UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Implicit Conflict Detection During Decision Making

Permalink
https://escholarship.org/uc/item/4rk7w3xn

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 29(29)

ISSN
1069-7977

Author
De Neys, Wim

Publication Date
2007

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Implicit Conflict Detection During Decision Making

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

Abstract

Popular dual process theories of reasoning and decision making have characterized human thinking as an interplay of an intuitive and analytic reasoning process. Although monitoring the output of the two systems for conflict is crucial to avoid decision making errors there are some widely different views on the efficiency of the process. Kahneman (2002) claims that the monitoring of the intuitive system is typically quite lax whereas others such as Sloman (1996) and Epstein (1994) claim it is flawless and people typically experience a struggle between what they "know" and "feel" in case of a conflict. The present study contrasted these views. Participants solved classic base rate neglect problems while thinking aloud. Verbal protocols showed no evidence for an explicitly experienced conflict. As Kahneman predicted, participants hardly ever mentioned the base rates and seemed to base their judgment exclusively on heuristic reasoning. However, a more implicit measure of conflict detection based on participants' retrieval of the base rate information in an unannounced recall test showed that the base rates had been thoroughly processed. Results indicate that although the popular characterization of conflict detection as an actively experienced struggle needs to be revised there is nevertheless evidence for Sloman and Epstein's basic claim about the flawless operation of the conflict monitoring process.

Introduction

Human reasoners seem to have a strong preference to base judgments on prior beliefs and intuition rather than on a logical reasoning process. Over the last decades numerous studies have shown that this tendency is biasing performance in many classic reasoning and decision making tasks (Evans, 2003; Kahneman & Tversky, 1973).

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; Goel, 1995; Kahneman, 2002; 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 fast and automatically whereas the operations of the analytic system would be slow and heavily demanding of people's computational resources. Dual process theories state that the heuristic and analytic system will often interact in concert. Hence, on these occasions the heuristic default system will provide us with fast, frugal, and correct conclusions. However, the prepotent heuristics can 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).

Although the dual process framework has been very influential (and with the work of D. Kahneman & A. Tversky even indirectly awarded a Nobel prize) it has also been criticized (e.g., Evans, in press; Gigerenzer & Regier, 1996; Osman, 2004; Stanovich & West, 2000). The characterization of the conflict detection process is a crucial case in point. Dual process theories generally state that the analytic system is monitoring the output of the heuristic system. When a conflict with analytic knowledge (e.g., sample size considerations) is detected, the analytic system will attempt to intervene and inhibit the prepotent heuristic response. However, if one looks at the literature it becomes clear that there are some widely different views on the efficiency of the conflict monitoring component during judgement and decision making. The classic work of Kahneman and colleagues, for example, claims that the monitoring of the heuristic system is quite lax (e.g., Kahneman & Frederick, 2005). It is assumed that by default people will tend to rely on the heuristic route without taking analytic considerations into account. In some cases people can detect the conflict and the analytic system will intervene but typically this will be quite rare. Most of the time people will simply not be aware that their response might be incorrect from a normative point of view. As Kahneman and Frederick (p. 274) put it: "People who make a casual intuitive judgement normally know little about how their judgment come about and know even less about its logical entailments". Thus, in this view people mainly err because they fail to detect a conflict.

In the work of Epstein (1994) and Sloman (1996) one finds a remarkably different view on conflict monitoring and the nature of reasoning errors. These authors assume that in general the heuristic and analytic routes are simultaneously activated and people typically do experience a conflict between two types of reasoning. People would "simultaneously believe two contradictory responses" (Sloman, 1996, p. 11) and therefore "behave against their better judgement" (Denes-Raj & Epstein, 1994, p. 1) when they err. Thus, people would be taking analytic considerations in mind and notice that they conflict with the heuristically cued belief. The problem, however, is that they do not always manage to override the compelling heuristics. In this view there is nothing wrong with the conflict detection process. Errors arise because people fail to inhibit the prepotent heuristic beliefs.

Based on the available data it is hard to decide between the different models and determine which conflict detection view is correct (e.g., Evans, in press). The most compelling evidence for successful conflict detection during decision

making comes from a number of intriguing anecdotes and spontaneous reports. Epstein (1994; Denes-Raj & Epstein, 1994; Epstein & Pacini, 1999), for example, repeatedly noted that when picking an erroneous answer his participants spontaneously commented that they did "*know*" that the response was wrong but stated they picked it because it "*felt*" right. Sloman (1996) cites evolutionary biologists Steven Jay Gould who relates experiencing a similar conflict between his logical knowledge and a heuristically cued stereotypical belief when solving Kahneman's and Tversky's infamous "Linda" problem. The problem, however, is that spontaneous self-reports and anecdotes are no hard empirical data. This is perhaps best illustrated by the fact that Kahneman (2002, p. 483) also refers to "casual observation" of his participants to suggest that only in "some fraction of cases, a need to correct the intuitive judgements and preferences will be acknowledged". It is clear that in order to conclude something about the efficiency of the conflict detection we need a straightforward empirical test to establish precisely how frequently people experience this conflict. The present study addresses this issue.

The present work adopted a thinking aloud procedure (e.g., Ericsson & Simon, 1980). Thinking aloud protocols have been shown to have a superior validity compared to interpretations that are based on retrospective questioning or people's spontaneous remarks (Payne, 1994).

Participants were asked to solve problems that were modeled after Kahneman and Tversky's classic (1973) base rate neglect problems. In these problems people first get information about the composition of a sample (e.g., a sample with 995 females and 5 males). People are told that short personality descriptions are made of all the participants and they will get to see one description that was drawn randomly from the sample. Consider the following example:

> In a study 1000 people were tested. Among the participants there were 4 men and 996 women. Jo is a randomly chosen participant of this study.
>
> Jo is 23 years old and is finishing a degree in engineering. On Friday nights, Jo likes to go out cruising with friends while listening to loud music and drinking beer.
>
> What is most likely?
> a. Jo is a man
> b. Jo is a woman

The normative response based on the group size information is (b). However, people will be tempted to respond (a) on the basis of heuristic beliefs cued by the description.

Given Kahneman and Tversky's (1973) classic findings one can expect that in the majority of cases people will err and pick the heuristically cued response in this task. The crucial question is whether people's verbal protocols indicate that they nevertheless take analytic considerations into account. In this task "analytic considerations" can be operationalized as referring to the group size information during the reasoning process (e.g., " … because Jo's drinking beer and loud I guess Jo'll be a guy, although there *were more women* …"). Such basic sample size reference during the reasoning process can be considered as a minimal indication of successful conflict monitoring. It shows that this information is not simply neglected. If Sloman and Epstein's idea about the parallel operation of the heuristic and analytic route is correct, such references should be found in the majority of cases. If Kahneman's idea about the lax nature of the conflict monitoring is correct, people will simply not be aware that the base rates are relevant and should hardly ever mention them during decision making.

It should be noted that both camps in the conflict monitoring debate, as the reasoning field at large, have conceptualized the conflict between the analytic and heuristic system as a consciously experienced, verbalizable event. Conflict monitoring is considered as a controlled process arising from the central executive aspect of working memory. Since James (1890) there is indeed a long tradition in psychology to consider such central, controlled (vs. automatic) processing as being consciously experienced (Feldman-Barrett, Tugade, & Engle, 2004). However, the available evidence from the cognitive literature suggests that this needs not always be the case. Although controlled processing can occur with a feeling of conscious deliberation and choice, it needs not (Feldman-Barrett et al., 2004).

While it is held that thinking-aloud is an excellent method to tap into the content of conscious thinking it cannot provide us with the information about cognitive processes that do not reach the conscious mind (Crutcher, 1994). Consequently, even if participants do not verbalize their experience of the conflict, one cannot exclude that the conflict monitoring might nevertheless have been successful. To capture such *implicit* detection participants were also presented with an unannounced recall test in the present study. After a short break following the thinking-aloud phase participants were asked to answer questions about the group sizes in the previous reasoning task. If people have successfully detected the conflict this implies that the group size has been taken into account and people spent some time processing it. Indeed, the detection of the conflict should trigger analytic system intervention which should result in some further scrutinising of the sample information. In sum, successful conflict detection should be accompanied by a deeper processing of the base rate information which should benefit recall. This recall index does not require that the conflict is consciously experienced or verbalizable.

To validate the recall hypothesis participants were also presented with additional control problems. In the classic base rate problems the description of the person is composed of common stereotypes of the smaller group so that base rates and description disagree. In addition to these classic problems we also presented problems where base rates and description both cued the same response. In these *congruent* problems the description of the person was composed of stereotypes of the *larger* group (e.g., Ferreira, Garcia-Marques, Sherman, & Garrido, 2006). Hence, contrary to the classic (i.e., *incongruent*) problems base rates and description did not

conflict and the response could be rightly based on the salient description without further analytic intervention.

Method

Participants. Twelve undergraduate students at York University (Toronto, Canada) participated in return for credit in a psychology course.

Material. *Decision-making task*. Participants solved a total of 18 problems that were modelled after Kahneman and Tversky's (1973) base rate neglect items. Six of these were the crucial *incongruent* problems where the description of the person was composed of common stereotypes of the smaller population group tested (i.e., the description and the base rates conflicted). There were also six *congruent* control items where the description and the base rates agreed. Finally, we also presented six *neutral* control items where the description only mentioned characteristics that were neutral with respect to group membership (e.g., 'Jo has blue eyes and black hair') while the base rates were indicating which group was larger.

The order of the two response options ('a' and 'b') was counterbalanced. Problems were printed one to a booklet. The problems were presented in pseudo-random order. Participants always started with an incongruent problem followed by a congruent and neutral problem. The remaining problems were presented in a randomly determined order.

Recall Task. Participants were asked to write down the base rates for each problem they previously solved. The following is an example of the recall question:

> One of the problems you just solved concerned Jo whose description was drawn at random from a sample of men and women. Try to answer the following questions.
>
> How many men were there exactly in the study? ____ (write down)
> How many women were there exactly in the study? ____ (write down)

After the base rate question followed two easy filler questions in multiple choice format that referred to the description of the problem. Performance on these filler problems was uniformly high. Each base rate question together with the two filler questions was printed one to a page in a booklet. Recall questions were presented in the same order as the decision making problems had been solved.

Procedure

Participants were first introduced to the thinking aloud procedure. Participants were told to start by reading the complete problem aloud, and say everything that they were thinking about while solving it. The complete session was tape-recorded. Coding of the verbal reports simply focused on whether the participants gave the correct answer[1] and whether they referred to the base rate information during decision making. A statement like " … because Jo's drinking beer and likes loud music I guess *Jo'll be a guy*, although there *were more women*" would be coded as an incorrect response since the participant did not pick the response (i.e., women) consistent with the largest sample group and as an instance of base rate mentioning. The following are some straightforward further illustrations of the protocol codings:

> … This guy is an engineer, because he likes computers and science fiction, and he seems like a loner...no wife. (Participant #12: incorrect response, base rates not mentioned)

> … It depends how you want to go **if you want to go according to the statistics there is a greater chance he is a lawyer** but because of the things he does, he is introverted, spends his time writing computer games it makes more sense that he is an engineer so … I don't know I will go with that. (Participant #1: incorrect response, base rates mentioned)

> … ok **5 engineers...** you would think he is an engineer **but cause there were more lawyers** he is a lawyer**.** (Participant #6: correct response, base rates mentioned)

After completing the decision-making task, participants had a short break and then were presented with the recall task. The recall task was not announced at the start of the experiment so participants did not know base rate recall would be tested until they had completed the decision-making task. Recall performance was scored in terms of whether the direction of the base rates was correctly recalled (i.e., which population group mentioned in the problem was larger and which group was smaller).

Results

Decision Making Task. On each problem we coded whether the participant gave the correct answer (i.e., accuracy) and whether the participant referred to the base rate information during decision making (i.e., base rate mentioning). Figure 1 present an overview of the mean performance on the different problem types.

As in Tversky's and Kahneman's classic studies, accuracy on the incongruent problems was very low. Participants were clearly biased by the salient description and selected the correct response in fewer than 20% of the cases. As expected, participants had far less difficulties with the neutral and congruent problems where the description was simply neutral or consistent with the base rates. An ANOVA established that the difference in accuracy between the problem types was significant, $F(2, 22) = 54.07$, $p < .001$.

[1] Consistent with previous dual process studies, responses that were in line with the base rates (i.e., selection of the largest group as most likely answer) were labelled as correct answers. It should be noted that especially in the case of the classic, incongruent problems the actual normative status of the 'correct' response is sometimes debated (Gigerenzer, Hell, & Blank, 1988).

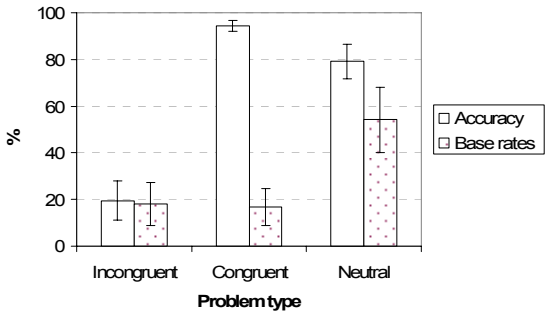

Figure 1. Mean proportion correct responses and explicit base rate mentioning in verbal protocols. Error bars are standard errors.

The more crucial question, however, is to what extent people take analytic considerations into account when solving these problems and refer to the base rates during decision making. An ANOVA established that the frequency of base rates mentioning depended on the type of problem, $F(2, 22) = 9.50$, $p < .005$. As Figure 1 shows, the verbal protocols indicate that on the majority of the neutral problems (54%) participants are considering the base rate information. However, once these same people are faced with a stereotypical description in the congruent and incongruent problems they seem to be completely discarding the base rates. On the crucial incongruent problems the base rates are mentioned only 18% of the time. People seem to be exclusively referring to the match between their response and the description without much evidence for a consciously experienced conflict.

Table 1 presents some interesting additional data. As Table 1 indicates, the few times participants did mention the base rates on the incongruent problems they also tended to solve the problem correctly most of the time (70%). The other way around, whenever participants did manage to give the correct response they typically (60% of the cases) also referred to the base rates. The same pattern was observed for the neutral problems. Indeed, accuracy and base rate mentioning correlated for the incongruent, $r = .92$, $p < .001$, and neutral problems, $r = .88$, $p < .001$. Not surprisingly, for the congruent problems where the description cues the correct response, accuracy did not depend on base rate mentioning, $r = .22$. In sum, whenever the classic incongruent problems were solved correctly, people successfully detected the conflict between the description and base rates. However, people erred on the vast majority of the problems and there was hardly any evidence for a consciously experienced conflict in these cases. Indeed, on the 80% of the incongruent problems that were solved incorrectly participants mentioned the base rates only 6% of the time. Consistent with Kahneman's claim about the lax nature of the conflict monitoring process, people do not seem to be aware that the base rates are relevant for solving the incongruent problems.

Table 1
Overview of Additional Performance Measures

	Problem Type		
Measure	Incongruent	Congruent	Neutral
% correct when base rates mentioned	70 (48)	88 (25)	100 (0)
% base rates mentioned when correct	60 (55)	16 (27)	56 (50)
% base rates mentioned when incorrect	6 (15)	25 (50)	0 (0)
r (base rate mentioning and accuracy)	.92*	.22	.88*
First Problem			
% Correct	0 (0)	92 (29)	83 (39)
% Base mentioned	0 (0)	8 (29)	50 (52)

* $p < .001$, standard deviations in parentheses.

One reason for the lack of base rate mentioning in the present experiment might be the repetitive nature of the problem presentation. Participants had to solve a total of six incongruent problems and they might have stopped verbalizing their processing of the base rates after a while because they became less motivated or because they figured they had already sufficiently clarified their reasoning on the previous trials. Such confound would have decreased the average performance. We therefore examined the data for the first three presented problems separately. The first one of these was always an incongruent problem, the second one a congruent, and the third one a neutral problem. As Table 1 shows, the general pattern was present right from the start. Contrary to the motivation hypothesis, performance on the first, incongruent problem was even worse. None of the participants solved it correctly or mentioned the base rates.

Recall Task. Figure 2 presents an overview of the recall findings. The verbal protocols already indicated that participants were taking base rates into account when solving the neutral problems. Accuracy was high and participants mentioned the base rates on the majority of the trials. As Figure 2 shows, the processing of the base rates during the neutral problem solving also resulted in a decent recall performance. Although participants did not know they had to memorize the base rates during decision making, they correctly identified which group was the largest on 66% of the neutral problems. For the congruent trials, where the description cued the correct response and base rates were hardly explicitly considered, participants only recalled 36% of the base rates correctly. The crucial finding, however, concerns the incongruent problems. Although the verbal protocols showed no evidence for a consciously experienced conflict and participants seemed to be almost completely

neglecting the base rates, recall performance did indicate that the base rates had been processed. With 69% correct identification recall was at par with the neutral problems and clearly superior to the recall for the congruent problems were there was no conflict to be detected. An ANOVA established that the recall performance significantly differed between the problem types, $F(2, 22) = 9.26$, $p < .001$.

The only difference between the congruent and incongruent problems was the conflicting nature of the description and base rates. If people would not be detecting the conflict and would simply neglect the base rate information on the incongruent problems, as the verbal reports suggested, recall performance for congruent and incongruent problems should not have differed.

Figure 2 also shows the results of a number of additional control analyses. One could argue that the better recall on the incongruent problems might have been inflated because of the few trials where the base rates were explicitly mentioned. A purer measure of implicit conflict detection would concern the recall performance on those trials where the base rates were not explicitly mentioned. Figure 2 presents the results of an extreme test of this hypothesis. Eight participants never mentioned the base rates on any of the incongruent problems they solved. As Figure 2 shows, they nevertheless showed a similar recall pattern. Although they never mentioned the base rates on the incongruent problems, recall was still at par with the neutral problems and clearly superior to the congruent problems, $F(2, 14) = 4.55$, $p < .05$.

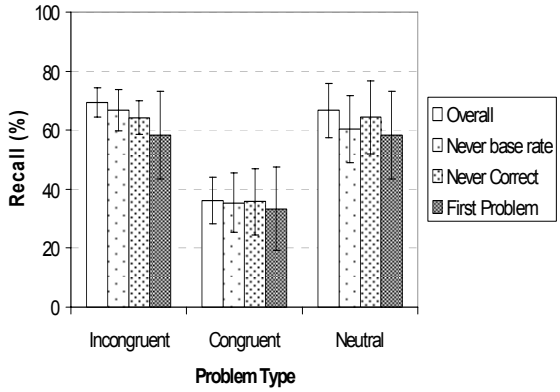

Figure 2. Mean overall proportion of correct base rate recall.

Similarly, one can look at accuracy and restrict the analysis to those participants who did not give a single correct response on any of the incongruent problems. This was the case for seven participants. As the recall findings in Figure 2 show, even people who always erred showed the superior recall for incongruent problems. The recall effect still reached marginal significance, $F(2, 12) = 3.39$, $p < .07$, in this small group.

One could also remark that the recall findings resulted from the repeated testing in the present experiment. The within-subject design might have made the conflict especially salient and cued a more profound conflict monitoring. To check this hypothesis we examined the recall data for the first three presented problems separately. The first one of these was always an incongruent problem, the second one a congruent problem, and the third one a neutral problem. As Figure 2 demonstrates, although correct recall for the first items tended to decrease somewhat the basic recall pattern was present right from the start. Base rates of the first incongruent (58%) and neutral problem (58%) were still memorized almost twice as well as the base rates of the first congruent problem (33%), $F(1, 11) = 11.96$, $p < .01$.

A final alternative explanation for the better base rate recall for incongruent and neutral problems vs. congruent problems might be the serial position of the presented problems. It is well established in memory studies that the first and last items on a list are better recalled than items that are presented closer to the middle. Although we used an unannounced recall procedure, the findings could have been affected if incongruent and neutral problems were presented more frequently in the beginning or at the end of the experiment. We therefore calculated the average distance of the 18 items from the middle position in the presentation order (i.e., the first problem received rank 1, the eighth and tenth problem rank 8 and so on). Incongruent and congruent problems had the same average distance (i.e., position 4.66) whereas the neutral items were actually presented somewhat closer to the middle (i.e., position 5.67). This shows that the presentation position factor cannot account for the recall pattern findings. Indeed, if the serial position would explain the better recall on the first (incongruent) over the second (congruent) problem, for example, recall on the thirdly presented neutral problem should have been even worse. As Figure 2 shows, this was clearly not the case.

Discussion

The present study showed that when people solve classic base rate problems there is hardly any evidence for an explicitly experienced conflict between problem solutions that are cued by the analytic and heuristic reasoning system. Only in 18% of the cases participants referred to the base rates and indicated they were taking analytic considerations in mind. However, the recall data showed that the base rates were not merely neglected. We might not be consciously experiencing an active struggle but our cognitive system does seem to be detecting the special status of the incongruent problems. Even when participants never mentioned the base rates and always erred on the incongruent problems they nevertheless managed to correctly identify which group was larger on the vast majority of the problems. For the congruent problems where the descriptions and base rates agreed this was not the case. If people were not detecting the conflict and were simply neglecting the base rate information on the incongruent problems, recall performance for congruent and incongruent problems should not have differed. In sum, while the study showed that the anecdotal characterization of conflict detection as an actively experienced struggle is far from prototypical, there is evidence for Sloman and Epstein's basic idea about the efficiency of the conflict monitoring process. Even when we err our reasoning engine seems to be picking

up that the description disagrees with the base rates. This suggests that the dominance of heuristic reasoning should not be attributed to a lack of conflict monitoring.

The evidence for the efficiency of the conflict monitoring during decision making has some important implications for the debate on human rationality (e.g., Stanovich and West, 2000). This rife debate centres around the question whether the traditional norms (such as standard logic and probability theory) against which the rationality of peoples decisions are measured are valid. It has been questioned for example why preferring base rates over beliefs would be more rational or "correct" than pure belief-based reasoning (e.g., Oaksford & Chater, 1998; Todd & Gigerenzer, 2000). One reason for criticizing the norm has been the consistent very low number of correct responses that has been traditionally observed on the classic reasoning and decision making tasks. If over 80% of well-educated, young adults fail to solve a simple decision making task, this might indicate that there is something wrong with the task scoring norm rather than with the participants. However, the debate, as the vast majority of dual process research, has often been characterized by an exclusive focus on people's response output (i.e., whether or not people manage to give the correct response) and not on the underlying cognitive processes (De Neys, 2006; Gigerenzer et al., 1988). The present data clarify that giving an erroneous belief-based response does not imply mere belief-based reasoning where people completely disregard the traditional norm. Results indicate that even people who consistently err detect the conflict between base rates and the description and allocate additional resources to a deeper base rate processing. If people did not believe that the group size information matters during problem solving, they would not waste time processing it. People might not always manage to adhere to the norm but they are clearly not simply discarding it or treating it as irrelevant. This should at least give pause for thought before rejecting the validity of the traditional norms. Clearly, people are more normative than their mere judgements show.

Acknowledgements

Wim De Neys is a post doctoral fellow of the Flemish Fund for Scientific Research. The study was conducted during a stay at Vinod Goel's lab at York University, Toronto (Canada). I would like to thank Tamara Glumicic for her help running the study.

References

Crutcher, R. J. (1994). Telling what we know: *Psychological Science*, *5*, 241-252.

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

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

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

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215-251.

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. (in press). On the resolution of conflict in dual process theories of reasoning. *Thinking and Reasoning*.

Feldman Barrett, L., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological Bulletin, 130*, 553-573.

Ferreira, M. B., Garcia-Marques, L., Sherman, S. J., & Garrido, M. (2006). Automatic and controlled components of judgment under uncertainty. *Proceedings of the Cognitive Science Society, 28*, 1293-1298.

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.

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

Goel, V. (1995). *Sketches of thought*. Cambridge, MA: MIT Press.

James, W. (1890). *The principles of psychology*. Oxford, England: Holt.

Kahneman, D. (2002, December). Maps of bounded rationality: A perspective on intuitive judgement and choice. Nobel Prize Lecture.

Kahneman, D. & Frederick, S. (2005). A model of heuristic judgement. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge Handbook of Thinking and Reasoning* (pp. 267-293). Cambridge, MA: Cambridge University Press.

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

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

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

Payne, J. W. (1994). Thinking aloud: Insights into information processing. *Psychological Science, 5*, 241-248.

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

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

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