

Conflict Sensitivity and the Conjunction Fallacy: Eye-tracking Evidence for Logical Intuitions in Conjunction Probability Judgments

Jenny Faure-Bloom, Gaëlle Vallée-Tourangeau, and Sabira Mannan

Department of Psychology, Kingston University
Kingston-upon-Thames UNITED KINGDOM KT1 2EE
j.faure/g.vallee-tourangeau/s.mannan@kingston.ac.uk

Abstract

Recent evidence shows that, contrary to what is commonly assumed, people who are pressured to think fast are also less likely to provide a heuristic judgment when heuristic and logical considerations point to conflicting answers in a conjunction fallacy task (Vallée-Tourangeau & Faure-Bloom, under review). The present study explores this finding using an eye-tracking methodology. Eye movements from 41 participants were recorded while they read a thumbnail description and made a judgment on a statement comparing the probability of a single-event and that of a conjunctive event. Results showed participants focused more on the comparative probability statement under logico-heuristic conflict while they focused more the task description in the absence of conflict. Additionally, longer judgment latencies predicted higher rates of heuristic responding, which contradicts the original dual-process assumption that heuristic thinking in conjunction fallacy tasks is fast.

Keywords: Conjunction fallacy, heuristics and biases, intuition, dual-processes, eye-tracking.

Introduction

People typically use heuristic thinking to make many of their day-to-day judgments and decisions. For example, judging the quality of a good based on its price, or judging a job candidate based on her level of assertiveness. In fact, most of the “educated guesses” we make are founded on heuristics. It would be very impractical to write exhaustive lists or use decision trees every time we needed to make a judgment. Heuristic thinking is often assumed to involve fast decision strategies that save time and effort (Kahneman, 2011). It is considered to be an efficient and reliable process in most instances (i.e., when the heuristic response is congruent with logical principles). However, when heuristic responses conflict with laws of logic and probability, heuristic thought has been known to lead to erroneous, so-called “biased” judgments.

A current debate in the literature asks the question: are people aware when their heuristic response conflicts with logical considerations? Default-interventionist models assume that heuristic processes are always activated first and that deliberate processes are activated only if

necessary to intervene, correct or support heuristic reasoning (Glöckner & Witteman, 2010). Erroneous judgments are assumed to result from the failure to detect the conflict between heuristic and logical considerations; hence heuristic responses are readily endorsed without scrutiny from the deliberative system (Kahneman, 2011). Parallel-competitive models argue that people are able to detect the conflict, but fail to inhibit the enticing heuristic response and are unable to replace it with a more deliberate logical response (Reyna & Brainerd, 2008). Both of these models assume that logical answers emerge from slow, deliberative, and rule-based processes whereas heuristic answers arise from faster, intuitive, and associative processes. It is assumed that the “fast thinkers” who rely on heuristic thinking by default, are also prone to making biased judgments and committing fallacies (Kahneman, 2011); however, few studies have empirically put this claim to the test. For example, in the original conjunction fallacy task (the Linda task, see next section), it is not possible to assess whether people’s judgments are solely informed by heuristic considerations or whether they are also sensitive to logic considerations even if their final answer is congruent with a heuristic assessment.

The Conjunction Fallacy

The conjunction rule of probability states that a conjunction cannot be more probable than one of its constituents. This is logically sound because conjunctions have an extensional nature, which means that the probabilities associated with the conjunction of events A&B are included in the probabilities associated with the event B. In other words, the likelihood of two events occurring simultaneously is always less than either one occurring alone. The representativeness heuristic is defined as a cognitive tool that uses past knowledge (e.g., stereotypes) to estimate probabilities and representativeness is an evaluation of the degree to which something corresponds to a stereotype or prototype (Tversky & Kahneman, 1983). Although this heuristic is fast, and at times efficient, it can lead to erroneous judgments, also known as conjunction fallacies. This was first showcased in the Linda task (Tversky & Kahneman, 1983), where

students were provided with a description of Linda designed to be highly representative of an active feminist. They then had to rank the likelihood of her being a) a feminist (F), b) a bank teller (B), and c) both a feminist and a bank teller (F&B). Over 80% of the participants ranked the conjunction as more likely than the unrepresentative constituent (i.e., $F\&B > B$). This is an error of judgment because the probability of the conjunction (F&B) can never be more than the probability of its constituents (either F or B); at best, they can equal one another.

Fast Logical Intuitions. Heuristic thinking has traditionally been assimilated with (i) intuitive processes and (ii) biased judgements: people provide erroneous judgements because their intuition fails to detect the conflict between logic and heuristic considerations which is only available to deeper, more effortful, more deliberative cognitive processing (e.g., see Kahneman, 2011). By contrast, a recent alternative theoretical position calls for distinguishing heuristic-biased thinking from intuitive processes. Namely, heuristic thinking may lead to biased judgements but the empirical evidence suggests that intuitive processes are not necessarily biased; people can be fast and logical and they can intuitively and implicitly detect the conflict between logical and heuristic considerations (e.g., De Neys, 2012; Villejoubert, 2009, 2011).

Thus, using a novel approach designed to disentangle logical and heuristic considerations, Vallée-Tourangeau and Faure-Bloom (under review) recently showed that people who were put under severe time pressure to provide a judgement about a conjunctive probability statement were almost six times more likely to provide a logical answer. This finding suggests that fast intuitive answers are not necessarily biased while heuristic and biased answers are not necessarily fast. Still, when logical and heuristic considerations were in conflict but participants were not pressured to make fast judgements, those who provided logical answers also took longer to respond. What is not clear is whether those longer response latencies were cued by heuristic considerations or logical considerations. Previous gaze and eye-tracking studies showed that longer latencies are associated with longer inspection of normatively critical problem information in syllogism tasks (Ball, Philips, Wade & Quayle, 2006) and in base-rate neglect tasks (De Neys & Glumicic, 2008) but the source of conflict in conjunction fallacy tasks remains to be identified.

The Present Experiment

The experiment reported here used a more fined-grained measure of conflict processing than response latencies, namely eye movements. Eye-tracking methodology has been proposed as the method of choice for tracing intuitive and deliberative processes (Glöckner & Herbold, 2011). Recent research suggests that automatic,

superficial levels of processing have shorter fixation latencies, while deeper processing and a deliberate consideration of information is related with longer fixations (Glöckner & Herbold, 2011). Furthermore, eye tracking permits the detection of conflict in processing as eye movements are known to be disrupted when readers' general knowledge conflicts with the text they are given to read (e.g., Duffy & Keir, 2004).

The aim of this experiment was to shed light on the processes involved in judging conjunctive probabilities, focussing in particular on the question of conflict detection between heuristic and logical considerations. It was hypothesized that eye movements would differ under conflict and non-conflict judgment settings. Specifically, it was expected that people would have longer total dwell times, and would spend longer re-fixating and the stereotypical description in non-conflict trials as a mean to confirm their initial judgment. By contrast, it was expected people would spend longer re-fixating the statement in conflict trials compared to non-conflict trials, as this statement conflicted with the mental model they had built after reading the description. In this task, the statement is the normatively critical problem. The statement controls the logicity, or illogicity, of the trial.

Method

Participants

Forty-three psychology students from Kingston University were recruited for this experiment. Two students were familiar with the task and therefore were removed from subsequent analyses. Of the remaining 41 participants (6 men and 35 women; mean age = 27.9 years, $SD = 3.48$), 29% were postgraduates and 71% were undergraduates. All the participants had background knowledge of psychology. The majority of the participants were enrolled in psychology courses (66%); however, some had dual majors which included criminology (7%), business (7%), human biology (5%), sociology (5%), journalism (2%), human rights (2%), English (2%), and creative writing (2%). The experiment was conducted in English; 61% of the participants were native English speakers. Participation was voluntary and anonymous, and the study took 45 minutes on average to complete. Participants were either paid £6 or received course credits for their time and participation.

Materials

Materials were adapted from those used in Vallée-Tourangeau and Faure-Bloom (under review). Participants were presented with 16 trials of a "Comparative Conjunction Probability Judgment" (CCPJ) task.

Sixteen Comparative Conjunction Probability Judgment (CCPJ) tasks were programmed in Matlab and

presented to each participant on a 21-inch CRT display screen. Participants' eye movements were recorded using a head mounted video-based Eyelink II eye tracker with a spatial resolution of 0.5 degrees and temporal resolution of 2ms while they assessed the validity of the statements presented.

Each CCPJ trial consisted of a thumbnail description and a statement sentence. The descriptions were short paragraphs that outlined the personality traits, hobbies and vocations of fictitious people. These portraits were intended to be congruent with a stereotype (e.g., Linda was presented as single, outspoken young woman who studied philosophy at university and who is deeply concerned with issues of discrimination and social justice to evoke the stereotype of feminists). Each description was followed by a unique statement; one sentence that contained both a single clause and a conjunction clause. Statements always featured a category that corresponded to the stereotype suggested by the thumbnail description (e.g., feminist in the Linda task) and an atypical category chosen to be at odds with the stereotype (e.g., bank cashier; the word "teller" was replaced with "cashier" as this is the term commonly used in the UK).

Representativeness was determined by the short description presented before the statement. Each of the 16 descriptions created were piloted to ensure that the person presented appeared both as a highly representative member of a stereotypical category and a highly unrepresentative member of an atypical category, without openly mentioning that this individual belonged to either category. Representative statements presented the stereotypical category as more likely than the atypical category and logical statements presented the single clause as more likely than the conjunction clause. Two main types of statements were used: conflict statements and no conflict statements (see Table 1 presenting all statements with the Linda task for illustration purposes; this scenario was only used once in the actual trials). Conflict was created by presenting statements where the conjunction of the stereotypical category and the atypical category was said to be more likely than the atypical category (a representative but illogical statement) or by presenting the atypical category as more likely than the conjunction of the atypical and the stereotypical categories (a logical but unrepresentative statement). Conversely, conflict was absent in statements where the stereotypical category was said to be more likely than the conjunction of the stereotypical and atypical categories (a representative and logical statement) or in statements where the conjunction of the atypical category and the stereotypical one was said to be more likely than the stereotypical one (an unrepresentative and illogical statement).

Conflict statements	
R/I	Lynn is more likely to be a bank cashier and an active feminist than she is to be a bank cashier.
U/L	Lynn is more likely to be a bank cashier than she is to be a bank cashier and an active feminist.
No conflict statements	
R&L	Lynn is more likely to be an active feminist than she is to be a bank cashier and an active feminist.
U&I	Lynn is more likely to be a bank cashier and an active feminist than she is to be an active feminist.

Table 1: Illustration of the four types of statements used to study the conjunction fallacy in CCPJ tasks.

Note. R/I: Representative but Illogical; U/L: Unrepresentative but Logical; R&L: Representative and Logical; U&I: Unrepresentative and Illogical.

On each trial, participants were required to either accept the statement as correct, or reject it as incorrect. Participants' answers were heuristic if they accepted representative statements as correct and rejected unrepresentative statements as incorrect. Conversely, answers were logical if they accepted logical statements as correct and rejected illogical statements as incorrect. The CCPJ tasks thus affords a new insight into the cognitive underpinnings of people's probability judgments under conflict as the answer given could suggest what type of process was employed to produce the final judgment (i.e., either a heuristic or logical one).

To ensure the quality of the eye-tracking data, we conducted manipulation checks on the CCPJ tasks. The number of words was counted for each thumbnail description and for each statement sentence, and an average word count was computed for each description and statement type (i.e., one value for each of: R&L, R/I, U/L, U&I). The mean number of words for the descriptions and statements were each subjected to a repeated-measures analysis of variance (ANOVA) with four levels (statement type). Results confirmed there was no significant difference in word count between the 4 types of description vignettes ($M_{R&L} = 37.50$, $SD_{R&L} = 8.23$, $M_{R/I} = 34.75$, $SD_{R/I} = 5.32$, $M_{U/L} = 38.75$, $SD_{U/L} = 5.62$, $M_{U&I} = 39.50$, $SD_{U&I} = 2.38$), $F < 1$; and no significant difference for the 16 statements ($M_{R&L} = 18.75$, $SD_{R&L} = .96$, $M_{R/I} = 19.50$, $SD_{R/I} = 1.29$, $M_{U/L} = 19.00$, $SD_{U/L} = .82$, $M_{U&I} = 19.00$, $SD_{U&I} = 1.41$), $F < 1$. Thus, any differences found in reading latencies or eye-movements could be assumed to result from our experimental manipulations and not from lack of experimental control in our design.

Design and Procedure

The experiment took place in a psychology laboratory on the Kingston University campus. Upon giving their

informed consent, all participants completed the 16 CCPJ trials.

CCPJ Task. Participants were sat in front of a 21-inch CRT display screen while the experimenter set up the Eyelink II. The eye-tracker was placed and secured on the head of the participants, and the video cameras were positioned to get a clear and accurate reading of the pupils. The experimenter then performed the calibration and instructed the participants to keep their head as still as possible for the remainder of the task. The first slide of the CCPJ task presented the instructions to the participants: “In this experiment, you will be presented with a series of scenarios describing a person followed by a statement. Your task is to read these descriptions and then indicate if you think the accompanying statement is true or false”. At this point the experimenter explained that if they wanted to answer true, and accept the statement, they must click on the left mouse button, and if their answer was false, and they wanted to reject the statement, they must click on the right mouse button. “YES” was positioned on the bottom left of the screen and coincided with the left mouse button, while “NO” was situated bottom right and coincided with the right mouse button. The next screen explained to the participants that they would be given three example trials before beginning the experiment. This allowed them to become familiar with the task and understand what was expected of them before beginning the experimental problems. They were prompted to take as much time as they needed to read the descriptions, and to consider the information carefully before reaching a decision. After completing three examples, they completed the 16 CCPJ problems in a random order while their responses and time latencies were recorded for each CCPJ problem as well as eye movements in two areas of interest (AoI). The first AoI was the description vignette, and the second was the statement sentence (see Fig. 1 for an example of a CCPJ trial with the AoI definitions). More specifically, eye movements were measured through number of refixations and the total dwell time for each AoI.

C-Span Task. Each participant also completed a computation span measure adapted from Salthouse and Babcock (1991) designed to measure working memory span. The task involved solving simple arithmetic problems whilst remembering the second digit of each equation for a recall test after each series of problems. Participants viewed the task stimuli on a computer monitor and recalled their answers verbally whilst the experimenter recorded and scored the participants’ answers, both for the arithmetic problems and the recall test, on an answer sheet. The arithmetic problems were presented as a video in Windows Media Player.

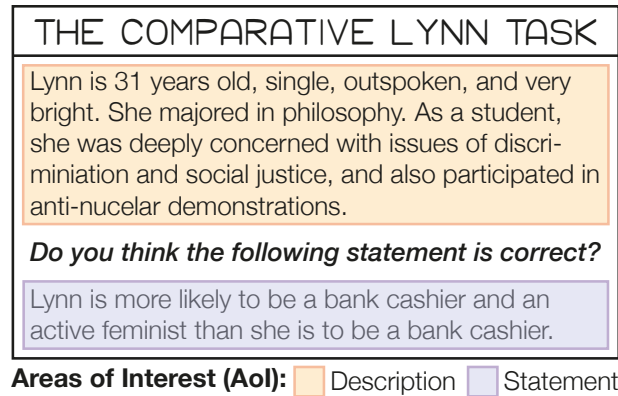


Figure 1: Design and layout of a Comparative Conjunction Probability Judgment problem.

Results

Behavioural Data

To examine patterns of heuristic responding, a heuristic score was computed for each statement type by calculating the proportion of time participants either accepted statements that were congruent with heuristic considerations (R&L and R/I statements) or rejected statements that were incongruent with such considerations (U/L and U&I statements). The heuristic score ranged from 0 (never followed the heuristic consideration) to 1 (always followed the heuristic consideration). As anticipated by the logical intuitionist account, statements for which heuristic considerations conflicted with logical ones led to lower rates of heuristic responding than non-conflict statements, $M_{conflict} = .83$, $SD = .22$ vs. $M_{no_conflict} = .87$, $SD = .16$. This difference, however, failed to reach a statistically significant threshold, $t(40) = -1.04$, $p = .15$, one-tailed. Thus, it was not possible to conclude that participants were sensitive to the conflict between heuristic and logical considerations on the sole basis of the behavioural data.

Eye Tracking Data

Eye movement patterns were examined to assess whether participants were sensitive to the conflict between heuristic considerations (i.e., representativeness) and logical considerations. The number of refixations (i.e., number of times participants revisited an AoI with at least one fixation outside the AoI) and the total dwell times (i.e., total amount of fixation time within an AoI) were examined across both of the conflict conditions (no conflict vs. conflict) and the two AoIs (description vs. statement).

Refixations. Participants redirected their eyes more often towards the description in the absence of conflict, $M_{no_conflict} = 3.74$, $SD = 1.46$, $M_{conflict} = 3.41$, $SD = 1.27$. However, in conflict trials, participants refixated more often on the statement, $M_{no_conflict} = 3.20$, $SD = 1.52$, $M_{conflict}$

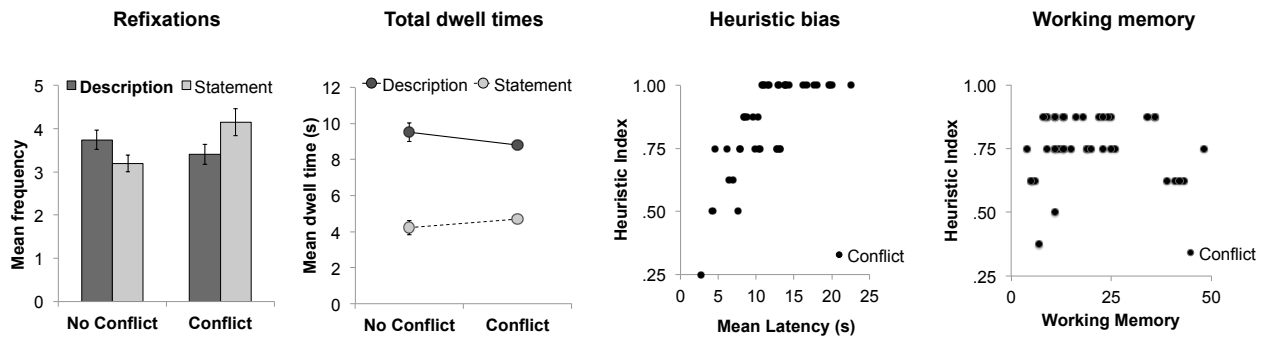


Figure 2: Mean number of refixations as a function of AoI (description vs. statement) and conflict (absent vs. present; left panel); mean dwell times as a function of AoI and conflict (middle left panel); rate of heuristic responding as a function of judgment latency (seconds; middle right panel); rate of heuristic responding as a function of working memory (C-Span score; right panel).

= 4.15, $SD = 2.00$. A 2 (description vs. statement) \times 2 (conflict vs. no conflict) ANOVA confirmed significant differences between the number of refixations as a function of the conflict conditions, $F(1,40) = 5.54, p < .05, \eta_p^2 = .12$. It also revealed a significant interaction between the AoIs and the conflict conditions, $F(1,40) = 32.96, p < .001, \eta_p^2 = .45$ (see Figure 2: Left Panel).

Dwell times. A 2 (description vs. statement) \times 2 (conflict vs. no conflict) ANOVA revealed that, on average, participants spent more time looking at the descriptions ($M_{no_conflict} = 9.51, SD = 3.26, M_{conflict} = 8.79, SD = 2.53$) compared to the statements ($M_{no_conflict} = 4.21, SD = 1.55, M_{conflict} = 4.71, SD = 1.50, F(1,40) = 209.27, p < .001, \eta_p^2 = .84$), regardless of the conflict condition. This is not surprising as the descriptions were short paragraphs containing 3-4 sentences, while the statements were only one sentence in length. More importantly however, the analysis revealed a significant interaction between AoIs and conflict condition for mean dwell times, $F(1,40) = 17.01, p < .001, \eta_p^2 = .30$ (see Figure 2: Middle Left Panel).

Altogether, the eye-tracking data revealed that, when logic conflicted with representativeness (i.e., conflict statements, see Table 1), participants spent more time looking at the statements compared to descriptions and they revisited statements more often. The reverse was true in the absence of conflict. These differences in eye-movement patterns clearly show that participants were sensitive to the conflict between representativeness and logic even if their final behaviour appeared to be solely influenced by heuristic considerations.

Judgment Latencies

Next, response latencies were analyzed to empirically test whether “fast thinkers” who are assumed to rely on heuristic considerations by default (e.g., Kahneman, 2011), were also prone to making biased judgments and committing conjunction fallacies. A heuristic index was computed for each participant for the conflict trials. This

entailed subtracting the unrepresentative answers from the representative answers (e.g., $RI1 + RI2 + RI3 + RI4 - UL1 - UL2 - UL3 - UL4$).

A regression analysis showed that longer judgment latencies predicted higher rates of heuristic responding in conflict statements, $R^2 = .52, F(1, 39) = 41.11, p < .001, \beta = .72, t(39) = 6.44, p < .001$ (see Figure 2: Middle Right Panel). These findings suggest that people who responded more logically to the conflict trials also responded significantly faster; heuristic responding in these trials were *not* fast.

Homogeneity of sample

Working memory scores were analysed to ascertain whether individual measures caused some participants to perform better (i.e., more logically) than others on the CCPJ trials.

A regression analysis showed lower working memory scores did not predict higher heuristic responding, $R^2 = .00, F(1, 39) = .01, p = .92, \beta = .02, t(39) = .11, p = .92$. These findings show that working memory played no role in participants logicity when making conjunction judgements (see Figure 2: Right Panel).

Discussion

Our aim with this experiment was to shed light on the processes involved in generating conjunction probability judgments. Specifically, we focussed on conflict detection between heuristic and logical considerations, and we provided empirical data to test the widely accepted assumption that heuristic thinking is fast and automatic while logical thinking is slower and more deliberate.

The evidence presented in this study goes against the default-interventionist interpretation since participants’ eye movements suggest they easily detected the conflict between logical and heuristic considerations. In the presence of conflict, participants looked for longer amounts of time at probability statements compared to

descriptions and they revisited statements more often. The reverse was true in the absence of conflict. Evidence also goes against the parallel-competitive view. Although this interpretation allows for conflict detection, it also assumes that in order to answer logically one must inhibit the persuasive heuristic answer. The data indicated that faster judgments were also more logical, which suggests that logical answers come first and are intuitive, but are overridden by the compelling heuristic answer when people take time to deliberate. When this happens, logical responding requires even more deliberation and effort to be reinstated. These findings support the notion of *logical intuitions* (De Neys, 2012; Villejoubert, 2009, 2011). This idea posits that people have an implicit knowledge of the laws of probability, and as a result, logical responses are an intuitive, effortless process. In other words, logical considerations can be fast and automatic. The bias, previously believed to exist in fast automatic considerations, in fact could exist in slow deliberate considerations.

The question remains, if people are able to detect conflict in conjunction probability judgments and they also have logical intuitions, why then do the majority of people still commit conjunction fallacies? A possible explanation for this could be that although people are able to generate fast logical intuitions about conjunctive probability judgments, when there is conflict they employ deliberate thinking and this is when the logical intuitions are overcome by slow heuristic deliberations. Nonetheless, despite the convincing evidence for conflict detection presented in this study, it remains true that the heuristic answer stayed prevalent even under conflict. Future research may also consider the factors that might sway participants towards the logical answer, such as an implicit learning task or a priming task. Future research may also consider the implications of these results for recent proposals attempting to cognitive biases as a consequence of random noise in otherwise rational judgements (see Costello, 2009).

Acknowledgements

Support from the Leverhulme Trust (grant RPG-228) is gratefully acknowledged. We thank Frédéric Vallée-Tourangeau for comments on earlier drafts of this article and James Lockwood for his assistance with data collection on this project.

References

- Ball, L. J., Phillips, P., Wade, C. N., & Quayle, J. D. (2006). Effects of Belief and Logic on Syllogistic Reasoning. *Experimental Psychology*, *53*, 77–86. doi:10.1027/1618-3169.53.1.77
- Costello, F.J. (2009). How probability theory explains the conjunction fallacy. *Journal of Behavioural Decision Making*, *22*, 213-234. doi: 10/1037/a0037010
- De Neys, W. (2012). Bias and Conflict: A Case for Logical Intuitions. *Perspectives on Psychological Science*, *7*, 28–38. doi:10.1177/1745691611429354
- De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition*, *106*, 1248–1299. doi:10.1016/j.cognition.2007.06.002
- Duffy, S. A., & Keir, J. A. (2004). Violating stereotypes: Eye-movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition*, *32*, 551-559.
- Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-Process Theories of Higher Cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*, 223-241. Doi: 10.1177/1745691612460685.
- Glöckner, A., & Herbold, A. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioural Decision Making*, *24*, 71-98.
- Glöckner, A., & Witteman, C. (2010). Beyond dual-process models: A categorisation of processes underlying intuitive judgment and decision making. *Thinking & Reasoning*, *16*, 1-25.
- Kahneman, D. (2011). *Thinking, fast and slow*. London: Penguin Books.
- Reyna, V.F., & Brainerd, C. J. (2008). Numeracy, ratio bias, and denominator neglect in judgments of risk and probability. *Learning and Individual Differences*, *18*, 89–107. Doi:10.1016/j.lindif.2007.03.011.
- Salthouse, T. A., & Babcock, R. L. (1991). Decomposing adult age differences in working memory. *Developmental Psychology*, *27*, 763-776. Doi: 10.1037/0012-1649.27.5.763.
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, *90*, 293–315. Doi: 10.1037/0033-295X.90.4.293.
- Vallée-Tourangeau, G., & Faure-Bloom, J. (2014). Intuitively logical, even under time pressure: rethinking the role of intuitive and deliberative processes when making social judgments under uncertainty. Manuscript under revision.
- Villejoubert, G. (2009). Are representativeness judgments automatic and rapid? The effect of time pressure on the conjunction fallacy. In N. Taatgen, H. van Rijn, L. Schomaker and J. Nerbonne (Eds.), *Proceedings of the 31th Annual Conference of the Cognitive Science Society* (pp. 2980-2985). Austin, TX: Cognitive Science Society.
- Villejoubert, G. (2011). Rethinking the Role of Intuition in the Conjunction Fallacy. In W. Brun, G. Keren, G. Kirkeboen, & H. Montgomery (Eds.), *Perspectives on Thinking, Judging, and Decision Making* (pp. 109–119). Oslo: Universitetsforlaget.