Klauer, Musch, & Naumer, 2000; Oakhill, Johnson-Laird, & Garnham, 1989; Newstead, Pollard, Evans, & Allen, 1992). For some syllogistic problems the logical status of the conclusion conflicts with background beliefs (i.e., conflict items, see Appendix for examples). The heuristic, belief-based system thus triggers an erroneous response and consequently many people fail to solve these problems correctly. For other syllogisms the logical status of the conclusion is consistent with the believability of the conclusion (i.e., no-conflict items). Correct solution rates on these no-conflict items are uniformly high.

Consistent with the dual process framework it has been observed (e.g., Kokis et al., 2002; Newstead et al., 2004; Stanovich & West, 1999, 2000) that individual differences in cognitive capacity predict performance on the conflict items but not on the no-conflict items. Indeed, the heuristic system is assumed to operate automatically, that is, it should not burden the limited executive resources. Hence, when the heuristic-belief based response is consistent with the logical response even the heuristically reasoning low spans will get the right answer. However, on the conflict items only the higher spans will manage to block the heuristic system and reason analytically to get the logically correct answer.

Despite the support from individual difference studies, dual process theories have been severely criticized (e.g., Gigerenzer & Regier, 1996; Stanovich & West, 2000). One fundamental critique concerns the fact that the framework has exclusively focused on the accuracy output (i.e., is a response correct or not) and not on the underlying cognitive processes. This may result in dramatic confounds: Even if one gives a correct response this does not imply that one has reasoned analytically. After all, you might have been lucky and guessed the correct answer. Likewise, giving an incorrect response does not imply that you did not reason analytically. Indeed, you might have noticed the belief-logic conflict and actively engaged in analytic processing but simply failed to complete the process.

The present study starts to address the processing shortcoming in individual differences studies by complementing reasoning accuracy data with latency data. The time needed to draw an inference can be used as indicator of the nature of a reasoning process: Dual process theories explicitly assume that the heuristic system operates much faster than the time-consuming analytic system. There is abundant evidence that automatically operating cognitive processes are faster than working memory resources demanding processes (e.g., Cowan, 1995; McElree, 2001; Shiffrin & Schneider, 1977). More specifically, reasoning under time pressure has been shown to result in increased heuristic responding (e.g., Roberts & Newton, 2001; Schroyens, Schaeken, & Handley, 2003). Likewise, De Neys (in press) established that analytic responses required more time when the response time for correct, analytic and incorrect answers was directly compared in a number of reasoning tasks.

In the present study participants were presented with a measure of working memory capacity and a syllogistic reasoning task with conflict and no-conflict problems. The dual process framework assumes that analytic reasoning on conflict items is characteristic of those highest in cognitive capacity. Low spans are expected to reason heuristically whereas high spans are assumed to block the heuristic response and reason analytically. Since the analytic system operates slower than the heuristic one, a higher span should therefore not only be associated with higher accuracy but also with longer response times. On the no-conflict items, the default heuristic computations will not need to be overridden by additional analytic computations. Both low and high spans can rely on the fast and frugal heuristic system to solve the problem. Therefore, everyone should draw relatively fast and correct conclusions. Hence, as with response accuracy, reasoning time should not be associated with working memory capacity on the no-conflict items. Finally, overall, one also predicts that reasoning time will be longer on the conflict than on the no-conflict items.

As many cognitive scientists, Stanovich and West (2000) and other dual process theorists assume that executive working memory resources are the quintessential component of computational cognitive capacity (e.g., Evans, 2003). However, so far, the individual difference studies in the dual process field have always adopted measures of general intelligence (e.g., SAT-scores). Although variations in working memory capacity are well captured by these measures the relation is not perfect (e.g., Engle, Tuholski, Laughlin, & Conway, 1999). Therefore, the present study adopted a test that was specifically designed to measure executive working memory capacity.

Experiment

Method

Participants

A total of 199 first-year psychology students from the University of Leuven, Belgium, participated in return for psychology course credit.

Material

Working memory measure. Participants' working memory capacity was measured using a version of the Operation Span task (Ospan, La Pointe & Engle, 1990) adapted for group testing (Gospan, for details see De Neys, d'Ydewalle, Schaeken, & Vos, 2002). Participants solve series of simple mathematical operations while attempting to remember a list of unrelated words. First, the operation from an operation-word pair is presented on screen (e.g., 'IS (4/2) – 1 = 5 ?'). Participants read the operation silently and press a key to indicate whether the answer is correct or not. Responses and response latencies are recorded. After the participant has typed down the response, the corresponding word (e.g., 'BALL') from the operation-word string is presented for 800 ms. Three sets of each length (from two to six operation-word pairs) are tested and set size varies in the same randomly chosen order for each participant. The Gospan-score is the sum of the recalled words for all sets recalled completely and in correct order.

Participants who make more than 15% math errors or whose mean operation response latencies deviate by more than 2.5 standard deviations of the sample mean are discarded.

Syllogisms. The syllogistic reasoning task was based on Sá, West, and Stanovich (1999). Participants evaluated eight syllogisms (see Appendix) taken from the work of Markovits and Nantel (1989). Four of the problems had conclusions whereby logic was in conflict with believability (i.e., conflict items, two items with an unbelievable-valid conclusion, and two items with a believable-invalid conclusion). For the four "no-conflict" items the believability of the conclusion was consistent with the logical status (i.e., two items with an unbelievable-invalid conclusion, and two items with a believable-valid conclusion).

The experiment was run on computer. Items were presented in the following format:

Premises: All mammals walk.
Whales are mammals.

Conclusion: Whales can walk.

a. The conclusion follows logically from the premises.
b. The conclusion does not follow logically from the premises.

Type down the letter that reflects your decision: _

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

Procedure

Participants were tested in groups of 18 to 47. Participants started with the Gospan task and completed the syllogistic reasoning task after a short break. Participants typed the letter reflecting their decision and pressed the Enter-key when finished. The next item was presented 750 ms after the Enter-key was pressed. The instructions made clear that there were no time limits, but it was stressed that once participants had made their final decision, they had to press the Enter-key immediately. The time between the presentation of the item and pressing the Enter-key was recorded together with the answer. The eight problems were presented in the same, randomly determined order to all participants in order to minimize any measurement error due to a participant by order interaction.

Results

Five participants were discarded because they did not meet the operation correctness or latency requirements of the working memory measure. Mean Gospan-score of the remaining 194 participants was 34.11 ($SD = 10.80$) words recalled correctly out of 60.

For each participant the accuracy score (i.e., the number of logically correct responses out of four) and mean inference latency on the four conflict and four no-conflict items was calculated. To prevent extreme reasoning times from unduly influencing the means, any latency more than 2.5 SD above a person's mean latency was replaced with that cutoff value. This procedure affected approximately 3.5% of all observations.

Table 1 presents the basic findings. Participants erred less on the no-conflict than on the conflict items, $t(193) = 9.99$, $p < .001$. Consistent with the prediction the no-conflict items were also solved faster, $t(193) = 2.09$, $p < .04$[1].

[1] Reported latencies were collapsed over correct and incorrect responses. The latency patterns were identical when only correct responses were analyzed: No-conflict items, $M = 12.46$ s ($SD = 4.01$), were solved faster than conflict items, $M = 13.44$ s ($SD = 4.98$), $t(179) = 2.74$, $p < .01$, and working memory capacity correlated with reasoning times on the conflict items, $r = .16$, $p < .03$, but not on the no-conflict items, $r = .11$, $p = .14$.

Table 1

Mean Reasoning Accuracy, Latency (s), and Correlations with Working Memory (WM) Capacity in Function of Belief-Logic Conflict

	Problem Type			
	Conflict		No-Conflict	
Measure	Accuracy	Latency	Accuracy	Latency
Mean	2.59	13.29 s	3.54	12.71 s
SD	1.26	4.25	.63	4.09
WM-correlation	.24	.20	.01	.11

As Table 1 shows, the correlations between working memory capacity and response accuracy replicated previous findings with general intelligence measures. In the event of a belief-logic conflict, higher spans were more likely to solve the problems correctly, $r = .24$, $p < .005$, but there was no association with working memory capacity for the no-conflict items, $r = .01$, $p = .97$. The correlation of .24 for the conflict items is in the range of the correlations reported by Newstead et al. (2004, i.e., $r = .18$ to $.23$) albeit somewhat lower than the associations observed by Stanovich and West (1999, i.e., $r = .33$ to $.50$).

The crucial finding is that, consistent with dual process predictions, the latency data shows a similar pattern: Higher spans spend more time solving the conflict items, $r = .20$, $p < .005$, but working memory capacity does not significantly predict latencies on the no-conflict items, $r = .11$, $p = .12$.

Table 2

Mean Latency Increase (ms) on Conflict Items (Compared with Correctly Solved No-Conflict Items) in Function of WM-Capacity

	Span group	
Conflict Response	High span	Low span
Conflict incorrect	2245 (7910)	-185 (3312)
Conflict correct	1079 (4815)	883 (4830)

Note. Standard deviations in parentheses.

Table 2 presents the results of two enlightening additional analyses. The analyses focused on the "reasoning time increase" introduced by the belief-logic conflict in function of working memory capacity and the accuracy on the conflict items. A first analysis examined the "reasoning time increase" in case the conflict item was solved erroneously. If participants draw an erroneous conclusion this does not necessarily imply that they reasoned purely heuristically. Especially for the higher spans it is plausible that they did notice the belief-logic conflict, engaged in analytic reasoning but simply failed to inhibit the prepotent heuristic response completely. To test this hypothesis participants' mean inference time for correctly solved no-conflict items was subtracted from their mean inference time for erroneously solved conflict items[2]. Correctly solved syllogisms for the no-conflict items can be assumed to be computed by the heuristic process. If an erroneous response on the conflict items is also completely based on similar belief-based, heuristic reasoning, one would not expect increased latencies. Results showed that overall there was a marginally significant 930 ms (SD = 5985) latency increase for the erroneous conflict items, $t(132) = 1.79$, $p = .076$. However, a correlational analysis indicated that the latency increase depended on working memory capacity, $r = .25$, $n = 133$, $p < .005$. To present a more specific picture the sample was split up in two span groups based on a working memory capacity median split. Consistent with the positive correlation, the mean latency increase for erroneously solved conflict items differed for the two capacity groups, between-subjects test, $t(131) = 2.37$, $p < .02$. As Table 2 shows, when high spans err on a conflict syllogism, they nevertheless reason about 2 s longer than when they correctly solve a no-conflict problem, $t(60) = 2.21$, $p < .03$. Low spans' latencies do not differ in the two cases, $t(71) = -.47$, $p = .64$.

The second analysis examined the "reasoning time increase" for different working memory capacity groups in case the conflict item was solved correctly. One could argue that the few times that lower spans manage to give a correct answer for a conflict item this is not based on analytic reasoning but merely results from a guessing strategy. To test this hypothesis participants' mean inference time for correctly solved no-conflict items was subtracted from the mean inference time for correctly solved conflict items[3]. If an erroneous response on the conflict items would be based on a heuristic guessing strategy one would not expect a latency increase. Results showed that overall there was a 984 ms (SD = 4810) latency increase for the conflict items, $t(179) = 2.74$, $p < .01$. What is crucial is that the increase did not depend on working memory capacity, $r = .07$, $n = 180$, $p = .34$. As Table 2 shows, both the top and bottom span group showed a rather similar latency increase, between-subjects test, $t(178) = .27$, $p = .79$. This contradicts the guessing hypothesis.

[2] Data of participants who solved all conflict items correctly (n = 61) was not included in the analysis.
[3] Data of participants who erred on all conflict items (n = 14) was not included in the analysis.

Discussion

The present study examined the impact of individual differences in working memory capacity on both reasoning accuracy and time. Given that the analytic system operates slower than the heuristic system, inferences that are assumed to be based on analytic reasoning should be slower than heuristic inferences. Thereby latency data provides straightforward evidence concerning the nature of a reasoning process. Results showed that in case of a belief-logic conflict high spans were indeed not only more likely to draw a correct conclusion but also reasoned longer. This supports the claim that a time and resource demanding process underlies correct reasoning in case of a belief-logic conflict. When believability and logical status of the conclusion were consistent, syllogisms were overall solved faster and more accurately. As with the accuracy data, working memory capacity did not predict reasoning time on the no-conflict items: High and low spans reasoned equally fast and well. These findings support the idea that a fast and undemanding process underlies reasoning when beliefs are not contradicted by logic.

One might claim that an apparent caveat of this study is that more cognitively gifted reasoners might also be generally more cautious. High spans could take more time, for example, because they read the problems more carefully than low spans. Thus, the longer latencies would not be associated with a slower operating reasoning process *per se*. However, such a "general cautiousness" should also show-up during the processing of the no-conflict items. The crucial point is precisely that the longer latencies (and higher accuracy) for high spans are specifically tied to the presence of a belief-logic conflict. This pattern is a priori predicted by the dual process framework.

The present findings stressed the link between working memory capacity and analytic reasoning. However, this does not imply that a large resource pool is all there is to analytic reasoning. The results clearly indicated that the relation between working memory capacity and reasoning performance is not perfect. Some low spans do manage to solve conflict items and they show the same latency increase as the high spans when they do so. Clearly, factors outside the cognitive ability spectrum will also affect performance (e.g., "epistemic thinking dispositions", see Stanovich & West, 1999). Thus, in pointing out the contribution of executive working memory resources for correct, analytic reasoning the present study does not minimize the role of other mediating factors.

The study also illustrated possible limitations of an exclusive focus on accuracy data in dual process studies. Contrary to low spans, high spans who failed to solve a conflict item nevertheless spent more time drawing the erroneous conclusion than they spent drawing a correct conclusion in the absence of a belief-logic conflict. Hence, despite the belief-dominated response, high spans' reasoning was affected by the conflict. This suggests that whereas an erroneous conflict response might be equated with mere heuristic reasoning for the low spans, an erroneous response for the high spans might be better characterized as a failure to complete an analytic reasoning process. In this case, individual difference studies that exclusively focus on reasoning accuracy will underestimate the link between working memory capacity and analytic reasoning processes: For the high spans even an erroneous response might have involved analytic reasoning.

Given that the present study only started complementing accuracy and latency data it is evident that the findings will need further refinement. It is intriguing, for example, that high spans' latency increase was more pronounced on the incorrect (2245 ms) than on the correct (1079 ms, see Table 2) conflict items. Thus, high spans actually needed about 1 s longer for an incorrect vs. correct conflict response. One possible explanation is that for high spans the incorrect conflict response does not result from an incomplete, unfinished analytic reasoning process but rather from additional heuristic considerations that follow a completed analytic process. Epstein (1994; see also Sloman, 1996) has noted that participants sometimes report they picked an erroneous responses although they clearly knew it was normatively inadequate. Hence, these participants did calculate the correct response but seemed to discard it afterwards. Such an "a posteriori" active discarding would be one explanation for the longer latencies for incorrect conflict responses. The present findings then indicate that this type of reasoning would be characteristic of higher span groups. In this respect, a combination of latency research with converging research methods (e.g., thinking-aloud studies) seems especially promising for the further development of the dual process framework.

Acknowledgements

Preparation of this manuscript was supported by a grant from the Fund for Scientific Research-Flanders (FWO).

References

Cowan. N. (1995). *Attention and memory: An integrated framework*. New York: Oxford University Press.

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.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General, 128*, 309-331.

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, UK: Psychology Press.

Goel, V., & Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition, 87*, B11-B22.

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

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

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

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

Kokis, J. V., Macpherson, R., Toplak, M. E., West, R. F., & Stanovich, K. E. (2002). Heuristic and analytic processing: Age trends and associations with cognitive ability and cognitive styles. *Journal of Experimental Child Psychology, 83*, 26-52.

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.

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 817-835.

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.

Newstead, S. E., Pollard, P., Evans, J. St. B. T., & Allen, J. L. (1992). The source of belief bias effects in syllogistic reasoning. *Cognition, 45*, 257-284.

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

Roberts, M. J., & Newton, E. J. (2001). Inspection times, the change task, and the rapid-response task. *Quarterly Journal of Experimental Psychology, 54A*, 1031-1048.

Schroyens, W., Schaeken, W., & Handley, S. (2003). In search of counter examples: Deductive rationality in human reasoning. *Quarterly Journal of Experimental Psychology*, *56A*, 1129-1145.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing : II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127-190.

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

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.

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. *Behavioral and Brain Sciences, 23*, 645-726.

Stanovich, K. E., & West, R. F. (2003). The rationality debate as a progressive research program. *Behavioral and Brain Sciences, 26*, 531-534.

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

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.

Appendix

Conflict items:

All things that have a motor need oil.
Automobiles need oil.
Automobiles have motors. (Believable–Invalid)

All unemployed people are poor.
David Beckham is not unemployed.
David Beckham is not poor. (Believable–Invalid)

All mammals can walk.
Whales are mammals.
Whales can walk. (Unbelievable–Valid)

All animals like water.
Cats do not like water.
Cats are not animals. (Unbelievable–Valid)

No-conflict items:

All birds have feathers.
Robins are birds.
Robins have feathers. (Believable–Valid)

All cows have four legs.
Snakes do not have four legs.
Snakes are not cows. (Believable–Valid)

All guns are dangerous.
Swords are dangerous.
Swords are guns. (Unbelievable–Invalid)

All things made of wood can be used as fuel.
Gasoline is not made of wood.
Gasoline cannot be used as fuel. (Unbelievable–Invalid