Research Report

Developmental frontal brain activation differences in overcoming heuristic bias

Katell Mevel [a,b], *Grégoire Borst* [a,b,c], *Nicolas Poirel* [a,b,c], *Grégory Simon* [a,d], *François Orliac* [a], *Olivier Etard* [d,e], *Olivier Houdé* [a,b,c] *and Wim De Neys* [a,b,f,*]

[a] *Laboratory for the Psychology of Child Development and Education, Sorbonne, CNRS UMR, 8240, Paris, France*
[b] *Paris Descartes University, Sorbonne Paris Cité, Paris, France*
[c] *Institut Universitaire de France, Paris, France*
[d] *Université de Normandie, ISTS, EA 7466, GIP Cyceron, Caen, France*
[e] *Centre Hospitalier Universitaire de Caen, Service des Explorations Fonctionnelles du Système Nerveux, Caen, France*
[f] *CNRS, France*

ABSTRACT

Since reasoning is often biased by intuitive heuristics, the development of sound reasoning has long been postulated to depend on successful bias monitoring and inhibition. The present fMRI study aimed to identify neural correlates of developmental changes in these processes. A group of adults and young adolescents were presented with ratio-bias problems in which an intuitively cued heuristic response could be incongruent (conflict item) or congruent (no-conflict item) with the correct response. Results showed that successfully avoiding biased responding on conflict items across both age groups was associated with increased activation in Anterior Cingulate Cortex (ACC) and the right Lateral Prefrontal Cortex (LPFC) regions of interest. Critically, the right LPFC activation decreased with age. Biased responding did not result in right LPFC or ACC modulation and failed to show any developmental activation changes. We discuss implications for ongoing debates on the nature of heuristic bias and its development.

© 2019 Elsevier Ltd. All rights reserved.

---

* *Corresponding author*. Sorbonne — Labo Binet, Paris Descartes University, 46 Rue Saint-Jacques, 75005, Paris, France.
E-mail address: wim.de-neys@parisdescartes.fr (W. De Neys).

1.   Introduction

Although human beings have unique cognitive abilities to reason, we do not always reason correctly. Influential research in the cognitive sciences has shown that in a wide range of tasks even educated adults often fail to respect basic logical and probabilistic principles (e.g., Kahneman, 2011). Consider the following example:

You are faced with two trays each filled with white and red jelly beans. You can draw one jelly bean without looking from one of the trays. The small tray contains a total of 10 jelly beans of which 1 is red. The large tray contains a total of 100 jelly beans of which 9 are red.

From which tray should you draw to maximize your chance of drawing a red jelly bean?

1. The small tray
2. The large tray

The example is an illustration of the infamous ratio-bias problem (Epstein, 1994). When presented with this problem many reasoners have a strong intuitive preference for the large tray. From a logical point of view, this is not correct of course. Although the large tray contains more red beans than the small tray (9 vs 1), there are also a lot more white beans in the large tray. If you take the ratio of red and white beans in both trays into account it is clear that the small tray is giving you a 10% chance of picking a red bean (i.e., 1/10) while the large tray only offers a 9% chance (i.e., 9/100). However, many educated reasoners are tricked by the absolute difference and fail to solve this basic "ratio" problem (e.g., Epstein, 1994). The fact that the absolute number of red beans is higher in the large tray has such a strong intuitive pull on people's thinking that they seem to neglect the ratio principle and end up being biased.

In general, it has been argued that human reasoners tend to base their judgment on fast intuitive impressions rather than on more demanding deliberative reasoning (e.g., Evans & Stanovich, 2013; Kahneman, 2011). Although this intuitive or so-called "heuristic" thinking might sometimes be useful, the problem is that it can also cue responses that conflict with more logical-mathematical considerations and bias our thinking.

Not surprisingly, reasoning theories have long stressed the importance of bias monitoring and inhibitory processing for sound reasoning (e.g., De Neys & Bonnefon, 2013; Evans, 2007; Houdé, 2000; Kahneman, 2011). The idea is that to avoid biased conclusions it is paramount to monitor for conflict between intuitively cued heuristics and more logical considerations and inhibit the heuristic response in case a conflict is detected. Consistent with these suggestions a large body of neuroimaging work has pointed to increased activation in cognitive control regions that are believed to be involved in conflict monitoring (e.g., Anterior Cingulate Cortex, ACC) and response inhibition (e.g., right lateral prefrontal cortex, right LPFC) when reasoners manage to avoid biased responding in logical and probabilistic reasoning tasks (e.g., De Martino, Kumaran, Seymour, & Dolan, 2006; De Neys, Vartanian, & Goel, 2008; Goel, Buchel, Frith, & Dolan, 2000; Houdé et al.,

2011; Leroux et al., 2009; Prado, Chadha, & Booth, 2011; Prado & Noveck, 2007; Stollstorf, Vartanian, & Goel, 2012; Tsujii, Masuda, Akiyama, & Watanabe, 2010; Tsujii & Watanabe, 2010; Vartanian et al., 2018). These studies typically contrast people's processing of classic reasoning problems in which a cued heuristic conflicts with the logically correct response (e.g., as in the introductory ratio bias problem) with control "no-conflict" versions. In the no-conflict versions the conflict is removed and the cued heuristic response is consistent with the correct logical response. For example, a no-conflict control version of the introductory ratio bias problem could simply state that the large tray contains 11 (instead of 9) red beans. In this case both the absolute number of red beans (i.e., 1 vs 11) and the ratio of red beans (i.e., 1/10 vs 11/100) would be higher in the large tray. Hence, both heuristic considerations based on the absolute number and logico-mathematical ratio considerations cue the exact same response. Hence, there is no conflict and no need to inhibit the heuristic responses.

Maturation of the monitoring and inhibition mediating brain regions and processes has also been assumed to be a key factor in our reasoning development (e.g., Amsel et al., 2008; Barrouillet, 2011; De Neys, 2013; De Neys & Feremans, 2013; Houdé, 2000; Houdé & Borst, 2015; Reyna, Chapman, Dougherty, & Confrey, 2012; Simon, Lubin, Houdé, & De Neys, 2015). However, few reasoning bias studies have directly contrasted how neural activation in these areas differs in adult and younger reasoners. The present study was designed to address this problem.

Given that cognitive control processes are believed to show key development during adolescence (e.g., Casey & Caudle, 2013; Luna, 2009; Luna, Padmanabhan, & O'Hearn, 2010; Santesso & Segalowitz, 2008) we recruited a sample of young adolescents and adults. Participants were presented with conflict and no-conflict problems modelled after the ratio bias task. We were specifically interested in age-group differences for both correctly and incorrectly solved conflict problems in the ACC and right LPFC regions of interest since these have been repeatedly linked with conflict monitoring and inhibition in the reasoning field (De Neys et al., 2008; Goel & Dolan, 2003; Stollstorff, Vartanian, & Goel, 2012; Vartanian et al., 2018). In line with previous studies, we expected to find increased activation in the ACC and right LPFC across both age groups for correctly solved conflict (vs no-conflict) problems. Our key interest was to examine how the activation in these regions differed between age groups.

In addition to correct responses, we also wanted to examine developmental changes in the neural correlates of incorrect (i.e., biased) conflict responses. As we noted, biased responding is believed to result from a failure to inhibit the heuristic responses. But the nature of this failure is debated (e.g., De Neys & Glumicic, 2008; see also; Mata & Ferreira, 2018; Travers, Rolison, & Feeney, 2016). One possibility is that biased reasoners detect that the heuristic response is logically incorrect but subsequently fail to successfully inhibit the salient heuristic response (i.e., a pure inhibition failure, e.g., Sloman, 1996; Epstein, 1994). Alternatively, people might simply fail to detect that the heuristic conflicts with logical considerations and see no need to inhibit the response (i.e., a detection failure, e.g., Evans & Stanovich, 2013; Kahneman, 2011).

In recent years behavioral and imaging studies have provided some evidence against the detection failure view (see De Neys, 2017, for a review). For example, at the behavioral level it has been shown that biased responders doubt their answer as reflected in longer latencies and decreased response confidence for conflict versus no-conflict problems (e.g., Ball, Thompson, & Stupple, 2017; Bonner & Newell, 2010; De Neys & Glumicic, 2008; Gangemi, Bourgeois-Gironde, & Mancini, 2015; Johnson, Tubau, & De Neys, 2016; Mevel et al., 2015; Pennycook, Trippas, Handley, & Thompson, 2014; Stupple, Ball, Evans, & Kamal-Smith, 2011; Thompson & Johnson, 2014; but see also Aczel, Szollosi, & Bago, 2016; Ferreira, Mata, Donkin, Sherman, & Ihmels, 2016). In addition, fMRI studies have found that biased reasoners—just as correct reasoners—show increased ACC (but not right LPFC) activation when solving conflict versus no-conflict problems (De Neys et al., 2008; Simon et al., 2015; Vartanian et al., 2018). These results have been taken to indicate that biased (adult) reasoners show sensitivity to the fact that their answer is logically incorrect and detect logic/heuristic conflict (De Neys, 2012, 2017). Critically, it has also been suggested that this conflict detection for incorrect responses would be late to develop and that this would be reflected in differential engagement of monitoring regions such as the ACC for younger reasoners than for adults (De Neys & Feremans, 2013; Simon et al., 2015). However, to date this neural hypothesis has not been tested. By contrasting the activation in our two age groups for incorrect conflict responses the current study could directly inform this debate.

2. Method

2.1. Participants

Forty-nine participants were scanned in this study, but brain imaging data from three adults and nine adolescents had to be discarded before analysis: Two adults and five adolescents exhibited a within-run maximal amplitude of translational or rotational between-volumes displacement above 2 mm or 1.5°, respectively; one adult and three adolescents had to be excluded due to technical scanner problems; finally, one adolescent was excluded because of a neurological abnormality. All remaining 37 participants (20 adults, 9 males, mean age = 21.4, SE = .4; 17 adolescents, 7 males, mean age = 13.1, SE = .3) were right handed native French speakers, scoring higher than + 40% to the Oldfield handedness scale. All participants or their parents/guardians gave written informed consent and received a 25 euro gift voucher. The study was approved by the research ethics committee (CPP Nord Ouest III; ID RCB: 2014-A00935-42) and conformed to the 1964 Declaration of Helsinki.

Sample size was determined on the basis of budget considerations. Our research budget allowed us to test 22 adolescents and 22 adults. The scanning center (Cyceron, Caen, France) allowed us to test 5 additional participants free of charge to compensate for sessions that had to be discarded due to technical scanner failure (4) and a detected brain anomaly (1). This resulted in a total number of 49 scanned participants.

2.2. Behavioral task and materials

The task format adopted in the present experiment was based on the ratio-bias task as previously published (e.g., Bonner & Newell, 2010; Mevel et al., 2015). Participants had to solve a total of 50 problems. The tray pairs were developed on the basis of percentage ranges used in previous research. The small tray contained 1, 2 or 3 red marbles out of a total of 5 or 10 (i.e., 20—60% or 10—30%, Rudski & Volksdorf, 2002). The large tray always had a total of 100 marbles and the proportion of red marbles differed from the small tray by a value of ±5%—9%. Twenty-five of the problems were Conflict problems and 25 were No-Conflict (control) problems. To make the task less repetitive, within each category 10 problems showed small trays containing 5 marbles (e.g., 1 red out of 5) and 15 problems showed a small tray containing 10 marbles (e.g., 1 red out of 10). Each problem was presented on a separate black slide, with the red and white marbles being ordered and the proportions given below in absolute written numbers (Mevel et al., 2015, see Fig. 1a for an illustration of the task format). The left tray was the small one for approximately half of the problems and the large one for the other half.

Before entering the scanner, participants were shown two problems as examples and were informed they had a maximum of 6500 msec to make their decision (pilot testing indicated that this deadline allowed ample time to reach a decision). Then, to further familiarize participants with the task format, they also solved 4 problems themselves (i.e., all No-Conflict problems with different proportions than in the fMRI task).

After the 50 Conflict and No-conflict items were presented, participants were also presented with one final attention check problem. This problem presented a choice between a large tray and a small tray with 0 red marbles in one or the other. Accuracy on the attention check reached 95% (adults: 100%; adolescents: 88%), suggesting that by and large participants were engaged in the task until the end of the study.[1]

2.3. Behavioral data analysis

For each individual we calculated the mean accuracy for Conflict and No-Conflict problems. Two adults did not err on any of the conflict problems, and two adolescents never succeeded in solving conflict problems correctly. They were discarded from the corresponding contrast analyses. When available, individual mean Response Times (RTs) were extracted for a) Incorrect Conflict (C0), b) Correct Conflict (C1) and c) Correct No-Conflict (NC1) problems. An analysis of variance (ANOVA) with one within factor (type of problem: conflict or no-conflict) and one between factor (age group: adolescents or adults) was run on the accuracy scores. For response times, ANOVAs with one within factor (type of problem) and one between factor (age group) were run for each comparison of interest: 1) C0 versus NC1, and 2) C1 versus NC1.

---

[1] Exclusion of two adolescent participants who failed the attention check control item did not alter the fMRI findings. Reported results concern the full sample without exclusion.
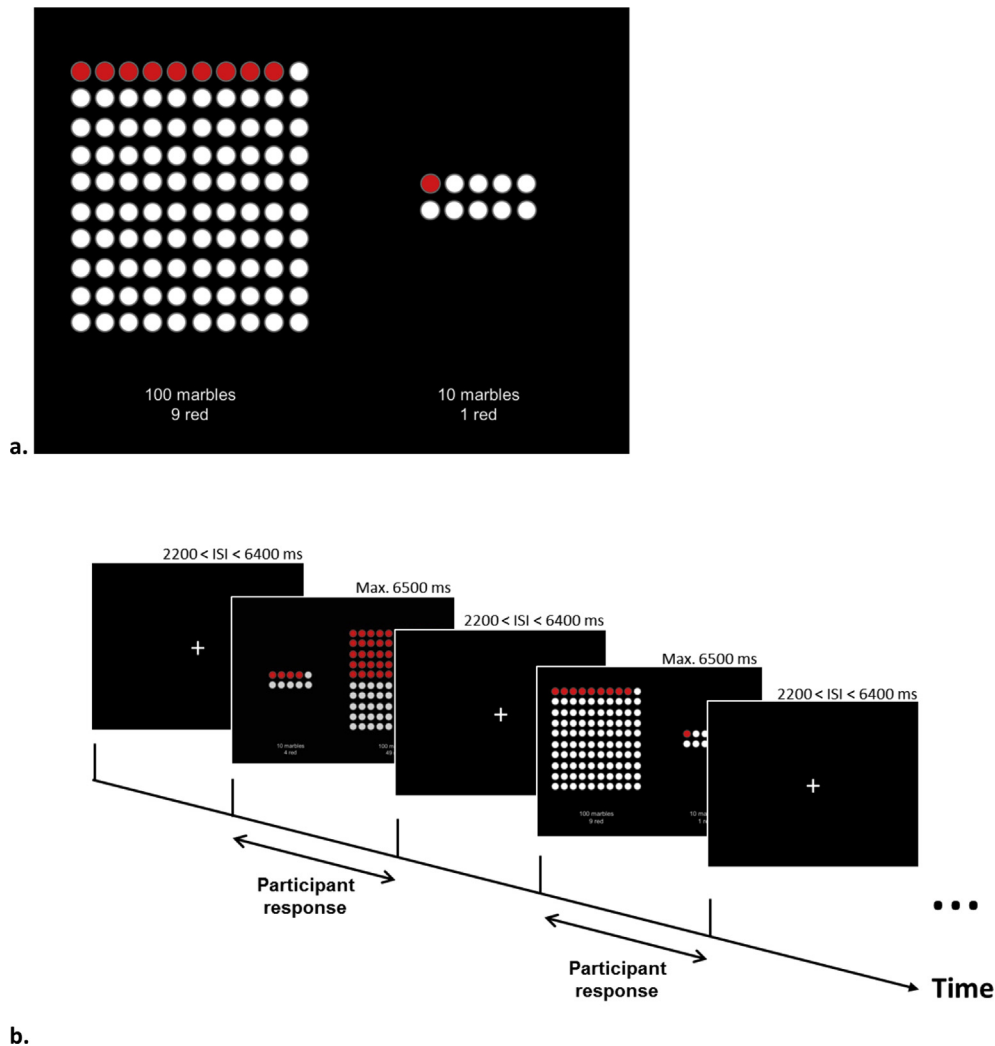
a.

b.

Fig. 1 — **a. Task illustration. Illustration of a conflict problem as presented during the fMRI task. b. Protocol. Each trial started with a white fixation cross in the middle of the screen, followed by the presentation of the item to be solved. Participants had a maximum of 6500 msec to make their decision. The inter-stimulus interval was randomly sampled between 2200 and 6400 msec.**

Missed conflict trials for which participants did not have time to respond (i.e., 12 trials in total, less than 1.5% of all conflict trials) were considered as failed and were set at the max RT (i.e., 6500 msec). For completeness, note that the incorrectly solved no-conflict trials were not considered in the response time (or fMRI, see further) analyses. In no-conflict trials, heuristic and logical considerations point to the same correct response. Consequently, incorrect responses are rare and cannot be interpreted unequivocally (De Neys & Glumicic, 2008; Pennycook, Fugelsang, & Koehler, 2015).

2.4.   fMRI event-related design

Fig. 1b illustrates the design. Each trial started with the presentation of a fixation cross in the middle of the screen, followed by the slide with the ratio bias item. Participants had to indicate within 6500 msec whether they chose the small or the big tray by pressing one of two buttons. As soon as they answered, the next trial started with the presentation of a fixation cross for a random duration sampled between 2200 msec and 6400 msec. Problems were presented in a jittered pseudo-randomized order for each participant: The sequence of Conflict and No-Conflict control problems was the same for all participants, but the specific problem within each category was randomly selected from the available stimuli, for each participant. Ten null events consisting of a fixation cross for a duration ranging from 2200 to 6400 msec were also added to the design to optimize signal detection. The overall optimization of the design was achieved with the Genetic Algorithm toolbox (Wager & Nichols, 2003).

Note that we initially also envisaged to invite participants to go through a second run of the same task right after the first one. The overall structure of the task was the same and the stimuli presented were identical, except that the right and left trays were inverted. However, especially in the adolescent group, participants frequently indicated they opted to refrain

from doing a second run. Because pre-processing of the available data also pointed to problematic within-run maximal displacements for the remaining participants, we decided to discard the second run data from further analysis altogether.

2.5. fMRI data acquisition

fMRI data were acquired at Cyceron (Caen, France) using a 3T scanner (Philips Achieva, Eindhoven, The Netherlands). Scanning was done as part of a larger one hour test session during which an unrelated task was also administered. For each participant, the MR data were collected using planes parallel to the anterior commissure − posterior commissure line. fMRI time series were collected using T2*-weighted FFE echo-planar imaging (EPI) sequence (31 axial slices; 3.75 mm thickness; no gap; reconstruction matrix = 640 × 640 mm; field of view = 240; repetition time = 2000 msec; echo time = 35 msec; flip angle = 80°; interleaved: bottom − up; max: 355 dynamics). The FOV covered the top of the cortex down to at least the base of the cerebellum. The first 6 volumes were discarded to account for spin saturation effects. A high resolution T1-weighted structural and a T2* non-EPI volumes were also acquired using 3D TFE (180 sagittal slices; resolution: 1 mm3; no gap; reconstruction matrix = 256 × 256 mm; field of view = 256; repetition time = 20 msec; echo time = 46 msec) and 2D FFE (70 axial slices; resolution: 2 mm3; no gap; reconstruction matrix = 256 × 256 mm; field of view = 256; repetition time = 20 msec; echo time = 46 msec) sequences, respectively.

2.6. fMRI data analysis

Image pre-processing and statistical analyses were performed using Matlab 2017a (Mathworks Inc., Natick, USA) and SPM12 (Wellcome Department of Imaging Neuroscience, London, UK). fMRI datasets were preprocessed as follows: individual fMRI time-series were (1) slice time corrected, (2) realigned to correct for head movement, (3) co-registered to the respective individual T2* non-EPI volume, which was already previously co-registered to the high-resolution structural image T1-weighted volume, (4) transformed into the MNI152 T1 template space using the parameters derived from the spatial normalization of the T1-weighted image in DARTEL, (5) resampled to 2-mm isotropic voxel size and (6) smoothed with a 8-mm_FWHM Gaussian Kernel.

We analyzed the data voxel-wise using the general linear model (GLM) approach. We defined four regressors, according to two crossed factors (1) a regressor related to the Type of Problem (i.e., the presentation of Conflict or No-Conflict problems): x (2) a regressor related to the Accuracy of the response (i.e., correctly solved or failed problems). For each regressor we modeled the stimulus-evoked neural response as a boxcar function, with each event starting at stimulus presentation onset, and lasting the time the participant took to make their decision. These boxcar functions were convolved with the default canonical hemodynamic response function of SPM12. In addition, we included one constant term and the six realignment parameters (3 translations, 3 rotations) as covariates into our model, to account for changes in signal level and influence of head motion on BOLD signal, respectively. We also applied a high-pass filter (cut off period = 128 sec) to remove low-frequency drifts. Of note, incorrectly solved No-Conflict problems were modeled but not investigated as these are known to be rare and are of no theoretical interest (e.g., De Neys et al., 2008; Mevel et al., 2015; Pennycook et al., 2015).

We first estimated the model at the individual level yielding two parameter estimates per condition (i.e., correctly solved or failed Conflict or No-Conflict problems). From these values, we computed, for each participant, the following statistical contrast maps: Failed Conflict > Correct No-Conflict (C0 > NC1), Correct Conflict > Correct No-Conflict (C1 > NC1), Correct Conflict > Failed Conflict (C1 > C0) and Failed Conflict > Correct Conflict (C0 > C1).

Our main interest was whether there was differential activation in Adolescents as compared to Adults. At the group level, we conducted two-sample t-tests, for each contrast described above, to test for any activation differences in the Adult and Adolescent groups (Adults > Adolescents, Adolescents < Adults). We determined a priori anterior cingulate cortex (ACC) and right lateral prefrontal cortex (right LPFC) regions of interest (ROI) based on the work of De Neys et al. (2008). Following De Neys et al. (2008), the ROIs were 12 mm radius spheres centered on the following coordinates ACC [0, 14, 42] and right LPFC [50, 20, 10] (see Fig. 2 for an illustration). We first performed analyses using the ROIs for which we applied both voxel-level and cluster-level thresholds at p FWE <.05, using the Small Volume Correction module in SPM12. Subsequently, we also ran whole-brain analyses for which we applied a voxel-level threshold at $p < .05$ (FWE corrected) and reported clusters larger than the k_{max} "Expected Voxels per Clusters". Finally, we performed supplementary analyses in which we combined adults and adolescents in one single group to highlight any potential common activation patterns. To this end, we performed one-sample t-tests using the same ROIs and whole-brain exploration approach.

3. Results and discussion

3.1. Behavioral data

3.1.1. Accuracy
Table 1 shows the accuracy results. A 2 (Age group) x 2 (Problem type) ANOVA indicated that there was a main effect of problem type such that mean accuracy was lower on conflict (63%) than on no-conflict problems (86.3%), $F(1,35) = 16.4$, $p = .0003$, $\eta p^2 = .032$. The main effect of age group was also significant, $F(1,35) = 14.3$, $p = .0006$, $\eta p^2 = .029$, indicating that adults (83.2%) reasoned more accurately than adolescents (66.1%). Although there was a trend towards a more pronounced age effect on the conflict problems, the interaction did not reach significance, $F(1,35) = .73$, $p = .40$, $\eta p^2 = .02$. By and large the accuracy performance is consistent with previous behavioral ratio-bias task studies (e.g., Bonner & Newell, 2010; Ferreira et al., 2016; Mevel et al., 2015).
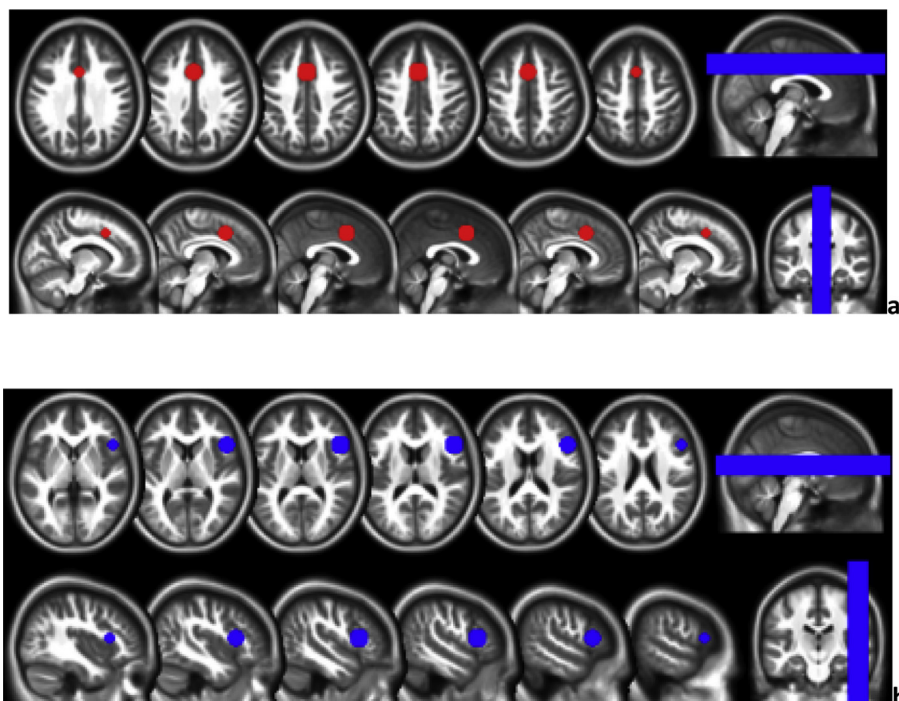
Fig. 2 − Regions of Interest (ROI). Overlay of the ROIs onto the Whole-Sample mean anatomical image, together with an orthogonal projection of the views. (a) Anterior Cingulate Cortex (ACC) in red, (b) Right Lateral Prefrontal Cortex (LPFC) in blue. Axial (up) and sagittal (down) views.

3.1.2. Response time

As in previous behavioral studies (e.g., De Neys & Glumicic, 2008; Pennycook et al., 2015), we contrasted the response times for both correctly and incorrectly solved conflict problems with response times for correctly solved no-conflict control problems. Table 1 shows the results. Consistent with the literature, for correct conflict trials, an ANOVA showed that conflict trials took longer than no-conflict trials, $F(1,33) = 8.6$, $p = .006$, $\eta p^2 = .21$. There was no further main effect of age, $F(1,33) = 2.8$, $p = .10$, $\eta p^2 = .08$, or interaction, $F(1,33) = 1.29$, $p = .26$, $\eta p^2 = .04$. For failed conflict trials, there was a main effect of age group, $F(1,33) = 4.28$, $p = .046$, $\eta p^2 = .11$, suggesting that adolescents were overall faster to respond. However, neither the effect of problem type, $F(1,33) = 3.06$, $p = .09$, $\eta p^2 = .08$, nor the interaction, $F(1,33) = 1.34$, $p = .26$, $\eta p^2 = .04$, reached significance. As we noted, increased latencies for conflict versus control no-conflict problems are typically taken as evidence for successful conflict detection among biased responders (De Neys &

Glumicic, 2008; Pennycook et al., 2015). Hence, at the behavioral level the response time results do not point to successful conflict detection among biased responders or a significant developmental modulation of this ability.

3.2. fMRI data

3.2.1. Developmental contrast: Adolescents versus Adults

We first contrasted adults and adolescents' brain activation in our ACC and right LPFC ROIs in the different contrasts of interest. With respect to correct conflict responses, results for the C1 > NC1 contrast showed that the right LPFC region [58 18 20] (k = 15, T = 3.93, peak-level p_{FWE} = .028, cluster-level p_{FWE} = .032) was differentially recruited in the adolescent and adult sample (see Fig. 3a). Critically, the LPFC activation was higher for adolescents than for adults (see Fig. 3b). Activation in the ACC region did not differ in the two age groups. With respect to incorrect conflict responses, results for the C0 > NC1 contrast showed that neither LPFC, nor ACC activation differed in the two age groups when participants were biased and failed to respond correctly to the conflict problem. The direct contrast of correct and incorrect conflict responses (C1 > C0) did also not point to differential developmental activation. None of the supplementary (see methods) reversed contrasts we explored pointed to any further significant activation differences in the ACC or LPFC ROIs.

In addition to our ROI analyses we also ran whole-brain analyses to explore any additional regions outside of our a priori ROIs that might show developmental modulation (P FWE < .05). However, none of the contrasts pointed to significant activation differences.

Table 1 − Overview of average (SEM) behavioral accuracy and response time.

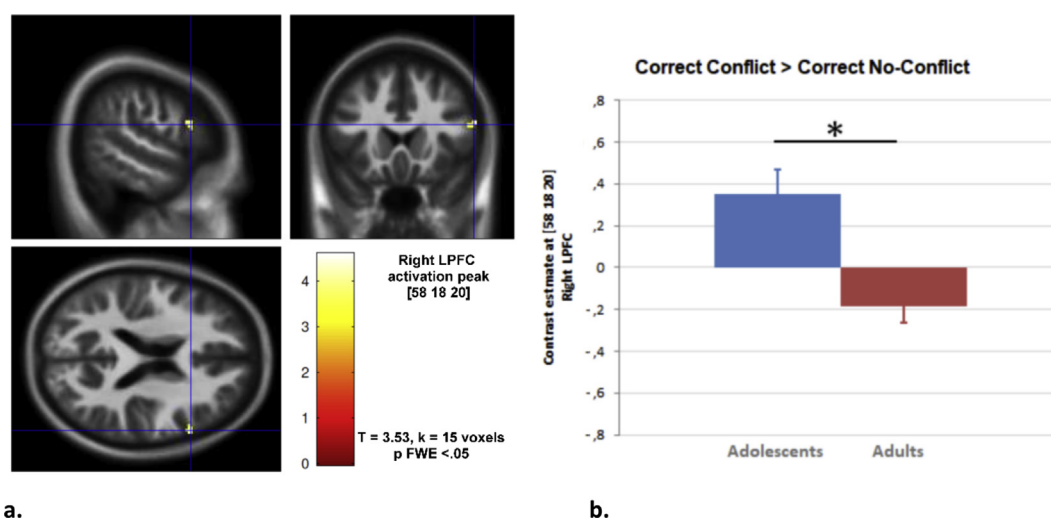|	Adolescents	Adults
Accuracy		
Conflict	52 ± 6.7	74 ± 6.1
No-Conflict	80.2 ± 3.7	92.4 ± 3.4
Response Time		
Correct Conflict	2424.9 ± 192	2982.7 ± 218
Failed Conflict	2196.8 ± 290	3114.4 ± 351
Correct No-Conflict	2219.4 ± 231	2729.6 ± 190

Fig. 3 − Main results of the developmental contrast. Adolescents show stronger right Lateral Prefrontal Cortex (LPFC) activation while correctly solving Conflict problems as compared to No-conflict problems [C1 > NC1]. a. Illustration of the peak activation at [58 18 20] given by the Small Volume Correction procedure, projected onto the whole-sample mean anatomical image. b. Plot of the BOLD individual contrast values as a function of the age group, extracted at the peak voxel. Error bars show one standard deviation.

3.2.2. Whole-group analysis

The developmental contrast indicated that right LPFC activation for correctly solved conflict problems differed in our two age groups. The ACC activation never differed. The lack of ACC activation differences might result from successful but equal conflict monitoring related ACC recruitment in both age groups. Alternatively, it might result from a lack of successful conflict monitoring and ACC activation in both age groups altogether. Bluntly put, either adolescents or adults detect conflict equally well or both groups fail to detect conflict. To get a maximally powerful test of these alternative hypotheses we combined adolescents and adults in a single group analysis and looked at differential activation for conflict and no-conflict problems across the whole group. We ran both ROI and whole-brain analyses and looked at both correct and incorrectly solved conflict trials.

With respect to correctly solved conflict trials (i.e., C1 > NC1 contrast), ROI results indicated that there was increased right LPFC [50 22 20] (k = 307, T = 5.35, peak-level p_{FWE} = .001, cluster-level p_{FWE} < .0001) and ACC [-6 24 44] (k = 13, T = 4.05, peak-level p_{FWE} = .021, cluster-level p_{FWE} = .035) activation when solving conflict versus no-conflict problems. These findings fit with previous fMRI findings in the reasoning field that pointed to the implication of these same two regions in successful conflict monitoring and inhibition of the heuristic response (e.g., De Neys et al., 2008; Prado et al., 2007; Simon et al., 2015; Stollstorff et al., 2012; Vartanian et al., 2018). With respect to incorrectly solved conflict problems (C0 > NC1 contrast), we observed neither right LPFC nor ACC recruitment. The lack of right LPFC recruitment fits with previous findings and theoretical predictions (i.e., biased responders are assumed to fail to inhibit). However, the lack of ACC activation directly contradicts previous findings (e.g., De Neys et al., 2008; Simon et al., 2015; Vartanian et al., 2018) and suggests that participants (adolescents and adults alike) fail to successfully register heuristic/logic conflict when they are

biased. The direct contrast of correct and incorrect conflict responses (i.e, C1 > C0) supported these conclusions and showed that correct responses were associated with both higher right LFPC [44 12 18] (k = 57, T = 4.80, peak-level p_{FWE} = .003, cluster-level p_{FWE} = .008) and ACC [0 16 50] (k = 111, T = 4.41, peak-level p_{FWE} = .008, cluster-level p_{FWE} = .002) activation.

In addition to the ROI analyses, we also ran whole-brain analyses on the combined whole group data to look for additional conflict-related activation outside the ACC and right LPFC regions. With respect to incorrect conflict responses (C0 > NC1), we failed to find any significant activation differences when solving conflict versus no-conflict problems. However, for correct responses (C1 > NC1), we observed increased activation for conflict versus no-conflict problems in the Inferior Frontal Gyrus (Pars Orbitalis, IFG) [-34 44 0] (k = 4, T = 6.11, peak-level p_{FWE} = .018, cluster-level p_{FWE} = .013). The direct contrast of correct and incorrect conflict (C1 > C0) responses pointed to additional activation in the Superior Parietal Cortex [-24 -62 60] (k = 9, T = 6.21, peak-level p_{FWE} = .014, cluster-level p_{FWE} = .007). Note that similar IFG and Parietal activations have already been linked to inhibitory processing during reasoning (e.g., Prado et al., 2007, 2011). For completeness, we also contrasted the activation in these two regions in our two age groups (i.e., 12 mm radius sphere centered on peak activation coordinates with small volume correction and p FWE p < .05). However, unlike the LPFC, activation in the IFG and Parietal ROI did not show significant developmental modulation.

4.   General discussion

Reasoning theories have long postulated that sound reasoning in the face of biasing heuristics depends on successful bias monitoring and inhibition. Imaging studies with adults and

younger reasoners have pointed to a key role of the Anterior Cingulate Cortex (ACC) and right Lateral Prefrontal Cortex (LPFC) in these processes. In the present study we directly contrasted neural activation in different age groups to identify potential developmental modulation. A group of adults and young adolescents solved ratio-bias problems in which an intuitively cued heuristic response could conflict with the correct response. Results showed that correct responding on conflict items was associated with an age related decrease in right LPFC activation. Hence, in those cases that the tempting heuristic response was successfully blocked, it required less right LPFC activation for adults then for adolescents. With respect to incorrectly solved conflict problems we did not observe any developmental activation differences.

Interestingly, the age-related activation decrease in our right LPFC region of interest fits with the pattern that was previously observed with elementary inhibitory processing tasks in the cognitive control field (Luna, 2009; Luna et al., 2010). In basic response interference tasks such as the Go/No-Go or Stroop it has also been observed that adolescents show a stronger activation in inhibition mediating regions than adults when responding correctly (e.g., Luna et al., 2010; Tamm, Menon, & Reiss, 2002). It has been suggested that this might reflect the decreased effort required to exert inhibitory control with age (Luna et al., 2010). Hence, adolescents can already show adult-like performance and inhibit successfully but this comes at the cost of extra effort. Our results indicate that we find the same developmental trend when inhibiting a prepotent heuristic responses in a high-level reasoning task. This tentatively suggests that the overall accuracy boost in the ratio-bias task performance (i.e., adults were overall less biased than adolescents) results from the less demanding nature of the inhibitory process among adults. Hence, both adolescents and adults can manage to inhibit biasing heuristics but this process will be less demanding (i.e., require fewer resources) for adults than adolescents. Bluntly put, adults might overall reason more accurately in the face of interfering heuristics because the task gets relatively easier and not necessarily because they "think harder".

Note that recent behavioral studies in the reasoning field have suggested that correct logical responses in "bias" tasks can sometimes be generated intuitively because adults have automated—through years of formal education—the necessary logical computations (e.g., Bago & De Neys, 2017, 2018; De Neys, 2017; Newman, Gibb, & Thompson, 2017). Although they would still be faced with a conflicting heuristic response, their "logical intuition" would be stronger and dominate. Hence, one speculative explanation for why adults show less right LPFC activation than younger reasoners might be that among adults there is less need for an effortful override because their intuitions are more accurate to start with.

In contrast with the findings for correctly solved conflict problems, our results with respect to incorrect responses fit less well with previous literature. Contrary to our expectations, we failed to observe any developmental effects here. Moreover, in contrast with previous fMRI work on other bias tasks (e.g., base-rate neglect, De Neys et al., 2008, Vartanian et al., 2018; number conservation, Simon et al., 2015) we failed to find any significant ACC activation across age. Consistent with the present fMRI findings, behavioral response times also failed to show evidence for a significant conflict detection effect or developmental modulation of this effect. Taken together, these findings suggest that biased reasoners are not successful at monitoring for conflict between heuristics and logical considerations in the ratio bias task.

We noted that the nature of biased responding has been debated (e.g., De Neys & Glumicic, 2008; Ferreira et al., 2016; Frey, Bago, & De Neys, 2017; Mata, Ferreira, Voss, & Kollei, 2017; Travers et al., 2016). Recently, the debate on the nature of biased responding has pointed to the importance of individual differences (Frey, Johnson, & De Neys, 2018; Mata & Ferreira, 2018; Mata, Schubert, & Ferreira, 2014; Pennycook et al., 2015). Although previous behavioral studies seem to indicate that many biased reasoners show bias detection effects (e.g., increased latency when solving conflict problems) in a range of bias tasks, it is also clear there are non-negligible individual differences (Frey et al., 2018; Mata et al., 2017; Pennycook et al., 2015). To identify such possible variation in the current study we correlated the size of the behavioral detection effect (i.e., RT C0 − RT NC1; the amount of slowing down in response to the presence of conflict) with the ACC average BOLD signal in the same contrast (C0 > NC1), in our full sample (n = 35). Results showed that despite the absence of an overall group effect, there is a close mapping between the behavioral detection effect and ACC activation. The stronger the behavioral effect, the larger the ACC activation increase, r = .41, p = .014 (see Fig. 4a). For illustrative purposes, using the SVC procedure on the C0 > NC1 contrast, Fig. 4b plots the activation in two identified ACC peaks for the group of biased participants who show the behavioral latency effect (i.e., "detectors", RT C0 > RT NC1, n = 24) and those who do not (i.e., "non-detectors", RT C0 ≤ RT NC1, n = 11). As Fig. 4 illustrates, those biased reasoners who show the expected behavioral detection effect, do tend to show increased ACC activation. But the point is that a substantial number of biased responders do not show the behavioral effect. Hence, the current findings do not necessarily question the role of the ACC in conflict detection during reasoning. They question the prevalence of such detection such that various reasoners who show ratio bias will not detect their bias. Taken together, these findings are in line with recent criticism that has questioned the generality of successful conflict detection among biased reasoners (Mata et al., 2017; Pennycook, Fugelsang, & Koehler, 2012; Pennycook et al., 2015; Singmann, Klauer, & Kellen, 2014; Travers et al., 2016).

Note that in further individual differences analyses, we also explored the correlation between the average ACC BOLD signal in both the C0 > NC1 and C1 > NC1 contrasts and the behavioral accuracy on conflict problems. Results pointed to negative associations (C0 > NC1, r = −.29, p = .09; C1 > NC1, r = −.64, p < .0001), especially for the correct conflict responses. Hence, the rare times that less accurate reasoners do solve conflict problems correctly this is associated with stronger ACC activation. This pattern further points to the importance of individual differences in ACC recruitment. Nevertheless, it should be clear that all our individual difference analyses concern but a limited sample and were exploratory in nature. Although the results are suggestive, it is clear that they need to be interpreted with caution.
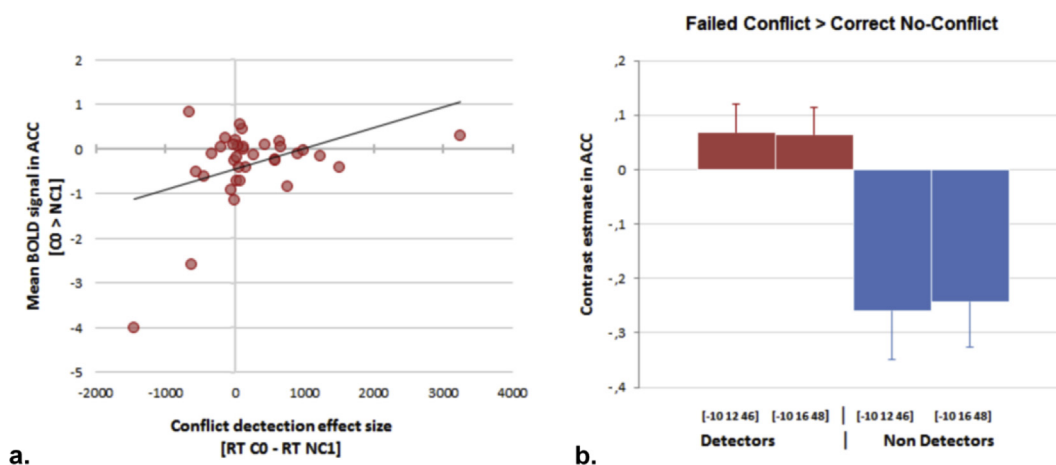
Fig. 4 — Plot of the correlation between the behavioral conflict detection effect and Anterior cingulate Cortex (ACC) BOLD signal on the [C0 > NC1] contrast. a. There is a significant positive correlation (r = .41, p = .014) between the mean BOLD signal in the ACC and the behavioral reaction time (RT) effect. b. Activation in two ACC peaks for the subgroup of reasoners who show (Detectors) and do not show (Non Detectors) the behavioral conflict detection effect.

To avoid confusion it is important to keep a number of further considerations and limitations into account when interpreting the current findings. First, we labeled the observed age-related increase in right LPFC activation as a developmental effect. It should be clear that we use the label "developmental" here in a purely descriptive sense to refer to an age-related difference. At no point do we imply that it necessarily reflects a structural change. Our study documents a critical age-related change in brain activation. It does not address the ultimate nature of this change. Second, our study focused on the right LPFC and ACC ROIs because these areas have been most extensively studied in related previous reasoning work (e.g., De Martino et al., 2006; De Neys et al., 2008; Goel et al., 2000; Houdé et al., 2011; Leroux et al., 2009; Prado & Noveck, 2007; Stollstorf et al., 2012; Tsujii et al., 2010; Tsujii & Watanabe, 2010; Vartanian et al., 2018). Clearly, this does not imply that other regions cannot contribute to sound reasoning. For example, our whole-brain analysis indicated that correct conflict responses were also associated with left lateralized prefrontal activation in the IFG. Moreover, recent studies also point to a possible role of, for example, the default mode network in biased responding (e.g., Li et al., 2017; Vartanian et al., 2018). Our ROI approach was conservative in nature in that we focused on two regions whose role in reasoning has been most extensively documented. But the focus on these two regions should not give rise to the false impression that other regions cannot be involved.

Finally, our discussion of the right LPFC activation focused on its well-established role in inhibition during reasoning. However, although our findings are consistent with this inhibitory account it is important to bear in mind that the results are also open to alternatives explanations. For example, the right LPFC activation could be associated with the use of an alternative calculation strategy rather than inhibitory processing per se. Likewise, the influential fuzzy-trace account of Reyna and colleagues (see Reyna, 2012, and Helm, Garavito, Rahimi-Golkhandan, & Reyna, 2017, for review) has suggested that reasoning development is characterized by an age related increase in the use of one type of representation over another (i.e., tendency to rely more on gist-based processing rather than verbatim representations). In theory, it is possible that the observed age-related decrease in right LPFC activation is associated with the preferential use of one type of representation over the other.

To conclude, the present study was the first to directly contrast adolescents' and adults' brain activation in a classic heuristics and biases reasoning task. Results point to a developmental right LPFC activation decrease when a prepotent biasing heuristic response is successfully overridden—which is consistent with the hypothesis that the effort of exerting inhibitory control decreases with age. With respect to biased responding our findings fail to find a developmental effect and suggest that biased adolescent and adult reasoners frequently fail to successfully monitor for conflict in the ratio bias task.

Transparency statement

Author contributions

KM and WDN conceptualized and designed the study. FO, GS, KM, and OE, ran the study. KM analyzed the data. GB, NP, OE, OH, and WDN, supervised the project. OH, NP, GS, and WDN, acquired project funding. KM and WDN wrote the manuscript.

Open practices

Acknowledgements

REFERENCES

Aczel, B., Szollosi, A., & Bago, B. (2016). Lax monitoring versus logical intuition: The determinants of confidence in conjunction fallacy. *Thinking & Reasoning, 22*(1), 99–117.

Amsel, E., Klaczynski, P. A., Johnston, A., Bench, S., Close, J., Sadler, E., et al. (2008). A dual-process account of the development of scientific reasoning: The nature and development of metacognitive intercession skills. *Cognitive Development, 23*(4), 452–471.

Bago, B., & De Neys, W. (2017). Fast logic?: Examining the time course assumption of dual process theory. *Cognition, 158,* 90–109.

Bago, B., & De Neys, W. (2018). The smart System 1: Evidence for the intuitve nature of correct responding on the bat-and-ball problem. *Thinking & Reasoning,* 1–43.

Ball, L. J., Thompson, V. A., & Stupple, E. J. (2017). Conflict and dual process theory: The case of belief bias. In W. De Neys (Ed.), *Dual process theory 2.0* (pp. 108–128). Oxon, UK: Routledge.

Barrouillet, P. (2011). Dual-process theories and cognitive development: Advances and challenges. *Developmental Review, 31*(2–3), 79–85.

Bonner, C., & Newell, B. R. (2010). In conflict with ourselves? An investigation of heuristic and analytic processes in decision making. *Memory & Cognition, 38*(2), 186–196.

Casey, B. J., & Caudle, K. (2013). The teenage brain: Self control. *Current Directions in Psychological Science, 22*(2), 82–87.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science, 313*(5787), 684–687.

De Neys, W. (2012). Bias and conflict a case for logical intuitions. *Perspectives on Psychological Science, 7*(1), 28–38.

De Neys, W. (2013). Heuristics, biases, and the development of conflict detection during reasoning. In H. Markovits (Ed.), *The developmental psychology of reasoning and decision making* (pp. 130–147). Hove, UK: Psychology Press.

De Neys, W. (Ed.). (2017). *Dual process theory 2.0.* Oxon, UK: Routledge.

De Neys, W., & Bonnefon, J.-F. (2013). The 'whys' and 'whens' of individual differences in thinking biases. *Trends in Cognitive Sciences, 17*(4), 172–178.

De Neys, W., & Feremans, V. (2013). Development of heuristic bias detection in elementary school. *Developmental Psychology, 49,* 258–269.

De Neys, W., & Glumicic, T. (2008). Conflict monitoring in dual process theories of thinking. *Cognition, 106*(3), 1248–1299.

De Neys, W., Vartanian, O., & Goel, V. (2008). Smarter than we think when our brains detect that we are biased. *Psychological Science, 19*(5), 483–489.

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

Evans, J. S. B. (2007). On the resolution of conflict in dual process theories of reasoning. *Thinking & Reasoning, 13*(4), 321–339.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition advancing the debate. *Perspectives on Psychological Science, 8*(3), 223–241.

Ferreira, M. B., Mata, A., Donkin, C., Sherman, S. J., & Ihmels, M. (2016). Analytic and heuristic processes in the detection and resolution of conflict. *Memory & Cognition, 44*(7), 1050–1063.

Frey, D. P., Bago, B., & De Neys, W. (2017). Commentary: Seeing the conflict: An attentional account of reasoning errors. *Frontiers in Psychology, 8,* 1284.

Frey, D., Johnson, E. D., & De Neys, W. (2018). Individual differences in conflict detection during reasoning. *Quarterly Journal of Experimental Psychology, 71,* 1188–1208.

Gangemi, A., Bourgeois-Gironde, S., & Mancini, F. (2015). Feelings of error in reasoning—in search of a phenomenon. *Thinking & Reasoning, 21*(4), 383–396. https://doi.org/10.1080/13546783.2014.980755.

Goel, V., Buchel, C., Frith, C., & Dolan, R. J. (2000). Dissociation of mechanisms underlying syllogistic reasoning. *Neuroimage, 12*(5), 504–514.

Goel, V., & Dolan, R. J. (2003). Explaining modulation of reasoning by belief. *Cognition, 87,* B11–B22.

Helm, R. K., Garavito, D. M., Rahimi-Golkhandan, S., & Reyna, V. (2017). The fuzzy-trace dual process model. In W. De Neys (Ed.), *Dual process theory 2.0* (pp. 90–107). Oxon, UK: Routledge.

Houdé, O. (2000). Inhibition and cognitive development: Object, number, categorization, and reasoning. *Cognitive Development, 15*(1), 63–73.

Houdé, O., & Borst, G. (2015). Evidence for an inhibitory-control theory of the reasoning brain. *Frontiers in Human Neuroscience, 9,* 148.

Houdé, O., Pineau, A., Leroux, G., Poirel, N., Perchey, G., Lanoë, C., et al. (2011). Functional MRI study of Piaget's conservation-of-number task in preschool and school-age children: A neo-Piagetian approach. *Journal of Experimental Child Psychology, 110,* 332–346.

Johnson, E. D., Tubau, E., & De Neys, W. (2016). The Doubting System 1: Evidence for automatic substitution sensitivity. *Acta Psychologica, 164,* 56–64.

Kahneman, D. (2011). *Thinking, fast and slow.* New York, NY: Farrar, Straus and Giroux.

Leroux, G., Spiess, J., Zago, L., Rossi, S., Lubin, A., Turbelin, M.-R., et al. (2009). Adult brains don't fully overcome biases that lead to incorrect performance during cognitive development: An fMRI study in young adults completing a Piaget-like task. *Developmental Science, 12,* 326–338.

Li, R., Smith, D. V., Clithero, J. A., Venkatraman, V., Carter, R. M., & Huettel, S. A. (2017). Reason's enemy is not emotion: Engagement of cognitive control networks explains biases in gain/loss framing. *Journal of Neuroscience, 37*(13), 3588–3598.

Luna, B. (2009). Developmental changes in cognitive control through adolescence. *Advances in Child Development and Behavior, 37,* 233–278.

Luna, B., Padmanabhan, A., & O'Hearn, K. (2010). What has fMRI told us about the development of cognitive control through adolescence? *Brain and Cognition, 72*(1), 101–113.

Mata, A., & Ferreira, M. B. (2018). Response: Commentary: Seeing the conflict: An attentional account of reasoning errors. *Frontiers in Psychology, 9,* 24.

Mata, A., Ferreira, M. B., Voss, A., & Kollei, T. (2017). Seeing the conflict: An attentional account of reasoning errors. *Psychonomic Bulletin & Review, 24,* 1980–1986.

Mata, A., Schubert, A. L., & Ferreira, M. B. (2014). The role of language comprehension in reasoning: How "good-enough" representations induce biases. *Cognition, 133*(2), 457—463.

Mevel, K., Poirel, N., Rossi, S., Cassotti, M., Simon, G., Houdé, O., et al. (2015). Bias detection: Response confidence evidence for conflict sensitivity in the ratio bias task. *Journal of Cognitive Psychology, 27*, 227—237.

Newman, I., Gibb, M., & Thompson, V. A. (2017). Rule-based reasoning is fast and belief-based reasoning can be slow: Challenging current explanations of belief -bias and base-rate neglect. *Journal of Experimental Psychology Learning, Memory, and Cognition, 43*(7), 1154—1170.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2012). Are we good at detecting conflict during reasoning? *Cognition, 124*(1), 101—106.

Pennycook, G., Fugelsang, J. A., & Koehler, D. J. (2015). What makes us think? A three-stage dual-process model of analytic engagement. *Cognitive Psychology, 80*, 34—72.

Pennycook, G., Trippas, D., Handley, S. J., & Thompson, V. A. (2014). Base rates: Both neglected and intuitive. *Journal of Experimental Psychology Learning, Memory, and Cognition, 40*(2), 544—554.

Prado, J., Chadha, A., & Booth, J. R. (2011). The brain network for deductive reasoning: A quantitative meta-analysis of 28 neuroimaging studies. *Journal of Cognitive Neuroscience, 23*(11), 3483—3497.

Prado, J., & Noveck, I. A. (2007). Overcoming perceptual features in logical reasoning: A parametric fMRI study. *Journal of Cognitive Neuroscience, 19*, 642—657.

Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in Fuzzy-Trace Theory. *Judgment and Decision Making, 7*, 332—359.

Reyna, V. F., Chapman, S. B., Dougherty, M., & Confrey, J. (2012). *The adolescent brain: Learning, reasoning and decision making*. Washington DC: American Psychological Association.

Rudski, J. M., & Volksdorf, J. (2002). Pictorial versus textual information and the ratio-bias effect. *Perceptual and Motor Skills, 95*(2), 547—554.

Santesso, D. L., & Segalowitz, S. J. (2008). Developmental differences in error-related ERPs in middle-to late-adolescent males. *Developmental Psychology, 44*(1), 205.

Simon, G., Lubin, A., Houdé, O., & De Neys, W. (2015). Anterior cingulate cortex and intuitive bias detection during number conservation. *Cognitive Neuroscience, 6*(4), 158—168. https://doi.org/10.1080/17588928.2015.1036847.

Singmann, H., Klauer, K. C., & Kellen, D. (2014). Intuitive logic revisited: New data and a Bayesian mixed model meta-analysis. *PLos One, 9*(4), e94223.

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

Stollstorff, M., Vartanian, O., & Goel, V. (2012). Levels of conflict in reasoning modulate right lateral prefrontal cortex. *Brain Research, 1428*, 24—32.

Stupple, E. J., Ball, L. J., Evans, J. S. B., & Kamal-Smith, E. (2011). When logic and belief collide: Individual differences in reasoning times support a selective processing model. *Journal of Cognitive Psychology, 23*(8), 931—941.

Tamm, L., Menon, V., & Reiss, A. L. (2002). Maturation of brain function associated with response inhibition. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(10), 1231—1238.

Thompson, V. A., & Johnson, S. C. (2014). Conflict, metacognition, and analytic thinking. *Thinking & Reasoning, 20*(2), 215—244.

Travers, E., Rolison, J. J., & Feeney, A. (2016). The time course of conflict on the cognitive reflection test. *Cognition, 150*, 109—118.

Tsujii, T., Masuda, S., Akiyama, T., & Watanabe, S. (2010). The role of inferior frontal cortex in belief-bias reasoning: An rTMS study. *Neuropsychologia, 48*, 2005—2008.

Tsujii, T., & Watanabe, S. (2010). Neural correlates of belief-bias reasoning under time pressure: A near-infrared spectroscopy study. *Neuroimage, 50*, 1320—1326.

Vartanian, O., Beatty, E. L., Smith, I., Blackler, K., Lam, Q., Forbes, S., et al. (2018). The reflective mind: Examining individual differences in susceptibility to base rate neglect with fmri. *Journal of Cognitive Neuroscience, 30*(7), 1011—1022.

Wager, T. D., & Nichols, T. E. (2003). Optimization of experimental design in fMRI: A general framework using a genetic algorithm. *Neuroimage, 18*(2), 293—309.