Feelings of error in reasoning – in search of a phenomenon

Amelia Gangemi[1,3,*] Sacha Bourgeois-Gironde[2] and Francesco Mancini [3]

[1]Department of Cognitive Science

University of Messina

Via Concezione, 6-8, 90128-Messina, Italy

[2] Institut Jean-Nicod, CNRS-EHESS-ENS

29, rue d'Ulm, 75005

Paris, France.

[3] Scuola di Specializzazione in Psicoterapia Cognitiva

Viale Castro Pretorio,116, 00185-Rome, Italy

[*]Corresponding author. Tel.:+39-06-3216801; fax:+39-06-3222743. E-mail address: gangemia@unime.it

Feelings of error in reasoning – in search of a phenomenon

Abstract

Recent research shows that in reasoning tasks, subjects usually produce an initial intuitive answer, accompanied by a metacognitive experience, which has been called Feeling of Rightness (FOR; Prowse Turner & Thompson, 2009). This paper is aimed at exploring the complimentary experience of Feeling of Error (FOE), that is, the spontaneous, subtle sensation of cognitive uneasiness arising from conflict detection during thinking. We investigate FOE in two studies with the "bat-and-ball" (B&B) reasoning task, in its standard and isomorphic control versions. Study 1 is a generation study, in which participants are asked to generate their own response. Study 2 is an evaluation study, in which participants are asked to choose between two conflicting answers (normative vs. intuitive). In each study, the FOE is measured by the FOE Questionnaire. Results show that the FOE is significantly present in the standard B&B task when participants give a wrong answer, that our questionnaire can measure it, and furthermore, that it is diagnostic of genuine error.

Keywords: Feelings of Error; Cognitive biases; Bat-and-Ball problem; Feelings of Error Questionnaire, Feelings of Rightness.

When faced with cognitive tasks, we are often characterized by the tendency to base our responses on fast intuitive impressions rather than on more deliberative reasoning. Although this intuitive "heuristic" thinking might sometimes be useful (cf. Harvey, Watkins, Mansell & Shafran, 2004), it will often cue biased responses that are in conflict with the normative ones (logical, or probabilistic or decisional) (cf. De Neys, Cromheeke & Osman, 2011; De Neys, 2012). The ease with which these biased responses are generated may create a strong intuition that they are correct, thus not requiring further analysis (Thompson, Evans & Campbell, 2013).

A number of cited research studies have focused on these intuitions, or better, on these sensations of "being right", and currently the interest on these monitoring and controlling processes in reasoning is growing. For example, Thompson and colleagues (e.g. Prowse Turner & Thompson, 2009; Thompson, Prowse Turner & Pennycook, 2011), have shown that in reasoning problems (such as syllogisms and conditionals), the production of an initial, intuitive answer is accompanied by a metacognitive experience, which has been called Feeling of Rightness (FOR). According to the Metacognitive Dual-Processes Theory (Thompson et al., 2011), the FOR (which exists along a continuum) is an affective response, which complements the fast and automatic "heuristic" response (Type 1 processes) to a task. It is assumed to be generated by automatic and largely implicit processes, whose origins are not likely available to conscious processes (see Koriat & Levy-Sadot, 1999). It has been argued (see, Thompson & Morsanyi, 2012), that a strong FOR presumably explains the compellingness of many cognitive illusions, because – as the heuristic output (Type 1 processes) is judged to be reliable – it discourages a deeper and slower processing by the second system of cognitive processing, the analytical one (Type 2 processes) (see also Thompson et al., 2013). This affective response arises from the fluency with which the initial answer is produced (i.e. the level of easiness or efficiency that an item requires to be processed), such that fluently produced answers give rise to a strong FOR. This FOR, in turn determines the extent and probability with which Type 2 processes will be engaged. Because many of the intuitions produced by Type 1

processes are fluent, it is common for them to be accompanied by a strong sense of rightness. By contrast, “more effortful, less efficient processes should produce a weaker FOR” (Thompson, 2009, p. 176), and consequently, an engagement of the analytic system to accurately process the output produced by the heuristic one.

In reasoning, FOR can be compared to similar metacognitive phenomena in the domain of perception and memory. For example, in memory recalling tasks, retrieving an answer from memory is accompanied by a Feeling of Familiarity (FOF), which is the cue that the retrieved item is the right one. Moreover, when recall is used to test episodic memory, people sometimes fail to retrieve previously encoded information, but express a Feeling of Knowing (FOK) (e.g., Koriat, 2000; Efklides, 2006), that they could recognize the information on a later test (e.g. Nelson & Narens, 1980). Analogously to FOR, this family of metacognitive experiences can produce erroneous judgements. For example, people can express high confidence in inaccurate memories, and this is due to properties of the retrieval processes that produce memories, rather than to the contents of memory per se.

All these metacognitive experiences are used primarily to describe situations where individuals believe they possess knowledge (i.e. FOR, see Thompson, 2009), but other knowledge states appear to be based on similar but *negative* metacognitive appraisals (i.e. *feelings of wrongness,* Thompson & Morsanyi, 2012). For example, when individuals conclude that they do not know (cannot even partially retrieve) the answer to a given question, they may refer a “Feeling of Not Knowing” (e.g. Glucksberg & McCloskey, 1981). Moreover, according to Thompson and Morsanyi (2012), negative emotional states would also contribute to judgments of confidence. For example, feelings of disgust affect moral judgments, even when the actions being judged do not cause any harm (e.g., incest between consenting siblings, Haidt 2001).

So far, very few empirical studies have investigated these *feelings of wrongness* in the context of reasoning. Yet, it is quite often that subjects solving problems that give rise to biased systematic

errors or cognitive illusions become conscious that something is wrong with their performance. They may not know what exactly, they may not precisely refer back to a specific moment in their recent performance, but they sense that they did not perform correctly (see, e.g., Piattelli-Palmarini, 1994).

In the present study, we try to expand the data on these spontaneous and implicit metacognitive processes in reasoning, through the exploration of Feeling of Error (FOE) (that can be viewed as the inverse and complementary experience to FOR). FOE is assumed to accompany the intuitive and erroneous answers usually evoked by those cognitive tasks that typically provoke biases or illusions. It might arise from conflict detection during reasoning. Recent empirical work on conflict detection during thinking (De Neys & Glumicic, 2008; De Neys, Cromheeke & Osman, 2011; De Neys, Moyens & Vansteenwegen, 2010), have indeed shown that adults and older adolescents are highly sensitive to the conflict between heuristic responses and traditional normative considerations in classic reasoning problems, at least at an implicit level ("feeling we are biased").

Our goal is both to explore and confirm the FOE phenomenon, in order to reach a deeper understanding of its nature and its possible role in promoting reflection upon and correction of the erroneous response. Like FOR, FOE might play a role in mediating the extent and quality of Type 2 thinking. However, in this paper, we restrict our investigation to the following questions:

- Is FOE actually present while performing certain reasoning tasks that generate systematic errors?
- And, in particular is FOE reliable, in the sense that it arises when people are actually mistaken?
- Finally, might FOE arise from the conflict between intuitive responses and traditional normative considerations in a reasoning task?

To this aim, in two different studies, we use an experimental task known to generate a large amount of errors: the bat-and-ball problem (Frederick, 2005):

A bat and a ball together cost $1.10. The bat costs $1 more than the ball. How much does the ball cost?

We chose this task because it usually yields an intuitive erroneous response, which is 10 cents, but the correct answer is 5 cents. If the ball were to cost 10 cents, the bat would cost $1.10 (i.e., $1 more), According to De Neys , Rossi & Houdé (2013) the explanation for the "10 cents" bias is that people substitute the critical relational "more than" statement by a simpler absolute statement: "The bat costs $1". This task is thus an example of the so-called attribute substitution. Individuals confronted with difficult questions often intuitively answer an easier one instead (see De Neys, Rossi & Houdé, 2013; Kahneman, 2011; Kahneman & Frederick, 2002), usually without being aware of the substitution. Although reasoners do not deliberately reflect upon their response, and do not know what the correct answer is, it is possible that they could detect the substitution process, showing a substitution sensitivity (De Neys et al., 2013)

To be sure that a lower FOE among participants who give the incorrect response does reflect error detection in the tasks, and not simply some sort of general, trait-like response doubt among cognitively poor reasoners, we also used a non-conflict version of the bat-and-ball problem (control problem):

A magazine and a banana together cost €2.90. The magazine costs €2. How much does the banana cost?

Recently presented by De Neys et al. (2013), it does not cue conflicting responses and is very easy to solve because it does not contain the relative statement ("*more than")* that usually triggers the substitution and thus the error. We expected few errors and hence participants should show a very floored FOE level with it.

In Study 1, for both the tasks, unlike most part of the studies on these metacognitive experiences, participants are asked to draw their own conclusions rather than to evaluate given conclusions (generation task). We assume that the FOE would be easily and more strongly generated in case of

incorrect answer, because of the conflict detection during reasoning (intuitive vs. correct answer), resulting diagnostic of actual errors.

In Study 2, in order to further investigate this conflict detection, participants were asked to choose between the two conflicting answers (evaluation task), as in the most part of studies (e.g. Shynkaruk & Thompson, 2006; Thompson et al., 2011). We assume that regardless of whether reasoners give an answer based on intuition or on normative principles, the conflict between responses would be detected. However, FOE would be diagnostic of actual errors, resulting higher in case of intuitive answers.

In both studies, subjects are given the Feeling of Error Questionnaire (FOE-Q, see below). It is a self-report questionnaire, and was developed to examine the FOE elicited by the tasks, specifically, to assess the affective state, in which our participants found themselves in while performing the tasks. This questionnaire has been borrowed from a questionnaire previously used with obsessive-compulsive disorder patients, the 'Not Just Right Experience–Questionnaire' (Coles, Frost, Heimberg & Rhéaume., 2003; Heimberg, Frost & Steketee, 2005). We have adapted it to assess the cognitive uneasiness felt by subjects performing reasoning tasks. Contrary to prior studies, in which these metacognitive feelings were measured with only one question (e.g. "at the time I provided my answer I felt…", followed by a likert scale ranging from 1: (guessing) to 7 (certain I'm right), we have chosen to assess it with a more articulated questionnaire.

Study 1

Method

Participants

A total of 145 undergraduates taking an introductory psychology course at the University of Messina (Italy) participated in return for course credit. Their mean age was 24 years (SD= 6 years); 100 (69%) were females. Participants provided written informed consent.

Material and procedure

Participants are given a booklet with written instructions, and the bat-and-ball (B&B) task, in its standard and non-conflict (control problem) versions. All the problems are printed on separate pages of the booklet. We translated them from English into Italian.

In the case of the bat-and-ball problem, culturally common items are used, a stamp and an envelope, as baseball is widely ignored in Italy:

An envelope and a stamp cost €1.10 together. The stamp costs €1 more than the envelope. How much does the envelope cost?

The correct answer is 5 cents, but the task typically elicits incorrect answers, which is 10 cents.

To make sure that the differential item content of the B&B problem did not affect the findings, the item content and control status of the problem were completely crossed. For half of the sample, we used the envelop/stamp/€1.10 content in the standard version and the magazine/banana/€2.90 content in the control version. For the other half of the sample, the contents of the two presented problems are switched. Moreover, in order to avoid carryover effects from one task to the others, presentation order of the tasks are counterbalanced. Participants are also given a short break among the two problems.

Participants are told to read the problems carefully and to solve them individually. After solving each task, participants are asked to fill in the Feeling of Error-Questionnaire (FOE-Q), assessing the level of cognitive discomfort arising from the just completed task. The FOE-Q begins by presenting the cognitive uneasiness sensation to participants:

While I was solving the task, I had the unpleasant sensation that I was not behaving exactly as I should or the way I would have liked to.

Afterwards, respondents are asked to indicate whether or not they experienced this feeling during the task and the degree of this feeling (from 0: “not at all” to 4: “extremely strong”). Participants are subsequently asked to rate this sensation in five dimensions, by using again a 5-point scale (0 = “not

at all", 4 = "extremely strong"):

- *How intense was this sensation?*
- *How unpleasant was this sensation during the task? How unpleasant is this sensation now?*
- *To what extent was this sensation due to the feeling that you didn't solve the task as you should?*
- *To what extent did this sensation mean there was something wrong or incomplete in the task?*

The FOE-Q total score can range from 0 to 24, with higher scores reflecting stronger FOE.

For the two tasks, each participant's performance in the task is coded as correct/incorrect.

Results

To examine how well the occurrence ratings for the FOE-Q items held together, we used Cronbach's alpha. This revealed that the internal consistency for the scale is high in the current sample (a = 0.86).

In line with previous studies, only 30% of participants managed to solve the standard B&B problem correctly. Incorrect responses were almost exclusively of the "10 cents" type (i.e. 100 out of 102 responses), suggesting that biased participants were not simply making a random guess but were, indeed, engaged in the hypothesised substitution process. As predicted, the control task was solved correctly by 97% of the participants.

As shown in Fig. 1, FOE was significantly higher in the standard versions of the B&B task, when participants performed badly, than in the control version where their performance was correct, $F(1, 101) = 658.46$, $p < .0001$, $\eta^2_p = .87$. (see Fig. 1).

Figure 1 also shows that FOE was significantly lower in reasoners who solved the standard problem correctly($F(1, 143) = 66.29$, $p < .001$, $\eta^2_p = .32$; M= 8.79, SD=4.51), than in those who performed badly in the same task (M=16.52, SD= 4.95; ($t(143)= 8.8$, $p<.0001$), while FOE does not significantly differ in the control problem between those who solved the standard problem correctly (M=2.39, SD=2.7) and incorrectly (M=2.01, SD=3.14; ($t(143)= .68$, $p=.49$) .

Please insert Fig. 1 about here

Study 2

In our first study, results indicate that participants who actually failed in the tasks experienced FOE, to a greater extent, than those who succeeded in them. This shows that FOE is diagnostic of genuine error. According to De Neys and colleagues (2011; 2013) and Thompson & Johnson (2014), we could explain this accuracy as determined by a detection of the conflict between the intuitive answer and the normative one, which impacts on the reasoning process. This conflict sensitivity might lead participants to acknowledge that their wrong response is questionable. In study 2, to further investigate this conflict detection, participants were asked to choose between the two conflicting answers (evaluation task), as in the most part of studies (e.g. Shynkaruk & Thompson, 2006; Thompson et al., 2011; Thompson & Johnson, 2014). We expected that participants would be less accurate in their negative metacognitive judgments when reading the response, because of the sense of familiarity produced by the intuitive answer (see De Neys et al., 2011; De Neys & Bonefon, 2013), which might increase the conflict between the correct and the incorrect answer. According to this, the correct answer could result questionable, too. We expected however that FOE would once again reveal itself diagnostic of actual error, reporting participants high level of it when wrong. We used the same task employed in the earlier study.

Method

Participants

A total of 148 undergraduates taking an introductory psychology course at the University of Messina (Italy) participated in return for course credit. Their mean age was 23.8 years (SD= 5.6 years); 94 (63.5%) were females. Participants provided written informed consent.

Material and procedure

As in the previous study, each participant is given a booklet with written instructions and the Italian version of the two versions of the B&B task. After having read each task, participants are

asked to choose between two conflicting answers. For the standard version of the B&B they are asked to choose between 5 cents (correct answer) vs 10 cents (incorrect but intuitive answer), for the control version of the B&B they can choose between 0.90 (correct answer) vs 0.95 (incorrect answer). The incorrect answer was chosen just because very similar to the correct one.

Participants are asked to read the problems carefully and to solve them individually. Participants are then asked to fill in the Feeling of Error-Questionnaire (FOE-Q).

Also in this study, to be sure that the content of the B&B problem did not affect the findings, the item content and control status of the problem were completely crossed. The presentation order of the two tasks are counterbalanced. For all tasks, each participant's performance in the task is coded as correct/incorrect.

Results

To examine how well the occurrence ratings for the FOE-Q items held together, we used Cronbach's alpha. This revealed that the internal consistency for the scale is high in the current sample (a = 0.85).

Only 34% of participants managed to solve the standard B&B problem correctly. The control task was solved correctly by 97% of the participants.

As shown in Fig. 2, FOE was significantly higher in the standard version of the B&B task when participants performed badly, than in the control version where their performance was correct, $F(1, 96) = 424.02$, $p < .0001$, $\eta^2_p = .81$. (see Fig. 2).

Figure 2 also shows that FOE was significantly lower in reasoners who solved the standard problem correctly($F(1, 146) = 6.55$, $p < .02$, $\eta^2_p = .05$; M= 11.3, SD=4.26), than in those who performed badly in the same task (M=14.18, SD= 5.53; ($t(146)= 3.15$, $p<.01$), while FOE does not significantly differ in the control problem between those who solved the standard problem correctly (M=1.2, SD=2.94) and incorrectly (M=1.35, SD=2.46; ($t(146)= .35$, $p=.73$) .

Please insert Fig.2 about here

Since participants of Study 1 and Study 2 were recruited from the same population, and did not differ in regards to sex (χ^2 (2, 148) = 2.89, p= .23) and age ($t(144) = .08$, $p = .94$), we compared their level of FOE in the two studies, rated with both the correct and incorrect answers to the standard version of the B&B task. A 2 (generation vs. evaluation study) X 2 (correct vs incorrect answer) ANOVA shows an interaction effect between the two factors, such that the difference between correct and incorrect answer in FOE ratings is larger in the generation task, than in the evaluation one, F (1, 289) = 15.5, p< .0001, η^2_p =.05. Participants giving the correct answer in the evaluation study (Study 2) reported a higher FOE (M=11.37, SD= 4.26), than those succeeding in our generation study (Study 1, M=8.79, SD= 4.51; *t* (291) 5.03, p<.0001. Participants giving the incorrect answer in the evaluation study (Study 2) reported a lower FOE (M=14.18, SD= 5.54), than those failing in our generation study (Study 1, M=16.52, SD= 4.95; *t* (291) 3.81, p<.0002).

Discussion

In the context of reasoning, few authors have drawn attention to negative self-appraisals by subjects who are victims of biases or of cognitive illusions. Some exceptions, precisely in the context of the heuristics-and-biases research program, are De Neys and colleagues (e.g. De Neys et al., 2011; 2013; De neys &Bonnefon, 2013), and Piattelli-Palmarini,who wrote, in 1994: “Those subjected to Tversky-Kahneman experiments, which are problems that raise cognitive illusions, are often conscious that something is wrong in their reasoning (…) anyone involved in one of these experiments or tests understands, in the intimacy of his reason, that something is tending to go away. He may not know what, but he knows something is not working” (pp. 40-41).

Our studies were designed to investigate just these feelings of error (FOE) experienced by individuals, when they produce incorrect responses in a reasoning task explicitly designed to elicit those responses. In both studies, we used the bat-and-ball problem. Because of its arithmetical nature, participants can easily recognize non-normative performance as errors.

In general, our data indicate that participants who actually failed in the task experienced FOE, as measured through the Sense of Error Questionnaire (FOE-Q), to a greater extent, than those who succeeded in it. This shows that FOE function as reliable signals when errors are objectively present.

Analogously to the FORs (Thompson et al., 2011; Thompson & Johnson, 2014), we could explain this accuracy as determined by a detection of the conflict between the intuitive answer and the normative one (see De Neys and colleagues, 2011; 2013; De Neys, 2012; Thompson & Johnson, 2014). Although reasoners do not deliberately reflect upon their response, and do not know what the correct answer is, it is possible that they could detect the conflict, showing a conflict sensitivity (De Neys et al., 2013), which impacts on the reasoning process. When individuals are called to produce the response, this impact is stronger in case of intuitive (wrong) answer, than in case of normative (correct) one, leading them to acknowledge that their wrong response is highly questionable. In our generation study, negative metacognitive judgments (FOE) are indeed more accurate, than in our evaluation study. When participants were asked to choose between the two conflicting answers (evaluation task), they were less accurate in their negative metacognitive judgments, probably because of the sense of familiarity produced by the intuitive answer to be evaluated (see De Neys et al., 2011; De Neys & Bonefon, 2013). This familiarity might be responsible for the increase of the conflict between the correct and the incorrect answer. As a result, the correct answer becomes questionable, too.

So what are the cues that give a person the feeling that the solution that s/he has just given is likely to be incorrect (FOE), and what is the role of these feelings? Some authors claim that the experience of errors is normally mediated by some non-conscious monitoring processes, in particular, conflict-monitoring, and error-detection processes (De Neys et al., 2008; 2010; 2011). Conflict is defined as the co-activation of two incompatible responses in one task (Carbonnell & Falkenstein, 2006), while error-detection processes target the specifically erroneous response when a conflict is present. In line with this perspective, we hypothesise that FOE, like noetic feelings (Metcalfe, 2000), are at least

partly conscious, and serve a crossover function: they mediate between implicit-automatic and explicit-analytic processes (Koriat, 2000). Presumably, in the bat-and-ball, the mechanisms of conflict monitoring and error-detection send conscious signals, such as FOE, to the subject when the resolution of conflict or the detection of error is not automatic and requires analytic cognitive strategies. FOE form a conscious experience even though, while experiencing them, participants may not yet have reached the correct answer. A participant has this sensation perhaps precisely because s/he is not yet able to explicitly detect and characterize her/his error (see Piattelli-Palmerini, 1994). In fact, people are not able to identify the exact normative principles that are being violated (in our task arithmetical rules). But, whenever their intuitive answer conflicts with the normative one, people do seem to acknowledge that their response is questionable. The fact that this conflict is affecting their judgments, implies that reasoners implicitly adhere to these normative principles (De Neys et al., 2011; De Neys, 2012; De Neys et al., 2013). Indeed, a characteristic of our substitution problem is that their solution is easily understood by participants, when explained (Frederick, 2005). These experiences thus have a critical role as they should motivate a more analytic approach (e.g. Type 2 thinking). Further studies are thus required to verify whether FOE might play a role in mediating the extent and quality of Type 2 thinking. According to the Metacognitive Reasoning Theory, the initial intuitive (Type 1) response would be accompanied by the FOE (with an increasing of its level), which in turn would determine the extent of analytic engagement (for example rethinking time, see Thompson et al., 2011) as well as the answer change (with a decreasing of its level) (for example, the outcome of analytic engagement). We should thus examine the possible role of FOE in promoting reflection upon and correction of the erroneous response, and thus verify whether the FOE actually triggers the Type 2 thinking when strongly present.

Of course, the present study has a number of limitations, and still leave unsolved a number of questions. First, contrary to our study, previous studies on the complimentary Feelings of Rightness find relatively little correspondence between confidence and accuracy (e.g. Shynkaruk & Thompson,

2006), although Thompson et al. (2011) found a small positive correlation with easier items. A possible reason for this apparent incongruity is that the task used in the current study is generally different from those used by Thompson et al., such as syllogistic or conditional reasoning. Our present data do not allow us to conclude that FOE is predictive of accuracy in a general way. It could be task specific. According to Stanovich and West (2008), tasks could differ along a number of dimensions, such as the probability of conflict detection. We should thus verify with other studies, whether the FOE keep on being diagnostic of error also with other classic and more difficult reasoning problems, such as syllogisms or conditionals, although the attribute substitution has also been proposed as an explanation for errors in other classic tasks, such as the conjuction fallacy task (Kahneman & Frederick, 2002). Some recent studies with the standard conjuction problem showed that biased reasoners are less confident about their response, than controls (e.g. De Neys & Feremans, 2012) Another reason for this apparent discrepancy between FOE and FOR could be that the cues to feeling right (like fluency) are quite different from those giving an intuition of error. Maybe different instructions could cause individuals to attend to different cues. In our study, participants are required to pay attention to the unpleasant sensation they felt during the task. In other studies participants are asked to express their Feelings of Rightness in their answer. Another reason could be the difference between our FOE questionnaire, and the one-item rating confidence scale usually employed in previous studies (e.g. Prowse Turner & Thompson, 2009).

Second, of course, the link between monitoring and performance is not always mediated by conscious feelings (Reder & Schunn, 1996), raising the questions of why one should experience any feelings at all, and in case whether s/he is aware of the source of those feelings. We did not directly tackle the issue of consciousness in this paper, but through the FOE- Questionnaire, used to account for the extent of FOE, we submit that they are conscious experiences, even though they do not consist in a full-fledged explicit account of errors and misleading cognitive procedures. Related to this, even though the FOE-Questionnaire requires participants to think about what their feelings were during the

task, it is not clear whether they report a lingering impression or suggested post-hoc feelings. It is also not clear whether they report a sense that they have committed an error rather than the sense that they have not been complying with a correct procedure.

The present paper represents a preliminary study whose results need to be extended. With it we hope to stimulate further research on negative self-appraisals by subjects who are victims of biases or of cognitive illusion.

References

Carbonnell, L. & Falkenstein, M. (2006). Does the error negativity reflect the degree of response conflict? *Brain Research, 1095,* 124-130.

Coles, M. E., Frost, R. O., Heimberg, R. G., & Rhéaume, J. (2003). Not just right experiences: perfectionism, obsessive–compulsive features and general psychopathology. *Behaviour Research and Therapy, 41,* 681–700.

Coles, M. E., Heimberg, R. G, Frost, R. O., & Steketee, G. (2005). Not just right experiences and obsessive–compulsive features: Experimental and self-monitoring perspectives. *Behaviour Research and Therapy, 43,* 153–167.

De Neys, W. (2012). Bias and conflict: A case for logical intuitions. Perspectives on Psychological Science, 7, 28–38.

De Neys, W., & Bonnefon, J. F. (2013). The whys and whens of individual differences in individual thinking biases. *Trends in Cognitive Sciences, 17,* 172-178.

De Neys, W., & Feremans, V. (2012). Development of heuristic bias detection in elementary school. Developmental Psychology.

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

De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and decision confidence. PLoS ONE, e15954. doi:10.1371/journal.pone.0015954

De Neys, W., Moyens, E., & Vansteenwegen, D. (2010). Feeling we're biased: Autonomic arousal and reasoning conflict. *Cognitive, Affective, & Behavioral Neuroscience, 10*, 208–216.

De Neys, W., Cromheeke, S., & Osman, M. (2011). Biased but in doubt: Conflict and Decision Confidence. PloS One, 6 (1), 1-10.

De Neys, W., Rossi, S., & Houdé, O. (2013). Bats, balls, and substitution sensitivity: Cognitive misers are no happy fools. *Psychonomic Bulletin& Review, 20,* 269-273

Efklides, A. (2006). Metacognitive and affect: What can metacognitive experiences tell us about the learning process? *Educational Research Review, 1,* 3-14.

Frederick, S. (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives, 19,* 25-42.

Glucksberg, S. & McCloskey, M. (1981). Decisions about ignorance: Knowing that you don't know. *Journal of Experimental Psychology: Human Learning and Memory , 7,* 311-325.

Haidt, J. (2001). The emotional dog and it's rational tail: a social intuitionist approach to moral judgment. *Psychoogcal Review, 108,* 814–834.

Harvey, A., Watkins, E., Mansell, W., & Shafran, R. (2004). *Cognitive behavioural processes across psychological disorders. A trasdiagnostic approach to research and treatment.* Oxford: Oxford University Press.

Kahneman, D. (2011). Thinking, Fast and Slow. New York: Farrar,Strauss, Giroux.

Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D.

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgements: Information based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 483–502).New York, NY: Guilford Press.

Koriat, A., (2000). The Feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, *9*, 149-171.

Metcalfe, J. (2000). Feelings and judgements of knowing: Is there a special poetic state? *Consciousness and Cognition, 9*, 178-186.

Nelson, T. O., & Narens, L. (1980). Norms of 300 general-information questions: Accuracy of recall, latency of recall, and feeling-of-knowledge ratings. *Journal of Verbal Learning & Verbal Behavior, 19*, 338-368.

Piattelli-Palmarini, M. (1994). *Inevitable Illusions*. New York: Wiley.

Prowse Turner, J. A., & Thompson, V. A. (2009). The role of training, alternative models, and logical necessity in determinin g confidence in syllogistic reasoning. *Thinking & Reasoning, 15*, 69–100.

Reder, L. & Schunn, C. (1996). Metacognition does not imply awareness: Strategy choice is governed by implicit learning and memory. In Reder, L. (Ed.) *Implicit Memory and Metacognition*. Mahwah, N.J.: Erlbaum, 45-77.

Shynkaruk, J. M., & Thompson, V. A. (2006). Confidence and accuracy in deductive reasoning. *Memory & Cognition*, *34*, 619–632.

Stanovich, K.E., & West, R.F. (2008). On the relative independence of thinking biases and cognitive ability. *Journal of Personality and Social Psychology, 94*, 672-695.

Thompson, V. A. (2009). Dual process theories: A metacognitive perspective. In J. Evans & K. Frankish (Eds.), In two minds: Dual processes and beyond. Oxford: Oxford University Press.

Thompson, V. A., & Morsanyi, K. (2012). Analytic thinking: Do you feel like it? *Mind & Society, 11*, 93-105.

Thompson, V., Prowse Turner, J.A., & Pennycook, G. (2011). Intuition, reason, and Metacognition. *Cognitive Psychology, 63*, 107–140.

Thompson, V.A., Evans, J. St. B. T., & Campbell, J. I. C. (2013). Matching bias on the selection task: It's fast and it feels good. *Thinking & Reasoning*, *19*, 431-452.

Thompson, V.A., & Johnson, S.J. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning*, *20*, 215-244.