Are representativeness judgments automatic and rapid? The effect of time pressure on the conjunction fallacy

Gaëlle Villejoubert (gaelle.villejoubert@univ-tlse2.fr)
Laboratoire CLLE-LTC (UTM, CNRS, EPHE)
Université de Toulouse Le Mirail
5 allées A. Machado
31058 Toulouse Cedex 9, France

Abstract

This study aimed to better understand the cognitive underpinnings of people's intuitions concerning the conjunction rule of probabilities. Using a novel methodology to study the conflict between representativeness and logic considerations in the conjunction fallacy, results revealed that, contrary to what is commonly assumed, assessments of representativeness are not automatic and rapid but are in fact most influential when participants are not pressured by time to provide a judgment. Implications for the different dual-system accounts of probability judgments will be discussed.

Keywords: Conjunction fallacy, Probability judgments, Intuition, Dual processes, Representativeness, Time pressure.

Introduction

The question of whether we can intuitively appreciate mathematical concepts has puzzled many psychologists and linguists (e.g., see Lakoff & Núñez, 2000). There is some strong evidence that this achievement is not beyond our capabilities. For example, Young Brazilian candy sellers with little formal education can outperform their non-sellers counterparts on complex ratio-comparison problems presented in familiar formats (i.e., using candies and bills) without using pen or paper (Saxe, 1988). Yet, the automatisation of mathematical operations can have its shortcomings. For example, erroneous intuitions about probabilities have been demonstrated numerous times. One striking example concerns intuitions about the conjunction rule, a basic probability law which does not seem to be intuitively grasped by individuals (Tversky & Kahneman, 1983). According to this rule, the probability of a conjunction of two events A and B, or $\Pr(A \wedge B)$, cannot exceed the probability of any of its constitutive events, namely $\Pr(A)$ or $\Pr(B)$. In certain situations, however, most individuals will erroneously judge that $\Pr(A \wedge B)$ is more probable than one of its constitutive probabilities, an error known as the "conjunction fallacy" (Tversky & Kahneman, 1983). Tversky and Kahneman (1983) provided a paradigmatic illustration of this error with the "Linda problem". In this problem, individuals are asked to read the following thumbnail description:

> Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations. (p. 297)

In the most transparent test of individuals' intuition for the conjunction rule, the authors asked participants to decide which of the two following alternatives was more probable:

(1) Linda is a bank teller.

(2) Linda is a bank teller and is active in the feminist movement.

Overall, 85% of respondents chose (2) over (1), thus revealing that their intuitive answer was at odds with the conjunction rule.

Why Does our Intuition Fail us?

There are two main types of accounts to explain the discrepancy between participants' intuitions and the objective mathematical rule in conjunction tasks. The first type of account focuses on *linguistic ambiguities* inherent in the task (e.g., see Hertwig, Benz, & Krauss, 2008; Politzer & Noveck, 1991). Researchers who emphasised the role of linguistic ambiguity proposed, notably, that people inferred that (1) implicitly negated the possibility that Linda was a feminist when asked to compare (1) and (2) (see e.g., Politzer & Noveck, 1991). In other words, when asked to compare $\Pr(A)$ and $\Pr(A \wedge B)$, individuals would instead compare $\Pr(A \wedge \text{not-}B)$ and $\Pr(A \wedge B)$. In this case, considering that $\Pr(A \wedge \text{not-}B)$ is greater than $\Pr(A \wedge B)$ can no longer be deemed to be a 'fallacy'. There is, however, evidence to show that removing the ambiguity is not sufficient to perfectly realign people's intuitions with mathematical rules. Thus, Tentori, Bonini, and Osherson (2004) asked individuals to evaluate $\Pr(A)$, $\Pr(A \wedge B)$, as well as $\Pr(A \wedge \text{not-}B)$[1]. Results showed that the majority of participants erroneously believed that $\Pr(A \wedge B)$ was more probable than $\Pr(A)$.

The second type of account, so-called the *attribute-substitution* account (Kahneman & Frederick, 2002; Kahneman, 2003), is presented as a more specific, process-oriented extension of the representativeness heuristic account (Tversky & Kahneman, 1983). According to Kahneman and Frederick (2002; but see also, Kahneman, 2003), people's intuitions deviate from normative expectations because individuals assess an explicit *target attribute* of a judgment by substituting a related *heuristic attribute* that comes more readily to mind. So, according to this account, people use representativeness or similarity as a heuristic attribute to gauge the probability that Linda is a bank teller and a feminist and the probability that she is a bank teller. Accordingly, since Linda appears more representative of a feminist bank teller than a

[1] where the explicit mention of $\Pr(A \wedge \text{not-}B)$ prevents participants to interpret A as referring to A-but-not-B.

bank teller, people believe that she is more likely to be both a feminist and a bank teller than a bank teller.

Does our Intuition *Always* Mislead us?

There is a confound in Linda problems that has yet to be addressed. People's intuitions have been assessed with one of two methodologies. One procedure consists in asking individuals to rank order different events according to their probability of being true. These events encompass—solely, or among others—the conjunction of two events and each of these two events in isolation. The other procedure consists in asking participants to give a numerical estimate of the probability that the conjunction is true, as well as the probabilities that each of its constitutive events are true[2].

In both cases, there are two ways in which participants' intuitions can be fallacious. They can believe that the conjunction is more probable than *any* of its constitutive events or they can believe that the conjunction is more probable than *one* of its constitutive events but not the other. Tversky and Kahneman (1983) reported that 85% participants believed Linda resembled more a typical feminist bank teller than a bank teller ($A \wedge B \succ B$) but also that she resembled more a typical feminist than a feminist bank teller ($A \succ A \wedge B$). Therefore, if individuals assess the explicit probability attribute by substituting a related representativeness attribute, they should only commit the conjunction fallacy when comparing the least representative event and the conjunction of the least representative and the most representative one. In other words, they should erroneously believe that $\Pr(A \wedge B) > \Pr(B)$ but they should also correctly infer that $\Pr(A) > \Pr(A \wedge B)$. Unfortunately, the authors did not report probability judgments for most probable event $\Pr(A)$. But in a recent study, Hertwig et al. (2008, Experiment 2) data suggesting that nearly 70% of participants who committed the conjunction fallacy considered that the conjunction was more probable than one of its constituting events, but nevertheless *correctly* inferred that the conjunction was less probable than the second constitutive event.

This result suggests that our intuition can sometimes be aligned with rules of mathematics, that it does not *always* mislead us. The issue, therefore, lies in identifying whether these correct intuitions are solely determined by heuristic processing (i.e., considerations of representativeness) or whether they reflect a sensititivity to the logic of sets which underpins the conjunction rule. The current procedures and available empirical results, however, do not allow a more precise determination of whether participants actually detect the conflict between what considerations of representativeness dictate and what the logic of sets entails. This is because the role of intuitions about the logic of sets and intuitions based on representativeness considerations are not experimentally manipulated in the tasks commonly used. One main contribution of the present study is to propose a new methodology for studying the conjunction fallacy which will make it possible to test for the separate effects of logic and representativeness assessments on judgments.

Two Modes of Cognitive Functioning

A key element of the attribute-substitution account is the fact that the result of the evaluation of the heuristic attribute is presented as effortless and immediate (Kahneman, 2003). In other words, individuals are assumed to be able to judge how representative Linda is of a feminist or a bank teller almost synchronously upon reading the problem. As such, Kahneman (2003) argues, judgments based on representativeness originate from an intuitive mode of cognitive function, within which judgments are made rapidly and automatically. The author then contrasts this intuitive mode of functioning with a controlled mode, described as deliberate and slower. This dual-system view of cognition has gained considerable influence in the past decade in research examining judgment, decision-making or reasoning (e.g., see Evans, 2008; Kahneman, 2003; Sloman, 1996). According to this view, the second, more deliberate, system will either be at the origin of the judgment provided or simply monitor its quality. From this theoretical perspective, the conjunction fallacy results from the generation of a representativeness-based impression by the first, more intuitive system, while the second, slower and more deliberative, system fails to detect and correct the inference based on similarity. Kahneman and Frederick (2002) argued the controlled mode of cognitive functioning only loosely monitors the output of the intuitive system. Kahneman (2003) added that individuals will often be satisfied with the first plausible judgment that comes to mind and only will only rarely think hard to produce their judgment. There are concrete, empirically testable, predictions that can be derived from this position. Namely, Kahneman and Frederick's *lax-monitoring* view suggests that judgments observed in conjunction problems should always be rapid and automatic and solely influenced by considerations of representativeness.

The *lax-monitoring* view is at odds with other conceptions of dual-system cognition. There is evidence pointing to the possibility that individuals are, to some extent, sensitive to the conflict existing between heuristic and more analytical considerations. In reasoning, this is best illustrated by the belief bias phenomenon (Evans, Barton, & Pollard, 1983). Belief bias is illustrated by asking participants to evaluate syllogisms which vary both according to their logical validity (i.e., whether their conclusion logically follows from the premises) and their believability (i.e., whether the conclusions is congruent with prior beliefs). Evans et al. (1983) showed believable conclusions were more often endorsed than unbelievable ones but *also* that valid conclusions were more often endorsed than invalid ones. Moreover, Evans and Curtis-Holmes (2005) showed that the main effect of logic disappeared under time pressure, which is consistent with the proposition that heuristic responses would be automatic and

[2] A variant of this procedure consists in asking *frequency* estimates. This can greatly reduce the proportions of erroneous judgments (see e.g., Hertwig et al., 2008) but does not address, in itself, the issue of the origin of the erroneous *probabilistic* intuitions.

immediate but that, when the slower controlled mode of functioning can intervene, such responses are modified.

This *conflict-monitoring* view has recently also gained support in research on probability judgments. Thus, individuals who do *not* judge that the conjunction is more probable than any of its constitutive events will spend more time on producing a judgment (De Neys, 2006). Additionally, De Neys and Glumicic (2008, Experiment 2) showed that individuals took longer to make a judgment based on conflicting analytical and heuristic pieces of information. They gave participants information about the composition of a sample (e.g., 996 women and 4 men) and a short personality description supposedly drawn at random from that sample. The description was either congruent with stereotypes of the larger group or with stereotypes of the smaller group and hence, incongruent with the stereotypes of the larger group. For example, an incongruent description for the above sample described Jo as a student in engineering, who likes to go cruising, listen to loud music and drink beer. Participants took longer to correctly infer that the individual most probably belonged to the larger group when the description was incongruent with stereotypes on that group. Here again, concrete predictions can be derived from the conflict-monitoring view for conjunction probability judgments. First, statements presenting the most representative outcome as the most probable should be more often endorsed than statements presenting an unrepresentative outcome as most probable. Second, statements presenting a single event (independently of its representativeness) as more probable than a conjunction should be more often endorsed than statements presenting the conjunction of events as more probable. Third, the latter effect should disappear under time pressure because the controlled mode of cognitive functioning will no longer be able to exert its slower-paced monitoring.

Experiment

The following experiment was designed to test the competing predictions of the lax-monitoring view (Kahneman & Frederick, 2002; Kahneman, 2003) and the conflict-monitoring view (De Neys & Glumicic, 2008) of cognitive functioning in order to better understand the cognitive underpinnings of people's intuitions concerning the conjunction rule of probabilities. It first aimed to establish whether those intuitions were solely driven by heuristic processing (lax-monitoring view) or whether they were resulting from a combination of heuristic and controlled processing (conflict-monitoring view). The methodology used was adapted from the procedure used in belief-bias studies of reasoning (Evans & Curtis-Holmes, 2005). A second objective of this study was to assess the effect of time pressure on those probabilistic intuitions. Based on the theoretical predictions of the two-system view of cognition, it was expected that participants required to respond rapidly would show an increased level of heuristic processing and a reduced level of logical responding.

Method

Participants Participants were recruited by second-year psychology students at the University of Toulouse, France, as a course requirement. Each student emailed or contacted several men and women who were older than 18 to invite them to take part in an on-line study. Of the 127 participants in the final sample, 24 already knew the Linda problem. Three participants took over two hours to complete the test, and another ten participants allocated to the limited time condition spent more than 12 seconds on average to provide an answer. All these participants were removed from subsequent analyses. Of the remaining 90 participants (43 men and 47 women; mean age = 29.85, *SD* = 12.52), 31% had post-graduate education, 36% had undergraduate education, 24% had graduated from high school, and the remaining 9% had left education before graduating from high school. Their background knowledge was diverse, ranging from maths and science (26%), social sciences (19%), services (21%), Technology and design (17%). The experiment was conducted in French and all participants were native French speakers.

Materials Two scenarios used were based upon the original Linda and Bill descriptions (Tversky & Kahneman, 1983), respectively. In addition, ten new descriptions were modelled after these original texts. A crucial difference between these twelve tasks and the original Linda problem, however, was the judgment task.

The materials used in this study presented participants with a unique statement which they could either accept as correct or reject as incorrect. The statements presented fell into four categories: Correct-Congruent (CC), Correct-Incongruent (CI), Incorrect-Congruent (IC), and Incorrect-Incongruent (II). Table 1 provides an illustration of the four types of conclusions for the Linda problem. Note, however, that different scenarios were used for each of the twelve tasks created in the experiment so that all statements referred to different contents.

Table 1: Illustration of the four types of statements used to study the conjunction fallacy

Possible statements	
CC	The fact that Linda is a feminist is more likely than the fact that she is a feminist and a bank teller.
CI	The fact that Linda is a bank teller is more likely than the fact that she is a feminist and a bank teller.
IC	The fact that Linda is a feminist and a bank teller is more likely than the fact that she is a bank teller.
II	The fact that Linda is a feminist and a bank teller is more likely than the fact that she is a feminist.

Note. CC: Correct-Congruent, CI: Correct-Incongruent, IC: Incorrect-Congruent, II: Incorrect-Incongruent.

Correct statements proposed that an event was more likely than a conjunction of this event and another as in (3) or, al-

Figure 1: Screen shot of the training task

ternatively, that the conjunction was less likely than one of its constitutive events.

(3) The fact that Linda is a feminist is more likely than the fact that she is a feminist and a bank teller.

Conversely, an incorrect statement proposed that the conjunction was more likely than its constitutive events or, alternatively, that one of its constitutive event was less likely than the conjunction, as in (4).

(4) The fact that Simon is a businessman is less likely than the fact that he is a businessman and an opera singer.

Congruency was determined by the description presented. Each description was designed so that the individual presented appeared as a highly representative member of a given category without explicitly mentioning that this individual actually belonged to that category. Congruency was manipulated by presenting either the representative event or the unrepresentative event as most likely.

Design and procedure Participants were randomly allocated to one of two experimental conditions: the free time or the limited time condition. A screen explained that they would be presented with thumbnail descriptions of a person followed by a statement. They were asked to read the description carefully and to indicate if they thought the statement was correct (free time condition) or indicate if they thought the statement was correct *as fast as possible* (italics added, limited time condition).

The next screen presented an example of a thumbnail description. Participants were asked to read the description attentively and to click on a "Continue" button when they were ready. Upon clicking, a probability statement appeared and they were instructed to click on a "yes" or "no" box to indicate whether or not they thought the statement was correct. In the free time group, participants could take as long as they wished to provide an answer. In the limited time group, a clock and a countdown appeared 5 seconds after the presentation of the statement and went down for another 7 seconds, followed by the message: "Respond now!" (see Figure 1 for a screenshot). The next screen informed participants that they would now enter the test. The experimental tasks were presented in a random order, following the same procedure as that of the example task. Each time, the program recorded the participant's choice (accept or reject the conclusion) and the time taken to respond from the moment the conclusion appeared on the screen until the participant clicked a box. After they had provided all their answers, participants were asked to fill in demographic information and were thanked for their participation. The experiment used a 2 × 2 × 2 mixed design. The within-subject factors were the logical validity of the statement (correct vs. incorrect) and the congruency between the event or conjunction presented as most likely and its level of representativeness (congruent vs. incongruent). The between-subject factor was the time condition (free time vs. limited time).

Results

Manipulation checks Participants in the free time condition spent more time completing the questionnaire and took more time to evaluate conclusions than those in the limited time condition (see Table 2). Independent samples t-tests confirmed observed differences were statistically significant both for the total time spent on the questionnaire and the response time; $t(88) = 3.18, p = .001$, one-tailed, and $t(88) = 5.74, p < .001$, one-tailed, respectively.

Table 2: Mean total time spent on the on-line questionnaire and mean taken to evaluate conclusions as a function of time condition.

	Free time	Limited time
	$M(SD)$	$M(SD)$
Total time (min)	00:11 (00:05)	00:08 (00:04)
Response time (sec)	19.52 (11.58)	09.04 (01.68)

Acceptance rates As the top panel of Figure 2 illustrates, acceptance rates were higher for statements that were congruent with the stereotype activated by the description than for those that were incongruent; $M = 69.25$ ($SD = 24.31$) vs. $M = 28.60$ ($SD = 23.18$), respectively. Individuals, however, were also sensitive to the logical validity of the conclusion presented as they accepted correct statements more often than incorrect ones; $M = 59.26$ ($SD = 26.27$) vs. $M = 38.59$ ($SD = 21.34$), respectively. A 2 (validity) × 2 (congruence) × 2 (time condition) mixed analysis of variance was conducted on these data. The main effect of congruence was significant, $F(1,88) = 136.04, p < .001, \eta_p^2 = .61$, as was the main effect of logical validity, $F(1,88) = 34.12, p < .001, \eta_p^2 = .28$, but the main effect of time condition was not, $F(1,88) = 1.93, p > .05$. None of the interactions were significant, largest non-reliable $F(1,88) = 2.83$ for the interaction between validity and congruence.[3].

Table 3: Mean values of the logic and representativeness index as a function of the response time condition.

	Free time	Limited time
	$M(SD)$	$M(SD)$
logic index	1.04 (1.99)	1.44 (2.03)
representativeness index	2.76 (1.98)	2.12 (1.97)

Finally, following Evans and Curtis-Holmes (2005), two indices were computed for each participant. A *logic* index (CC + CI - IC - II) measured the difference between acceptance of correct and incorrect conclusions and ranged from +6 (perfectly logical) to -6 (absolutely illogical). A *representativeness* index (CC + IC - CI - II) measured the difference between acceptance of conclusions that were congruent or incongruent with the stereotype activated. This index ranged from +6 (always heuristic) to -6 (absolutely not heuristic). Table 2 summarises the descriptive results for these indices. Contrary to dual-systems predictions, participants' responses were *not* less logical under time pressure; $t(88) = -0.94$, $p = .35$. Moreover, and also contrary to theoretical predictions, under time pressure, responses were not more influenced by representativeness assessments; $t(88) = 1.51$, $p = .14$. Finally, whereas acceptance rates were more influenced by representativeness than by logic under free time ($t(48) = -3.59, p < .005$), this was no longer true under time pressure; $t(40) = -1.40, p = .17$.

Response times Average response times for accepting each type of conclusion presented (CC, CI, IC, II) were screened for multivariate outliers using the *Mahalanobis* distance measure (D^2). No outliers were identified (Smallest $D^2 = .54, p = .03$) and all participants' response times were retained for the analyses. The acceptance time data were analysed with a $2 \times 2 \times 2$ mixed ANOVA. The main effect of congruence was significant, $F(1,88) = 6.00$, $p < .02$, $\eta_p^2 = .06$ as well as the main effect of time condition, $F(1,88) = 11.91$, $p < .002$, $\eta_p^2 = .12$, while the main effect of validity was not, $F < 1$. Of greater interest was the significant interaction between validity and congruence, $F(1,88) = 72.64$, $p < .001$, $\eta_p^2 = .45$, accounting for most of the variance. No other interactions were significant, largest non-reliable $F(1,88) = 3.33$ for the interaction between congruence and time condition. As the bottom panel of Figure 2 illustrates, individuals were quicker to accept conclusions that were congruent and correct compared to congruent but incorrect ones; $M = 6.73s$ ($SD = 5.15$) vs. $M = 15.57s$ ($SD = 14.21$), respectively. By contrast, they took longer to accept conclusions that were incongruent but correct compared to incongruent and incorrect ones; $M = 12.35s (SD = 12.78)$ vs. $M = 4.31s$ ($SD = 8.58$), respectively. This pattern of results remain unchanged by time pressure. Individuals responded more quickly when pressured by time; $M = 11.94s$ ($SD = 8.17$) vs. $M = 7.54s$ ($SD = 8.93$), respectively.

[3]The careful reader may notice that participants with limited time seemed more likely to accept correct but incongruent statements than those with free time; $M = 49.59, SD = 35.84$ vs. $M = 33.33, SD = 35.83$. Although potentially meaningful, this difference is not statistically significant if we correct for familywise error; $t(85) = -2.11$, $p = .04 > .0125(.05 \div 4)$.

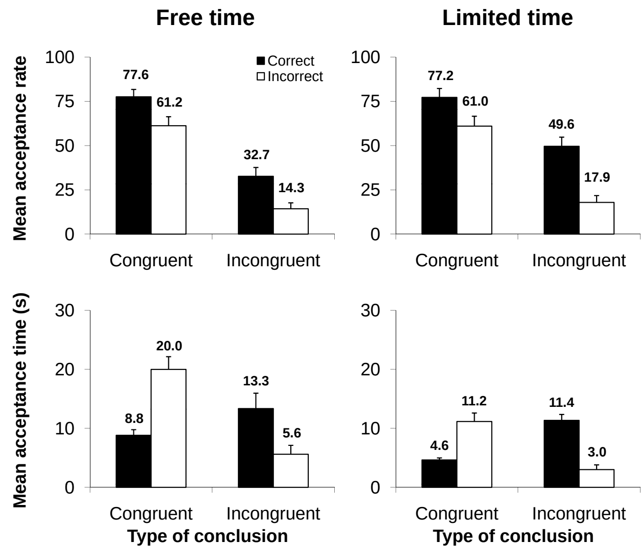

Figure 2: Mean acceptance rate and acceptance time as a function of conclusion validity, congruency with the stereotype activated, and response time condition

Discussion

According to the *lax-monitoring* view of cognitive functioning, conjunction fallacies occur because people solely rely on representativeness attributes to make their judgments. Moreover, those judgments are assumed to be hasty and individuals are assumed to be reluctant to 'think hard' (Kahneman & Frederick, 2002; Kahneman, 2003). The present experiment provided the first empirical test of this claim. Results confirmed that judgments were heavily influenced by considerations of representativeness as statements presenting the most representative outcome as the most probable were more often endorsed than statements presenting an unrepresentative outcome as most probable. However, those results also pointed to the fact that individuals are sensitive to the logic of sets which underpins the conjunction rule since statements presenting a single event (independently of its representativeness) as more probable than a conjunction were more often endorsed than statements presenting the conjunction of events as more probable. This result is at odds with predictions of the *lax-monitoring* view and, instead, lends support to the *conflict-monitoring* view of cognitive functioning (De Neys & Glumicic, 2008). They reveal that, even when thinking about probabilities—as opposed to easier formats such as frequencies (Hertwig et al., 2008)—individuals' cognitive processing is not solely based on heuristic considerations. They *do* appear to think hard when heuristic and logic responses are in conflict. In fact, they spent twice as long thinking about their answer in such instances.

The present study also allowed testing an assumption shared by both of the above-mentioned views: namely, that attribute substitution is a rapid and automatic response originating from the intuitive mode of cognitive functioning. Reliance on heuristic processing was therefore expected to increase under time pressure while logical responding was expected to decrease. Results did not confirm these predictions. In fact, when people did not have the leisure to think as long as they wished, they tended to be *less* influenced by representativeness considerations although time pressure did not affect their sensitivity to the logic of sets. If anything, it lessened the influence of representativeness considerations on judgments compared to that of logical considerations. Even when pressured by time, individuals spent longer responding to items which where logically correct but incongruent with representativeness considerations or incorrect but congruent with such considerations. If replicated, such preliminary results would bear important consequences for our understanding of the type of processing underpinning attribute-substitution. Future research could seek to triangulate those results by burdening executive resources with an attention-demanding secondary task (e.g., see De Neys, 2006).

One possible explanation for these finding is that sensitivity to logical considerations arises from rapid rule-based processing (Sloman, 1996). Accordingly, the logical response would come first and depend solely on the formal structure of the statement, and not on its the meaning. The data collected in this experiment, however, offered no support for this explanation: under time pressure, the more valid responses individuals gave, the longer they thought about each answer they gave; $r = .83, p < .001$, suggesting they detected the conflict between representativeness and logic and therefore processed content. This relationship was still positive under free time, although it was only marginally significant; $r = .27, p = .06$.

Conclusions

This study presented a novel methodology for studying the cognitive underpinnings of the conjunction fallacy, which allowed a more precise determination of whether participants actually detect the conflict between what considerations of representativeness dictate and what the logic of sets entails. This in turn allowed demonstrating that intuitions regarding the conjunction rule of probabilities are not solely influenced by heuristic processing and that people think longest when there is a conflict between the heuristic-based response and the logic-based response. Moreover, contrary to what is commonly assumed, representativeness assessments were not rapid and automatic. In fact, it is when people had time to think that they were most influenced by representativeness. These results apply to intuitions arising from subjective probability considerations. Future research could extend on these results to see if they hold when judgments are frequency-based since frequency formats foster higher rates of logical responding (Hertwig et al., 2008).

Acknowledgments

I thank students for their contribution to scenario design and recruitment, Frédéric Vallée-Tourangeau for his helpful comments on an earlier version of this paper and Jean-François Bonnefon for helpful discussions of this work. This research was supported by grant ANR-07-JCJC-0065-01 from the Agence Nationale de la Recherche.

References

De Neys, W. (2006). Automatic-heuristic and executive-analytic processing during reasoning: Chronometric and dual-task considerations. *The Quarterly Journal of Experimental Psychology*, *59*, 1070-1100.

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*, 1248-1299.

Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*, 255-78.

Evans, J. S. B. T., Barton, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, *11*, 295-306.

Evans, J. S. B. T., & Curtis-Holmes, J. (2005). Rapid responding increases belief-bias: Evidence for the dual-process theory of reasoning. *Thinking & Reasoning*, *11*, 382-389.

Hertwig, R., Benz, B., & Krauss, S. (2008). The conjunction fallacy and the many meaning of *and*. *Cognition*, *108*, 740-753.

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*, 697-720.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (p. 49-81). Cambridge, UK: Cambridge University Press.

Lakoff, G., & Núñez, R. E. (2000). *Where mathematics comes from: How the embodied mind brings mathematics into being.* New York: Basic Books.

Politzer, G., & Noveck, I. A. (1991). Are conjunction rule violations the result of conversational rule violations? *Journal of psycholinguistic research*, *20*, 83-103.

Saxe, G. (1988). The mathematics of child street vendors. *Child Development*, *59*, 1415-1425.

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

Tentori, K., Bonini, N., & Osherson, D. (2004). The conjunction fallacy: a misunderstanding about conjunction? *Cognitive Science*, *28*, 467-477.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review.*, *90*, 293-315.