

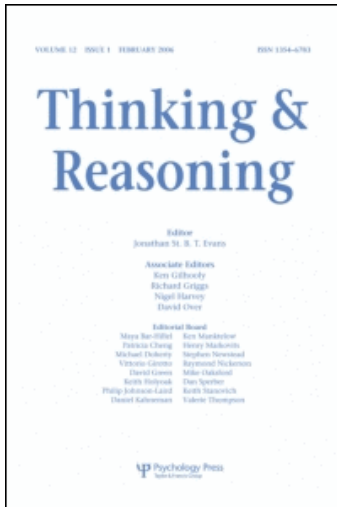
This article was downloaded by: [Franssens, Samuel]

On: 7 May 2009

Access details: Access Details: [subscription number 910970385]

Publisher Psychology Press

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Thinking & Reasoning

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title-content=t713685607>

The effortless nature of conflict detection during thinking

Samuel Franssens ^a; Wim De Neys ^a

^a University of Leuven, Belgium

First Published: May 2009

To cite this Article Franssens, Samuel and De Neys, Wim (2009) 'The effortless nature of conflict detection during thinking', *Thinking & Reasoning*, 15:2, 105 — 128

To link to this Article: DOI: 10.1080/13546780802711185

URL: <http://dx.doi.org/10.1080/13546780802711185>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

The effortless nature of conflict detection during thinking

Samuel Franssens and Wim De Neys
University of Leuven, Belgium

Dual process theories conceive human thinking as an interplay between heuristic processes that operate automatically and analytic processes that demand cognitive effort. The interaction between these two types of processes is poorly understood. De Neys and Glumicic (2008) recently found that most of the time heuristic processes are successfully monitored. This monitoring, however, would not demand as many cognitive resources as the analytic thinking that is needed to solve reasoning problems. In the present study we tested the crucial assumption about the effortless nature of the monitoring process directly. Participants solved base-rate neglect problems in which heuristic and analytic processes cued a conflicting response or not. Half of the participants reasoned under a secondary task load. A surprise recall task was used as an implicit measure of whether the participants detected the conflict in the problems. Results showed that, even under load, base-rate recall performance was better for conflict problems than for no-conflict problems. Although participants made more reasoning errors under load, recall of the conflict problems was not affected by the working memory load. These findings support the claim about the successful and undemanding nature of the conflict detection process during thinking.

Keywords: Reasoning; Decision-making; Conflict monitoring; Cognitive control; Dual-task.

Human judgements are often based on intuition and prior beliefs rather than on a logical reasoning process. Sometimes this leads to bad decisions. A well-known reasoning error is the base-rate fallacy. The following

Correspondence should be addressed to Samuel Franssens, Lab Experimental Psychology, University of Leuven, Tiensestraat 102, 3000 Leuven, Belgium.
E-mail: samuel.franssens@gmail.com

© 2009 Psychology Press, an imprint of the Taylor & Francis Group, an Informa business
<http://www.psypress.com/tar> DOI: 10.1080/13546780802711185

example is an adaptation of the classic base-rate neglect problem (Kahneman & Tversky, 1973):

In a study 1000 people were tested. Among the participants there were 5 sixteen-year-olds and 995 forty-year-olds. Els is a randomly chosen participant of the study.

Els likes to listen to techno and electronic music. She often wears tight sweaters and jeans. She loves to dance and has a small nose piercing.

What is most likely?

1. *Els is 16 years old.*
2. *Els is 40 years old.*

In this problem people are presented with two types of knowledge concerning a hypothetical person. First they get information about the number of members of two social groups in a sample. Then they read a description that clearly matches a stereotypical member of the smallest group. When asked which social group the person most likely belongs to, people tend to choose the group that is cued by the description. However, because the person was chosen randomly, the normative response should be the biggest group in the sample. People underuse or ignore prior probabilities when making their judgement. This reasoning flaw is called the base-rate fallacy and has been demonstrated in numerous studies (Tversky & Kahneman, 1974).

Dual process theories of thinking explain this tendency to rely on descriptive information by assuming two kinds of reasoning processes (Epstein, 1994; Evans, 2006; Kahneman, 2002; Sloman, 1996; Stanovich, 1999). A general distinction is made between *Type 1* or heuristic processes that are fast and automatic, and *Type 2* or analytic processes that are slow and controlled. Type 1 processes have high capacity, do not require much effort, and can operate in parallel. Type 2 processes, however, have limited capacity, require much cognitive effort, and operate sequentially (Evans, in press). Type 2 processes are related to individual differences in working memory and general intelligence, but Type 1 processes are not (Kokis, MacPherson, Toplak, West, & Stanovich, 2002). Faced with certain problems, the two types of processes will cue the same response. Heuristic processes then quickly deliver a correct answer, and laborious analytic reasoning becomes superfluous. However, in cases like the above example, the heuristic and the analytic response will be different: Heuristic processes will claim Els is 16 years old, but analytic processes will declare she is 40 years old. Choosing the normative answer then requires analytic processes to override the heuristic response (Stanovich, in press; Stanovich & West, 2000).

Clearly, sound reasoning and successfully overriding the heuristic response require that the output of the two processes is monitored

continuously. Conflict monitoring is a key component of any dual process theory of thinking, but the process is poorly understood. De Neys and Glumicic (2008) recently started examining the process and suggested that the monitoring is quite flawless. They argued that people would have little difficulty in detecting that heuristic responses conflict with analytic considerations. In the De Neys and Glumicic study participants solved a set of classic base-rate neglect problems (Kahneman & Tversky, 1973) while thinking aloud. Verbal reports indicated that the participants almost never referred to the base-rates in the tasks. Although this finding seemed to suggest that they did not detect a conflict between base-rates and descriptions, De Neys and Glumicic noted that it is possible the detection of a conflict is not easily verbally expressed. They reasoned that, even though participants almost never report experiencing a struggle between the two options, it is possible that conflict monitoring is successful on a more implicit level. To capture such implicit detection, participants were also presented with a surprise recall task that tested the recollection of the base-rates after they had finished solving all the base-rate problems. There were more correct recall answers on questions about classic base-rate neglect problems (the description refers to a member of the smallest group) compared with questions about control problems in which there was no conflict (a member of the biggest group is described). This suggests that the participants somehow noticed the conflicting nature of these items, which prompted them to further scrutinise them, yielding a better memory for the group sizes. Even the participants who failed to correctly solve any of the classic base-rate neglect problems displayed this enhanced memory for base-rates of classic problems. Other implicit measures such as problem-processing times and visual inspection times also indicated that these items were inspected more thoroughly.

De Neys and Glumicic (2008) therefore characterised conflict monitoring as a successful implicit process. In the present study we further explore the characteristics of the monitoring process by focusing on its processing requirements. Given the apparent omnipresence of successful monitoring, De Neys and Glumicic assumed that the process would be undemanding. This assumption helped to explain why even the least gifted reasoners showed almost flawless performance on the conflict-monitoring measure: People's limited cognitive resources might not always suffice to complete the reasoning process, but they would be sufficient for the less demanding monitoring. However, the crucial hypothesis that the monitoring process does not require much cognitive resources is only a post-hoc argumentation. The present research attempts to test this hypothesis directly. Dual process theories state that analytic thinking is slow and resource demanding. A cognitive load that burdens central processing resources should therefore lead to a performance drop on reasoning problems that require analytic

thinking. However, if the monitoring process itself indeed demands only a minimal amount of resources, conflict detection should be largely unaffected by such a load.

In the present study participants solved three types of base-rate neglect problems. In the *incongruent* problems the hypothetical person was portrayed as a stereotypical member of the smallest social group in the sample (i.e., base-rates and description conflicted). In the *congruent* problems the base-rates and the description pointed towards the same answer. These congruent problems could be answered correctly by taking the heuristic route. In other problems the description was not stereotypical (*neutral* problems). Because there were no heuristic cues in these items, we expected people to reason analytically and focus on the base-rates. A high number of correct responses was expected for the congruent and the neutral items. For the incongruent problems, however, the prediction from the literature (Kahneman & Tversky, 1973) was that people would follow their intuition and err in the majority of cases.

Half of the participants solved the base-rate neglect problems under a secondary task load. Dual process theories assume that correct answers on incongruent and neutral items depend on the activation of resource-demanding analytic processes. Congruent items, however, can be solved by heuristic processes, which operate automatically. Therefore the expectation was that the secondary task would lead to a drop in reasoning performance for incongruent and neutral items, but not for congruent items.

The second part of the experiment was a surprise recall task. The participants had to answer questions about each of the problems they had just solved. There were easy questions about the descriptions and a crucial question about the base-rates in each problem. As in De Neys and Glumicic's (2008) study, successful conflict detection in the judgement task should result in a better memory for base-rates of items in which there was a conflict (incongruent items), compared with items in which there was no conflict (congruent items), because of a deeper processing of these items.

The aim of the load manipulation was to limit the cognitive resources the participants could allocate to processing information from the reasoning problems. Therefore an overall poorer recall was expected for participants who performed the judgement task under load. However, if the hypothesis about the effortless nature of the conflict-monitoring process is correct, participants who made their judgements under load should still have managed to detect the conflicts in the reasoning problems. This should have resulted in a better processing of base-rates for items in which there was a conflict. Consequently, recall should be better for incongruent items than for congruent items, whether or not the participants reasoned under load.

METHOD

Participants

A total of 74 second-year psychology students at the University of Leuven (Belgium) participated in return for course credit.

Pilot study

We created 53 stereotypical and neutral personality descriptions. Eight second-year psychology students (who did not participate in the main study) then judged how likely it was that the described person belonged to each of two social groups. The raters used an 11-point scale ranging from 0, extremely unlikely, to 10, extremely likely. After they had finished rating all 53 items, the students had to answer two questions about the descriptions in these items. These were multiple-choice questions with four options. From this we selected two sets of four stereotypical items (as explained below, both sets functioned as congruent items as well as incongruent items in the judgement task) and one set of four neutral items. The descriptions in the stereotypical items moderately but consistently cued one of two groups; the descriptions in the neutral items were rated as equally likely to refer to each group. Mean ratings for the descriptions in the two sets of stereotypical items were 7.96 ($SD = 0.64$) and 8.03 ($SD = 0.41$) for the most likely group and 3 ($SD = 0.61$) and 3.21 ($SD = 1.06$) for the least likely group. For the neutral items the mean ratings were 6.71 ($SD = 0.85$) for the most likely group and 6.06 ($SD = 0.96$) for the least likely group. Performance on the recall task was the same for both sets of stereotypical items, with an average of 6.83 out of 8 questions correct for each set (2 questions for each of 4 items). Recall was a bit lower for the neutral items (5.83 out of 8 questions correct). This result was expected because the descriptions in the neutral items by definition did not fit any stereotype, which made them less salient and harder to encode.

Materials

In the judgement task four stereotypical items served as incongruent problems. The other four served as the congruent items. However, a second version of the task was created, in which the base-rates of all items were switched. The incongruent items in the first version now became the congruent items, and vice versa. In the actual experiment half of the participants solved the first version and the other half solved the second version. Results for each item type were averaged over the two versions. Together with the pilot study, this extra control ruled out the possibility that differences in performance between incongruent and

congruent items could be explained in terms of differences in item stereotypicality.

The different problems were presented with slightly varied base-rates. More precisely, for each item type two problems were presented with a 995/5, one with a 994/6, and one with a 996/4 base-rate ratio. This variation was included to make the subsequent recall task more engaging. De Neys and Glumicic (2008) have already shown that there is no difference in performance for the three base-rate levels.

Note that we specifically opted to use base-rates that were extreme and descriptions that were moderate. This was done to ensure that the items would evoke a conflict between heuristic and analytic processes. Items with moderate base-rates or extreme descriptions would leave room for discussion about which answer is correct, if the participants adopt a Bayesian approach to solve the items (Gigerenzer, Hell, & Blank, 1988). In order to calculate the probability that the person described belongs to one or the other social group, information in the description will then be combined with the information from the base-rates. This leads to difficulties when the base-rates are too moderate (e.g., when a sample consists of 30 men and 70 women, and the person described is a football fan. Saying that the person is a man cannot unequivocally be considered as wrong) or the descriptions are too extreme (e.g., when a sample consists of 995 women and 5 men, and the person described is the leader of the Catholic Church, i.e., the Pope. No matter what the base-rates are, the person can never be a woman). The combination of extreme base-rates with moderately stereotypical descriptions guaranteed that the response cued by the base-rates was the normatively correct one. Note that we do not wish to associate heuristic processes with wrong answers and analytic processes with right answers. Evans (2006, 2007a, 2007b) has recently argued that analytic processes can be responsible for cognitive biases as much as heuristic processes. We acknowledge that, to elicit a conflict between heuristic and analytic processes, it was not strictly necessary to include a correct answer in the items. However, without the extreme base-rates and the moderate descriptions, it was possible that both types of processes pointed towards the same answer if the participants made their decisions in the aforementioned Bayesian manner. In that case there would not be a conflict between heuristic and analytic processes. Previous work with similar items (De Neys & Glumicic, 2008; De Neys, Vartanian, & Goel, 2008) showed that the extreme base-rates do not boost reasoning performance. As in the classic study of Kahneman and Tversky (1973), people select the heuristic response on the vast majority of the conflict problems.

The order of the two response options (1 and 2) was counterbalanced. For half of the problems the correct response (i.e., the response consistent with the base-rates) was option 1, for the other half the second option

(2) was the correct one. The problems were based on a variety of stereotypes related to gender, race, age, and lifestyle. The following are examples of the three problem types (an overview of all items can be found in Appendix 1):

Incongruent:

In a study 1000 people were tested. Among the participants there were 994 Swedes and 6 Italians. Mario is a randomly chosen participant of the study.

Mario is 25 years old. He is a charming young man and is a real womaniser. His favourite dish is the spaghetti his mother makes.

What is most likely?

1. *Mario is a Swede.*
2. *Mario is an Italian.*

Congruent:

In a study 1000 people were tested. Among the participants there were 5 women and 995 men. Dominique is a randomly chosen participant of the study.

Dominique is 32 years old and is a confident and competitive person. Dominique's goal is building a career. Dominique does a lot of sport and is well muscled.

What is most likely?

1. *Dominique is a woman.*
2. *Dominique is a man.*

Neutral:

In a study 1000 people were tested. Among the participants there were 994 people from Antwerp and 6 people from Amsterdam. Bart is a randomly chosen participant of the study.

Bart is sixteen years old and still goes to school. He weighs 80 kilos and has a little sister of 14 years old and an older brother who has already been attending university for two years.

What is most likely?

1. *Bart is from Antwerp.*
2. *Bart is from Amsterdam.*

Procedure

The experiment was run on a computer and started with a welcoming screen. The participants received information about the content of the reasoning problems, but any hints on how to solve the problems were avoided:

In a big research project a number of studies were carried out where short personality descriptions of the participants were made. In every study there were participants from two population groups (e.g., carpenters and policemen). In each study one participant was drawn at random from the sample. You'll get to see the personality description of this randomly chosen participant. You'll also get information about the composition of the population groups tested in the

study in question. You'll be asked to indicate to which population group the participant most likely belongs.

Judgement task. The 12 base-rate neglect problems were presented in random order. In each problem the participants first saw a screen with the following introductory information: *In a study 1000 people were tested. Els [or any of the 12 hypothetical persons] is a randomly chosen participant of the study.*

After reading this they had to press the space bar to go to the next screen, on which the first base-rate neglect problem appeared:

In the study there were 5 sixteen-year-olds and 995 forty-year olds.

Els likes to listen to techno and electronical music. She often wears tight sweaters and jeans. She loves to dance and has a small nose piercing.

What is most likely?

1. *Els is 16 years old.*
2. *Els is 40 years old.*

Judgements had to be made by pressing the “1” or “2” key. The participants then proceeded with the next problem. After they had finished solving all 12 problems the recall task was introduced.

In the *load condition* the participants had to make their judgements while performing a visuospatial secondary task (De Neys, 2006a; 2006b; De Neys & Verschueren, 2006; Verschueren, Schaeken & d'Ydewalle, 2004). After the screen with the introductory information about the base-rate neglect problem, a pattern of four dots in a three-by-three grid (see Figure 1 for an example) was presented for 900 milliseconds. The participants were told that they had to reproduce this pattern after solving the base-rate neglect problem that followed. Upon completion of the base-rate neglect problem, they saw an empty grid that they could fill up with dots by clicking in the grid. It was emphasised that it was important to get the pattern right every time. Previous research has established that this dot memory task efficiently taps executive processing resources (Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001).

Recall task. The questions were presented in a fixed order, with questions about the same problem grouped together. Two questions about the descriptions were followed by two questions concerning the sizes of both groups. The questions about the descriptions were included to make the task less repetitive and more engaging; however, our main interest was participants' recollection of the base-rates. Details about the problem at hand (name of the person, social groups involved) remained visible on the screen: *Among the problems you just solved there was one about Els whose*

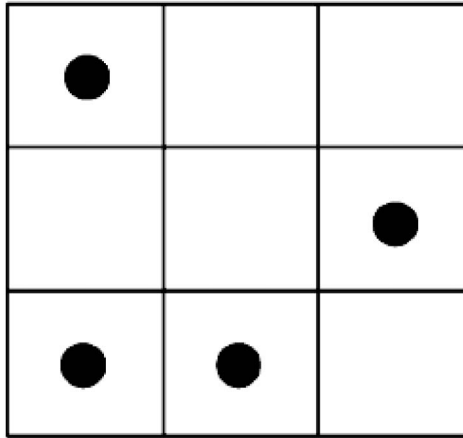


Figure 1. Example of a dot pattern in the visuo-spatial task.

description was randomly chosen from a study in which there were sixteen-year-olds and forty-year-olds.

When questioned about the description, the participants saw four statements about the person and were asked to indicate which one was correct (an overview of all questions can be found in Appendix 2):

Which of the following statements is correct?

1. *Els likes to listen to classical music.*
2. *Els likes to listen to Studio Brussel.* [a popular radio station in Belgium]
3. *Els likes to listen to techno music.*
4. *Els drinks two litres of water every day.*

On the next screen, the participants had to type in the sizes of the two social groups.

Exactly how many sixteen-year-olds were there in the study?

Answer:

Exactly how many forty-year-olds were there in the study?

Answer:

RESULTS AND DISCUSSION

Load task

On average the participants filled in 3.48 ($SD = 0.38$) dots out of 4 correctly, which indicates that the secondary task was properly performed. Scores on

the load task did not correlate with the judgement task ($r = .103, p = .58$) nor with the recall task ($r = -.186, p = .31$), which means there was no trade-off between the load task and judgement or recall.

Judgement task: Accuracy

As in the classic studies of Tversky and Kahneman (1974), the participants selected the heuristic answer on the majority of the incongruent items. However, accuracy was high on congruent and neutral problems. The difference in accuracy between the problem types was significant, $F(2, 84) = 58.90, p < .001, \eta_p^2 = .58$. As Figure 2 shows, the load manipulation only affected performance on items that required analytic reasoning (incongruent and neutral items), $F(1, 72) = 4.18, p = .04, \eta_p^2 = .05$. When heuristic reasoning sufficed (congruent items), there was no effect of load, $F(1, 72) < 1$. This result supports the crucial assumption from dual process theories that analytic processes demand cognitive resources, but heuristic processes operate automatically.

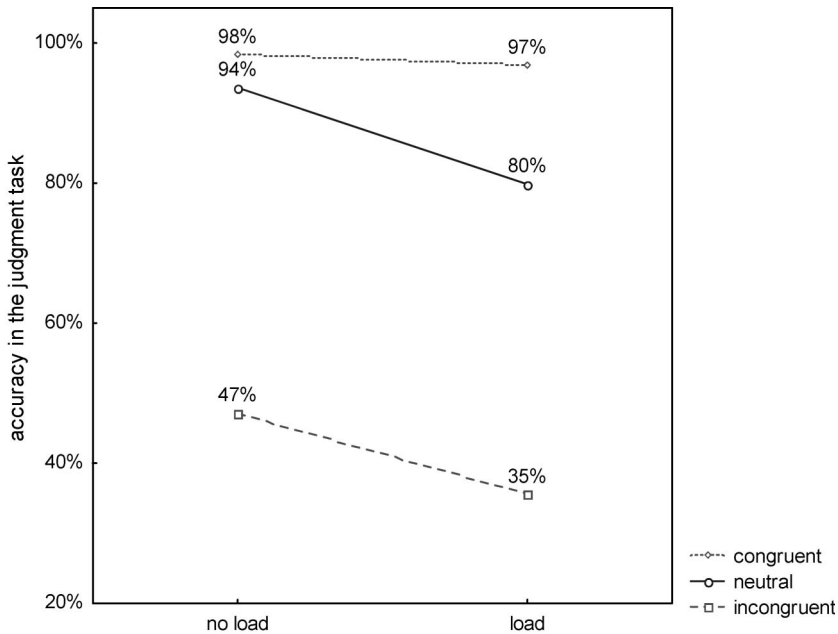


Figure 2. Mean proportion of correctly solved problems in the judgement task.

Judgement task: Response times

There was a main effect of item type on response time, $F(2, 84) = 11.98$, $p < .001$, $\eta_p^2 = .22$ (see Table 1). More specifically, items that required analytic thinking (incongruent and neutral items), took longer to solve than items for which heuristic thinking (congruent items) sufficed, $F(1, 42) = 30.77$, $p < .001$, $\eta_p^2 = .42$. This result is consistent with the dual process theory assumption that analytic processing is more time consuming than heuristic processing. There was no significant effect of load on incongruent and neutral items nor on congruent items, all $F_s(1, 72) < 1$.

Recall task

Accuracy on the two questions about the descriptions was aggregated. Answers on questions about the base-rates were coded as correct when the participants recalled which one of the two groups was the largest (i.e., when the order of magnitude of the base-rates was correctly recalled).¹ Accuracy on the two questions about the base-rates was also aggregated.

Figure 3 presents an overview of the results. First we focus on the data from the participants in the no-load condition. The effect of item type on the recall of the descriptions was significant, $F(2, 84) = 11.22$, $p < .001$, $\eta_p^2 = .21$. Recall of descriptions was lower for neutral items than for incongruent and congruent items, $F(1, 42) = 18.66$, $p < .001$, $\eta_p^2 = .30$. This was expected, because the descriptions in these neutral items were less salient than the descriptions in the other items. There was no difference in recall performance for incongruent and congruent items, $F(1, 42) < 1$.

Our main interest was the performance on questions about the base-rates. The effect of item type was significant, $F(2, 84) = 5.29$, $p = .006$, $\eta_p^2 = .11$.

TABLE 1
Mean response times in seconds (*SD*)

	<i>No-load</i>	<i>Load</i>
Congruent	13.8 (3.5)	15.2 (5.3)
Neutral	16.5 (4.2)	17.1 (4.2)
Incongruent	16.5 (4.8)	17.4 (6.5)

¹If a number higher than 500 was given for the largest population group in the problem and a number lower than 500 for the smallest group, the answer was coded as correct. On the vast majority of trials (95%+) people answered with the base-rates actually presented in the task (i.e., 995/5, 994/6, 996/4). Errors consisted of switching the base-rates around.

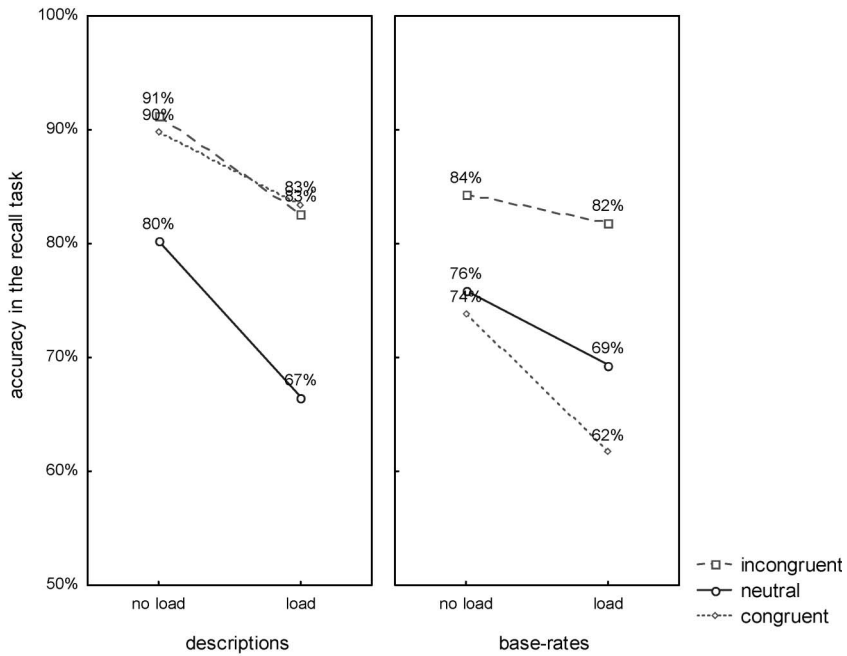


Figure 3. Mean proportion of correctly answered questions in the recall task.

As Figure 3 shows, recall performance was higher for incongruent items than for congruent items, while neutral items fell in between. The crucial difference between the incongruent and congruent items was significant, $F(1, 42) = 7.72$, $p = .008$, $\eta_p^2 = .15$. This means that incongruent items were processed better than congruent items. The only difference between incongruent items and congruent items was the presence of a conflict between base-rates and description. Therefore, the participants must have detected this conflict, which urged them to further inspect these items, and which in turn resulted in a better recall of these items. This finding is in line with the results from De Neys and Glumicic (2008).

The data from the load condition show that the overall recall of descriptions was lower than in the no-load condition, $F(1, 72) = 9.6$, $p = .002$, $\eta_p^2 = .11$, but we find the same pattern of results for the different item types. Again, recall was lower for neutral items than for incongruent and congruent items, $F(1, 30) = 28.65$, $p < .001$, $\eta_p^2 = .48$, and there was no difference between incongruent and congruent items, $F(1, 30) < 1$.

The crucial question in this experiment was whether the load manipulation could prevent the participants from successfully detecting the conflicts in the judgement task. If conflict detection requires many cognitive resources

it should no longer be successful under load, and base-rate recall performance should no longer differ for items in which there was a conflict and items in which there was no conflict. However, if conflict detection operates automatically then it should also be successful under load, and recall of conflict items should be better than recall of no-conflict items. As Figure 3 shows, even under load, base-rates of incongruent items were better recalled than base-rates of congruent items, $F(1, 30) = 9.94$, $p = .003$, $\eta_p^2 = .24$. Recall of neutral items fell in between that of incongruent and congruent items. Furthermore, recall of base-rates of incongruent items did not decrease under load, $F(1, 72) < 1$. However, recall performance for congruent items was lower in the load condition than in the no-load condition, $F(1, 72) = 4.19$, $p = .04$, $\eta_p^2 = .05$. Recall for neutral items was not significantly affected $F(1, 72) = 1.20$, $p = .27$, $\eta_p^2 = .01$. The load manipulation thus resulted in a lower performance on items in which there was no conflict, but when there was a conflict (or when the description did not cue a response), recall performance was unaffected. This finding is further evidence for the hypothesis that conflict detection is an effortless process. Although analytic reasoning performance deteriorated because of the load manipulation, the conflict-monitoring process itself was unaffected by this load.

We argued that the absence of a load effect on our measure of conflict detection, base-rate recall performance, indicated that the monitoring process operates rather automatically. In theory, this lack of a load effect on recall performance could, however, also be attributed to a trade-off between processing time and conflict detection accuracy. Indeed, even if conflict detection were demanding, people might still be successful at it under load by taking additional time to complete the monitoring process. Because people would take more time to make a judgement, recall might also benefit from this extra processing time and remain unaffected. Hence, successful detection and better recall would be bought at the cost of extra processing time. However, such a trade-off implies, that an increase in response times under load should be substantially larger for incongruent problems than for congruent items. The latencies (see Table 1) clearly indicate that this was not the case. The increase in response times was not significant under load and even tended to be somewhat more pronounced for congruent than for incongruent items. This alternative interpretation in terms of a trade-off between secondary task and processing times can therefore be discarded.

GENERAL DISCUSSION

The present study tested the hypothesis that the conflict-monitoring process uses only a minimal amount of cognitive resources and operates quite automatically. In our study half of the participants had to reason under a

load that burdened their cognitive resources. Afterwards they were questioned about the base-rates in these problems. Results from the judgement task showed that the load task affected performance on problems that required analytic reasoning, but it did not affect performance on problems for which heuristic reasoning sufficed. This supports the key assumption from dual process theories that analytic thinking is resource demanding, whereas heuristic thinking is not. Results from the recall task showed that participants' memory was better for base-rates of problems in which there was a conflict, than for problems in which there was no conflict. This result replicates the finding from De Neys and Glumicic (2008) concerning the successful nature of the conflict-monitoring process. Our main interest, however, was the recall performance from the participants who had to make their judgements under load. If the conflict-monitoring process does indeed require only a minimal amount of resources, it should not be affected by a cognitive load. This is exactly what we observed. In the load condition base-rate recall was still better for conflict problems than for no-conflict problems. Furthermore, recall performance on conflict items was equally high in the load condition as in the no-load condition. Even though the load task impeded the analytic thinking that was needed to solve the reasoning problems in the judgement task, it did not prevent the participants from detecting the conflict in these reasoning problems. This supports the hypothesis that conflict monitoring during judgement is an effortless process.

The finding of conflict detection as a successful and effortless process contributes to a much-needed specification of the processing characteristics of dual process theories. De Neys and Glumicic (2008) noted that conflict detection as a successful process implies that, whenever people are confronted with a reasoning problem, heuristic processes are accompanied by at least some minimal analytic activation. Other authors defend this same position (Evans, 2006, 2007b; Kahneman & Frederick, 2002; Stanovich, in press). The idea of an effortless monitoring process can also be found in other researchers' work. Thompson (in press) recently suggested that heuristic responses to reasoning problems are accompanied by a certain *feeling of rightness*, an intuition that the answer is correct. The strength of this feeling of rightness determines whether or not analytic processes rethink the heuristic response. A low feeling of rightness triggers analytic intervention. This process might be linked to conflict monitoring. Thompson's characterisation of this process also seems to imply a non-demanding, automatic nature. Evans (in press) recently made an interesting distinction between two kinds of Type 1 processes. *Preattentive* processes, on the one hand, provide working memory with content and thus determine what information analytic processes are working with. *Autonomous* processes, on the other hand, bypass working memory and "can control

behaviour directly without need for any kind of controlled attention". Evans argued that the conflict between autonomous Type 1 processes and Type 2 processes is settled by preconscious *Type 3* processes. Analytic processes cannot both monitor the output of heuristic processes and at the same time also engage in demanding computations to work out their own response. In this sense dual process theories have traditionally overburdened analytic processes with the responsibility of monitoring for conflict. The recent views of De Neys and Glumicic, Evans, and Thompson seem to share the basic assumption that processes that monitor for the need of analytic intervention operate in a rather automatic way. The crucial contribution of the present paper is that it supports this assumption with empirical findings.

In this study we presented some of the first direct evidence for the automatic nature of the monitoring process. Caution is needed, however, when interpreting the present results. For example, the results come from a single experiment with one specific judgement task (base-rate neglect problems with extreme base-rates and moderate descriptions). It is clear that the findings need replication with other tasks. One could also question whether the findings are also valid for the population at large. For younger or less-gifted populations (e.g., children or people with specific brain damage), monitoring might be far from effortless. Furthermore, it should be clear that the present results do not necessarily imply that the conflict-monitoring process operates in a completely automatic manner (i.e., that it does not require *any* cognitive resources at all). The point is that the monitoring process was unaffected by a secondary task that did have an effect on analytic reasoning. It is therefore safe to say that the monitoring process is far less demanding than the analytic reasoning process that is needed to complete the reasoning problem.

If we are right about the successful and effortless nature of the conflict-monitoring process, one could wonder why people still make reasoning errors. De Neys and Glumicic (2008) have already suggested that reasoning errors need to be attributed to a failure to inhibit heuristic responses rather than to a monitoring failure (see also Evans, 2007a; Houdé & Moutier, 1996; Stanovich, in press). Although people might detect that the heuristic response conflicts with the analytically appropriate response, they will not always manage to inhibit the tempting heuristic. A recent neuroimaging study by De Neys et al. (2008) presented evidence for this claim. In their study, participants solved base-rate neglect problems while the activation of brain regions believed to be involved in conflict detection (anterior cingulate cortex) and response inhibition (lateral prefrontal cortex) was monitored. Scanning results showed that the conflict detection area was always activated when the participants solved incongruent problems, but this activation was unrelated to the performance on these problems. This finding confirms the hypothesis that conflict detection is successful most of the time.

The inhibition area, however, was only activated when conflict problems were solved correctly. Taken together, this line of research indicates that reasoning errors are the result of a failure to inhibit heuristic processes rather than the result of a failure to detect the erroneous nature of the heuristic response in itself. The present study clarifies that this monitoring for heuristic–analytic conflict during judgement is a successful process that operates rather automatically.

Manuscript received 4 October 2007

Revised manuscript received 26 November 2008

First published online 21 February 2009

REFERENCES

- De Neys, W. (2006a). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science*, *17*, 428–433.
- De Neys, W. (2006b). Automatic-heuristic and executive-analytic processing in reasoning: Chronometric and dual task considerations. *Quarterly Journal of Experimental Psychology*, *59*, 1070–1100.
- De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of reasoning. *Cognition*, *106*, 1248–1299.
- De Neys, W., Vartanian, O., & Goel, V. (2008). Smarter than we think: When our brain detects that we are biased. *Psychological Science*, *19*, 483–489.
- De Neys, W., & Verschueren, N. (2006). Working memory capacity and a notorious brain teaser: The case of the Monty Hall Dilemma. *Experimental Psychology*, *53*, 123–131.
- Epstein, S. (1994). Integration of the cognitive and psychodynamic unconscious. *American Psychologist*, *49*, 709–724.
- Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin and Review*, *13*, 378–395.
- Evans, J. St. B. T. (2007a). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning*, *13*, 321–329.
- Evans, J. St. B. T. (2007b). *Hypothetical thinking: Dual processes in reasoning and judgement*. Hove, UK: Psychology Press.
- Evans, J. St. B. T. (in press). How many dual process theories do we need: One, two or many? In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond*. Oxford, UK: Oxford University Press.
- Gigerenzer, G., Hell, W., & Blank, H. (1988). Presentation and content: The use of base-rates as a continuous variable. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 513–525.
- Houdé, O., & Moutier, S. (1996). Deductive reasoning and experimental inhibition training: The case of the matching bias. *Current Psychology of Cognition*, *15*, 409–434.
- Kahneman, D. (2002, December). *Maps of bounded rationality: A perspective on intuitive judgement and choice*. *Nobel Prize Lecture*. Retrieved 30 April 2007 from: http://nobelprize.org/nobel_prizes/economics/laureates/2002/kahnemann-lecture.pdf
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgement. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics & biases: The psychology of intuitive judgment* (pp. 49–81). New York: Cambridge University Press.

- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237–251.
- 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.
- 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.
- Slovan, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, 119, 3–22.
- Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.
- Stanovich, K. E. (in press). Distinguishing the reflective, algorithmic, and autonomous minds: Is it time for a tri-process theory? In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond*. Oxford, UK: Oxford University Press.
- Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate. *Behavioral and Brain Sciences*, 23, 645–726.
- Thompson, V. (in press). Dual process theories: A metacognitive perspective. In J. St. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond*. Oxford, UK: Oxford University Press.
- Tversky, A., & Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, 185, 1124–1131.
- 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*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

APPENDIX 1

The 12 base-rate neglect problems in the first version of the task are presented below (translated from Dutch). In the second version all base-rates were switched. Incongruent problems in version one became congruent problems in version two, and vice versa. For example, the reasoning problem featuring Els became the following in version two:

In a study 1000 people were tested. Among the participants there were 995 sixteen-year-olds and 5 forty-year-olds. Els is a randomly chosen participant of the study.

Els likes to listen to techno and electronic music. She often wears tight sweaters and jeans. She loves to dance and has a small nose piercing.

What is most likely?

1. *Els is 16 years old.*
2. *Els is 40 years old.*

Base-rates and descriptions no longer conflict, this is now a congruent problem.

Incongruent problems

Els:

In a study 1000 people were tested. Among the participants there were 5 sixteen-year-olds and 995 forty-year-olds. Els is a randomly chosen participant of the study.

Els likes to listen to techno and electronic music. She often wears tight sweaters and jeans. She loves to dance and has a small nose piercing.

What is most likely?

1. Els is 16 years old.
2. Els is 40 years old.

Etienne:

In a study 1000 people were tested. Among the participants there were 4 people who drive a second-hand Nissan and 996 people who drive a BMW. Etienne is a randomly chosen participant of the study.

Etienne is 38 years old. He works in a steel plant. He lives in a small apartment on the outskirts of Charleroi. His wife has left him.

What is most likely?

1. Etienne drives a BMW.
2. Etienne drives a second-hand Nissan.

Mario:

In a study 1000 people were tested. Among the participants there were 994 Swedes and 6 Italians. Mario is a randomly chosen participant of the study.

Mario is 25 years old. He is a charming young man and is a real womaniser. His favourite dish is the spaghetti his mother makes.

What is most likely?

1. Mario is a Swede.
2. Mario is an Italian.

Sarah:

In a study 1000 people were tested. Among the participants there were 995 Muslims and 5 Buddhists. Sarah is a randomly chosen participant of the study.

Sarah is 19 years old. She likes to philosophise and has an aversion to materialism. She wears second-hand clothes and would love to go to India one day.

What is most likely?

1. Sarah is a Buddhist.
2. Sarah is a Muslim.

Congruent problems

Aline:

In a study 1000 people were tested. Among the participants there were 996 people who like to watch Canvas and 4 people who like to watch VTM. Aline is a randomly chosen participant of the study.

Aline is 35 years old. She writes reviews for a magazine. Her husband works at the university. She loves painting and photography.

What is most likely?

1. Aline likes to watch Canvas.
2. Aline likes to watch VTM.

Debby:

In a study 1000 people were tested. Among the participants there were 994 people who live in the country and 6 people who live in the city. Debby is a randomly chosen participant of the study.

Debby is 22 years old. She rides a horse. After school she takes care of the animals at home. In the weekends she rises early and visits her grandparents.

What is most likely?

1. Debby lives in the country.
2. Debby lives in the city.

Jeanine:

In a study 1000 people were tested. Among the participants there were 5 people who vote for Groen! and 995 people who vote for Vlaams Belang. Jeanine is a randomly chosen participant of the study.

Jeanine is 67 years old. She's always worked on the assembly line. She attaches great importance to traditional values and lives in a residential area where there's a lot of crime.

What is most likely?

1. Jeanine votes for Vlaams Belang.
2. Jeanine votes for Groen!

Dominique:

In a study 1000 people were tested. Among the participants there were 5 women and 995 men. Dominique is a randomly chosen participant of the study.

Dominique is 32 years old and is a confident and competitive person. Dominique's goal is building a career. Dominique does a lot of sport and is well muscled.

What is most likely?

1. Dominique is a woman.
2. Dominique is a man.

Neutral problems

Hugo:

In a study 1000 people were tested. Among the participants there were 995 fifty-year-olds and 5 sixty-year-olds. Hugo is a randomly chosen participant of the study.

Hugo is very curious about new cultures. He likes to try food from other countries. He just returned from a vacation in Hungary.

What is most likely?

1. Hugo is sixty years old.
2. Hugo is fifty years old.

Jan:

In a study 1000 people were tested. Among the participants there were 4 people who play the saxophone and 996 people who play drums. Jan is a randomly chosen participant of the study.

Jan is 19 years old. He studies in Brussels and doesn't have a girlfriend. He's just bought an old, second-hand car with the money he saved.

What is most likely?

1. Jan plays the saxophone.
2. Jan plays drums.

Pablo:

In a study 1000 people were tested. Among the participants there were 5 Club Brugge fans and 995 Anderlecht fans. Pablo is a randomly chosen participant of the study.

Pablo is 39 years old. He's a dedicated football fan. His week starts off badly when his team loses. He takes his son to watch every home game.

What is most likely?

1. Pablo is an Anderlecht fan.
2. Pablo is a Club Brugge fan.

Bart:

In a study 1000 people were tested. Among the participants there were 994 people from Antwerp and 6 people from Amsterdam. Bart is a randomly chosen participant of the study.

Bart is sixteen years old and still goes to school. He weighs 80 kilos and has a little sister of 14 years old and an older brother who has already been attending university for two years.

What is most likely?

1. Bart is from Antwerp.
2. Bart is from Amsterdam.

APPENDIX 2

There were two questions about the descriptions from each problem. Details about the problem at hand (name of the person, social groups involved) remained visible on the screen. For example: *Among the problems you just solved there was one about Els whose description was randomly chosen from a study in which there were sixteen-year-olds and forty-year-olds.*

The participants saw four statements about the person, and were asked to indicate which one was correct (“*Which of the following statements is correct?*”). The statements were numbered, and the participants had to press the number that corresponded with their answer. The response options for every question are presented below.

Els:

Question 1

1. Els likes to listen to classical music
2. Els likes to listen to Studio Brussels
3. Els likes to listen to techno
4. Els drinks two litres of water every day

Question 2

1. Els has a birth mark on her belly
2. Els has a nose piercing
3. Els has a tattoo
4. Els wears a diamond necklace

Etienne:

Question 1

1. Etienne lives in Liège
2. Etienne lives in Charleroi
3. Etienne lives in Hasselt
4. Etienne lives in Knokke

Question 2

1. Etienne works at a lawyer’s office
2. Etienne works at a steel plant
3. Etienne works for the government
4. Etienne works for a building industry

Mario:

Question 1

1. Mario is a womaniser
2. Mario has been in a relationship for two years
3. Mario is very punctual

4. Mario gives a lot of compliments

Question 2

1. Mario likes pizza
2. Mario likes spaghetti
3. Mario likes fish
4. Mario likes readings novels

Sarah:

Question 1

1. Sarah likes to philosophise
2. Sarah lives with her parents
3. Sarah is non-violent
4. Sarah follows the rules of her religion

Question 2

1. Sarah wants to go to Thailand
2. Sarah wants to go to Saudi Arabia
3. Sarah wants to go to Spain
4. Sarah wants to go to India

Aline:

Question 1

1. Aline is a cook
2. Aline writes reviews
3. Aline is a cleaning lady
4. Aline is a doctor

Question 2

1. Aline loves her husband very much
2. Aline likes to paint
3. Aline gossips a lot
4. Aline speaks French

Debby:

Question 1

1. Debby often plays squash
2. Debby rides a horse
3. Debby wants to become a vet
4. Debby is a fan of the local football team

Question 2

1. Debby knows how to milk a cow
2. Debby loves babies
3. Debby has animals at home
4. Debby doesn't have room for animals

Jeanine:

Question 1

1. Jeanine is a housewife
2. Jeanine works on the assembly line
3. Jeanine is a psychologist
4. Jeanine is a cashier

Question 2

1. Jeanine lives in a home for the elderly
2. Jeanine shares an apartment with friends
3. Jeanine lives in a residential area
4. Jeanine lives in Oostende

Dominique:

Question 1

1. Dominique is emotional
2. Dominique is aggressive
3. Dominique is self-assured
4. Dominique is careless

Question 2

1. Dominique is not overweight
2. Dominique is well muscled
3. Dominique recently had an accident
4. Dominique is worried about his or her weight

Hugo:

Question 1

1. Hugo likes to try food from other countries
2. Hugo is a chef in a Chinese restaurant
3. Hugo wears traditional African clothes
4. Hugo is about to move to Kenya

Question 2

1. Hugo is unreliable
2. Hugo is a miser
3. Hugo is very funny
4. Hugo is very curious about new things

Jan:

Question 1

1. Jan is married
2. Jan is divorced
3. Jan doesn't have a girlfriend
4. Jan lives with his girlfriend

Question 2

1. Jan doesn't have a car
2. Jan drives a second-hand car
3. Jan drives a mini-van
4. Jan drives a grey car

Pablo:

Question 1

1. Pablo goes to watch every home game of his team
2. Pablo only goes to watch his team when they play on location
3. Pablo only watches football on TV
4. Pablo watches every game of his team

Question 2

1. Pablo takes his cousin to watch football games
2. Pablo takes his dad to watch football games
3. Pablo watches football games with his friends
4. Pablo takes his son to watch football games

Bart:

Question 1

1. Bart works in a video store
2. Bart studied law
3. Bart is unemployed
4. Bart still goes to school

Question 2

1. Bart has a brother and a sister
2. Bart is an only child
3. Bart has two brothers
4. Bart has one sister