bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Metacognitive awareness of difficulty in action selection: the role of the cingulo-opercular network

Martyn Teuchies[1], Kobe Desender[1,2], Wouter de Baene[3], Jelle Demanet[1], Marcel Brass[1]

[1] **Department of Experimental Psychology, Ghent University, Belgium**

[2] **Department of Neurophysiology and Pathophysiology, University Medical Center Hamburg-Eppendorf, Germany**

[3] **Department of Cognitive Neuropsychology, Tilburg University, The Netherlands**

Corresponding author at: Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, Belgium. Fax: +32 9 264 64 96.
E-mail address: m.teuchies@ufl.edu (M. Teuchies).

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

The question whether and how we are able to monitor our own cognitive states (metacognition) has been a matter of debate for decades. Do we have direct access to our cognitive processes or can we only infer them indirectly based on their consequences? In the current study, we wanted to investigate the brain circuits that underlie the metacognitive experience of fluency in action selection. To manipulate action-selection fluency we used a subliminal response priming paradigm. On each trial, participants additionally engaged in the metacognitive process of rating how hard they felt it was to respond to the target stimulus. Despite having no conscious awareness of the prime, results showed that participants rated incompatible trials (during which subliminal primes interfered with the required response) to be more difficult than compatible trials (where primes facilitated the required response) reflecting metacognitive awareness of difficulty. This increased sense of subjective difficulty was mirrored by increased activity in the rostral cingulate zone (RCZ) and the anterior insula, two regions that are functionally closely connected. Importantly, these reflected unique activations and were not explained by reaction times or prime-response compatibility. We interpret these findings in light of a possible grounding of the metacognitive judgement of fluency in action selection in interoceptive signals resulting from increased effort.

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1. Introduction (504/650 words max)

To what extent are humans able to monitor their own cognitive processes? Looking at the literature there is some controversy surrounding this question. Some research suggests that humans are poor judges of their own cognitive processes (Nisbett & Wilson, 1977; Wilson & Dunn, 2004; Johansson, Hall, Sikström, & Olsson, 2005), whereas others have shown that humans are remarkably good at monitoring their own cognition, as participants are often aware when they made an error (Murphy, Robertson, Harty, & O'Connell, 2015), and can provide very precise estimates of the probability of being correct (Boldt & Yeung, 2015). This process of self-monitoring is also referred to as metacognition, as it describes insights into our own cognitive processes (Metcalfe & Shimamura, 1994; Brown, 1978). The neuro-cognitive mechanisms that underlie this self-monitoring ability are still poorly understood. Self-monitoring of behavior has been the focus of a number of recent brain imaging studies in the domain of error detection and decision making (Klein et al., 2007; Ullsperger, Harsay, Wessel, & Ridderinkhof, 2010; Fleming, Weil, Nagy, Dolan, & Rees, 2010). The anterior insula has been found to be involved in error detection (Klein et al., 2007; Ullsperger et al., 2010), whereas anterior prefrontal regions were found to be involved in decision confidence (Fleming et al., 2010). More recently, the process by which humans monitor the difficulty in action selection has attracted increasing attention (see e.g. Desender, Van Opstal, & Van den Bussche, 2014; Desender, Van Opstal, Hughes, & Van den Bussche, 2016; Questienne, Atas, Burle, Gevers, 2017). An interesting aspect of action selection is that one can manipulate its difficulty without participants becoming aware of the manipulation. It is known that subliminal response conflict hampers performance: it slows down response speed and it increases error rates (see e.g. Vorberg et al., 2003). By using a subliminal response priming paradigm, one can thus manipulate conflict between two response options outside participants' awareness. Consequently, a metacognitive representation of this manipulation cannot be based on a conscious interpretation of the events (i.e., the visually conflicting information), but has to be based on the interpretation of internal signals caused by these events. Brain imaging studies have revealed that subliminal response conflict is registered in the brain by increased activity

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in both the rostral-cingulate zone (RCZ) and the anterior insula (Teuchies et al. 2016). Even though participants are typically unaware of the subliminal conflict-inducing stimulus, previous research has shown that they nevertheless report increased levels of subjective difficulty when responding to trials with subliminal conflict (Desender et al., 2014; 2016). This raises the question about the source of such metacognitive judgements. Previous research indicates that the experience of subjective difficulty does not simply reflect a read-out of response speed (Desender et al., 2016), but rather seems to be based on motor conflict induced by the subliminal primes (Questienne, Atas, Burle, & Gevers, 2017). This leads to the prediction that the metacognitive judgement is related to brain processes that are involved in subliminal conflict processing, namely the RCZ and the anterior insula (Teuchies et al., 2016).

2. Method Section

2.1 Participants

Participants in this study were 30 Dutch-speaking students from Ghent University (19 female, mean age = 23.77 years, SD = 3.20); each reported as healthy and with no history of neurological, pain, or circulatory disorders and normal or corrected-to-normal vision. All participants gave written informed consent, and the study was approved by the Medical Ethical Review Board of the Ghent University hospital, in accordance with the declaration of Helsinki. All participants were right-handed, as assessed by the Edinburgh Inventory (Oldfield, 1971), and were compensated thirty euros for their participation.

2.2 Stimuli

Stimulus presentation and response registration was done using Tscope software (Stevens, Lammertyn, Verbruggen, & Vandierendonck, 2006). In the scanner room the task was presented using a Brainlogics 200MR digital projector that uses digital light processing (DLP) running at a refresh rate of 60 Hz with a viewing distance of 120 cm. Using DLP it took 1 ms to deconstruct the image on

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the screen allowing our subliminal primes to be presented with great precision. The mean presentation time was 18.00 ms (*SD* = 0.24; range 15.91 – 18.91 ms). Three types of grey colored primes were used: left or right pointing arrows or a neutral prime (which consisted of overlapping left and right pointing arrows). The primes were followed by superimposed metacontrast masks of the same luminance. The metacontrast masks were embedded within target arrows that pointed left or right. Primes subtended visual angles of 0.8° x 1.86°, and the targets 1.09° x 3.47°. Prime and target stimuli could appear randomly above or below a fixation cross at a visual angle of 1.38°. The unpredictable location was included to enhance the masking effect (Vorberg et al., 2003). A circular rating scale was adapted from Kahnt, Heinzle, Park, & Haynes (2011). The x and y coordinates of the mouse response were converted into polar coordinates ranging from 0 degrees (easiest) to 360 degrees (most difficult). The thickness of the scale increased with difficulty. The easiest point on the scale was the tail of the circle; the most difficult point was the thickest point of the circle. The orientation of the scale was randomly chosen on each trial so that the starting point of the scale was unpredictable. This prevented participants from preparing a motor response before seeing the actual scale.

2.3 Procedure

Except for the ratings, the experimental design was identical to Teuchies et al. (2016). Primes were presented for 16.7 ms (1 refreshrate at 60Hz), followed by a blank screen for 33.3 ms and a target that also functioned as a mask. Target duration was 250 ms. The response window was set to 1500 ms. Participants were instructed to respond as fast and accurate as possible to the direction of the target arrows with their left middle finger (left pointing targets) and left index finger (right pointing targets) using an MR compatible response box. If participants failed to respond within this time window, they saw “te laat” (too late) for 1000 ms after the trial. After each response a blank screen was shown for 1500 ms followed by the rating part of the trial during which the rating scale was shown until participants had given their response with their right hand using an MR compatible

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

optical track-ball mouse to select a point on the rating scale that matched their subjective sense of difficulty. The response was registered only when the mouse was actually on the rating scale. Mouse clicks outside of the rating scale were not registered. Participants were instructed to use the entire scale and they were instructed that the extremities of the scale represented their personal most difficult and easiest points. Once they clicked on the scale a blank screen was shown for the inter-trial-interval. The inter-trial-interval was jittered with values ranging between 1000ms and 5250ms. The jitter values followed a distribution with pseudo-logarithmic density (range, 1000–5250 ms, in steps of 250 ms; mean jitter, 2625 ms). Before doing the experiment in the scanner participants carried out two training blocks of 48 trials each. In the first training block they were only presented with the response priming task, without the rating to let them experience the response priming task. When asked, all participants indicated that they made mistakes and that some trials felt more difficult than others. In the second training block the rating was added after every individual trial and participants were instructed to rate on each trial how difficult they found it to respond as fast and accurate as possible to the target stimulus. Participants were never alerted to the possibility of primes being presented. The main task inside the fMRI scanner consisted of three blocks of 72 trials each. Within each block, each prime-response compatibility condition (compatible, incompatible and neutral) occurred equally often. At the end of the task participants were asked whether they noticed anything unusual about the stimuli during the task. None of the participants noticed the primes, but three of them reported seeing a “flash” before the target was presented. These participants were included in the final sample. Following the test phase, participants were explicitly told about the presence of the primes, and performed a prime-visibility test. This test allowed us to check if the prime stimuli were indeed presented subliminally, or not. In this test participants were asked to identify the direction of the primes (left or right) on each individual trial by using the same left and right response buttons as used during the test phase. During this test, participants remained in the scanner, so environment and apparatus were identical to the main experiment. To minimize indirect priming effects on the recognition of the primes, participants were required to respond at least 600

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

ms after the mask was presented. A visual cue ('*') signaled when they were allowed to respond. The test consisted of two blocks of 50 trials each.

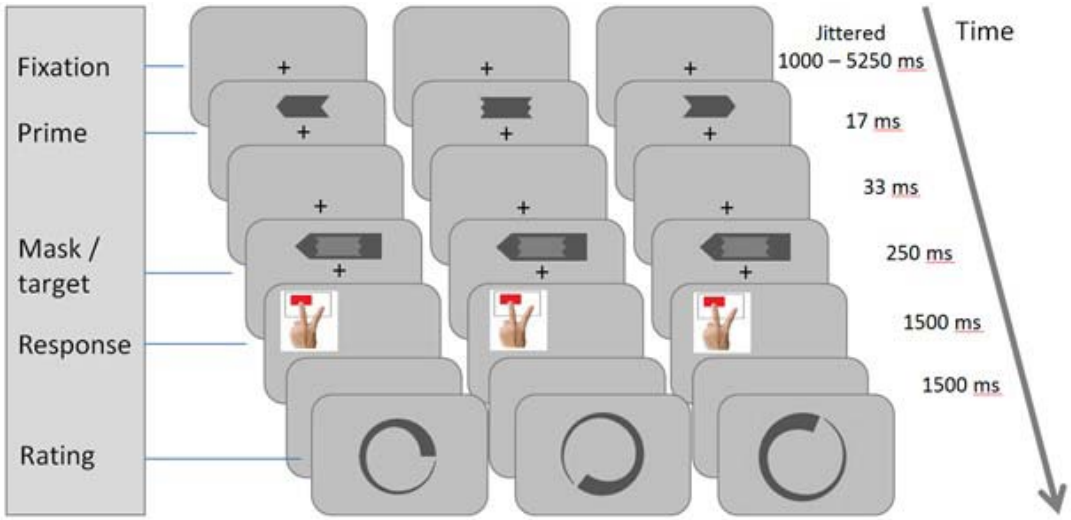

Figure 1. Schematic of an Experimental trial. Three possible combinations of the factor prime-response compatibility (compatible: left panel; neutral: middle panel; incompatible: right panel). Participants were instructed to respond to the target stimuli (with the left hand), and were unaware of the presence of the arrow primes. Primes and targets could appear randomly above or below fixation on each trial. After their response participants indicated their subjective feeling of difficulty using a circular rating scale. The thin tail is the easiest point and the scale continuously increases in thickness and difficulty up to the thick end representing the most difficult point. Participants were instructed to use the whole scale.

2.4 fMRI data acquisition and preprocessing

Data were acquired with a 3T Siemens Magnetom Trio MRI system (Siemens Medical Systems, Erlangen, Germany) using a 32-channel radiofrequency head coil. Participants were positioned head first and supine in the magnet bore. First, 176 high-resolution anatomical images were acquired using a T1-weighted 3D MPRAGE sequence (TR = 2,250 ms, TE = 4.18 ms, TI = 900 ms, image matrix = 256 x 256, FOV = 256 mm, flip angle = 9°, and voxel size = 1 x 1 x 1 mm). Whole-brain functional images were then collected using a T2-weighted echo-planar imaging (EPI) sequence, sensitive to blood-oxygen-level dependent contrast (TR = 2,000 ms, TE = 35 ms, image matrix = 64 x 64, FOV = 224 mm, flip angle = 80°, slice thickness = 3.0 mm, distance factor = 17%, voxel size 3.5 x

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3.5 x 3.0 mm, and 30 axial slices). A varying number of images were acquired per run due to individual differences in choice behavior and reaction times. All data were preprocessed and analyzed using Matlab and the SPM8 software (Wellcome Department of Cognitive Neurology, London, UK). To account for possible T1 relaxation effects, the first four scans of each EPI series were excluded from the analysis. The ArtRepair toolbox for SPM was used to detect outlier volumes concerning global intensity or large scan-to-scan movement (Mazaika, Whitfield-Gabrieli, & Reiss, 2007). First, a mean image for all scan volumes was created, to which individual volumes were spatially realigned using rigid body transformation. Thereafter, they were slice time corrected using the first slice as a reference. The structural image of each participant was co-registered with their mean functional image after which all functional images were normalized to the Montreal Neurological Institute (Montreal, Quebec, Canada) T1 template. Motion parameters were estimated for each session separately. The images were resampled into 3 x 3 x 3 mm voxels and spatially smoothed with a Gaussian kernel of 8 mm (full-width at half maximum). A high-pass filter of 128 Hz was applied during fMRI data analysis to remove low-frequency drifts.

2.5 Behavioral Data Analysis

Mean reaction times (RTs), error rates and subjective ratings were submitted to a repeated-measures ANOVA, with prime-response compatibility (compatible vs. incompatible vs. neutral) as factor. The responses to the primes in the visibility check were categorized using signal detection theory (Green & Swets, 1966). Measures of prime discriminability (d') for each participant were computed. We then used a one-sample t-test to see whether the mean d' of the sample deviated from zero.

2.6 General GLM analyses

The participant-level statistical analyses were performed using the general linear model (GLM). In a first analysis, compatibility conditions (compatible, incompatible and neutral) were modelled in a single regressor of interest and raw subjective rating values for each trial were added

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

as an extra parameter allowing us to look at brain activity related to the raw subjective difficulty ratings. In a second analysis, we wanted to capture variance in brain activity that was *unique* to the subjective difficulty ratings, independent of the variables prime-response compatibility and reaction times (which are both known to affect subjective difficulty ratings; Desender, Van Opstal, & Van den Bussche, 2017; Questienne, Atas, Burle, & Gevers, 2017). In order to capture variance related to compatibility, three different regressors of interest (compatible/incompatible/neutral) were modelled for this variable. In order to look at brain activation uniquely attributed to subjective ratings independent of reaction times (i.e., both variables showed modest negative relation: mean r = -.30, sd = .16, range = -.545 and .011), we regressed out the influence of one variable on the other. Specifically, for each participant we regressed reaction times onto subjective ratings and used the residuals of this regression model to test the brain regions that were sensitive to subjective ratings. The advantage of this approach is that these residuals are perfectly de-confounded from reaction times, so their parametric regressor will only capture variance that is unique to subjective ratings. Henceforth, we will refer to these as *subjective rating residuals*. Similarly, we used the residuals from a regression model with raw subjective ratings as predictor variable, to obtain a regressor for reaction times that was perfectly de-correlated from the raw ratings (henceforth referred to as reaction times residuals). Because the order of the parametric regressors matters (i.e., the second regressor will only capture activation that has not been captured yet), one model was specified in which we first entered reaction time residuals as a parametric regressor and subjective rating residuals as the second parametric regressor and and a second model was specified in which the opposite order was used.

In both analyses, erroneous trials and the first trials of each block were always modeled as separate regressors of no interest (4.9% of the trials). The events of interest were the periods after the onsets of the different targets in the response priming task. Vectors containing the event onsets were convolved with the canonical hemodynamic response function (HRF) to form the main regressors in the design matrix (the regression model). Motion parameters for each individual

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

subject were added. No derivatives were added to the model for this analysis. The statistical parameter estimates were computed separately for each voxel for all columns in the design matrix. Contrast images were constructed for each individual to compare the relevant parameter estimates for the regressors containing the canonical HRF. The group-level random effects analysis was then performed. Using one-sample t-tests we looked at the effects of the subjective difficulty ratings and reaction times across prime-response compatibility conditions. The subjective difficulty ratings and the reaction times had been added as parametric regressors during the first-level analysis. Only clusters corrected for familywise error rate with a peak-level threshold of $p < .05$ are reported. The resulting maps were overlaid onto a structural image of a standard MNI brain, and the coordinates reported correspond to the MNI coordinate system.

2.7 Regions of interest analyses

In the region of interest (ROI) analyses, we focused on the RCZ and the anterior insula as these were our principal ROI's based on our previous study (Teuchies at al., 2016). Accordingly, the peak coordinates were taken from this previous study. To create ROI's we created spheres with a 5mm radius around the peak coordinates of the RCZ [MNI 6 20 43] and the anterior insula [MNI -36 20 -2]. To visualize the effects of rating per compatibility condition, we created a first-level model where we performed a median split on the subjective rating data. This way we ended up with 6 regressors, compatible-easy and compatible-difficult, incompatible-easy and incompatible-difficult and neutral-easy and neutral-difficult, allowing us to directly contrast trials that were rated as easy with trials that were rated as difficult. The percent beta change of each ROI was submitted to a 2 X 3 repeated-measures ANOVA, with factors subjective rating (difficult vs. easy) and prime-response compatibility (compatible vs. incompatible vs. neutral). When a significant main effect of subjective rating or prime-response compatibility (compatible vs. incompatible vs. neutral) was found, the differences between the conditions were analyzed post-hoc using two-tailed paired samples t-tests.

2.8 Mediation analyses

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

As described below, activity in both the RCZ and the anterior insula was related to difficulty ratings. To shed light on the direction of these effects, we performed causal mediation analyses. First, we tested whether the influence of the RCZ on difficulty ratings was mediated by the anterior insula, conditional on congruency. A mediator and an outcome model were fitted on the data using mixed regression modelling, as implemented in the lme4 package (Bates, Maechler, Bolker, & Walker, 2015). Using this type of analysis, both ratings and brain activations can be fit at the single-trial level. Random slopes were added for each variable when this increased the model fit, as assessed by model comparison. For these models, *F* statistics are reported and the degrees of freedom were estimated by Satterthwaite's approximation. A mediator mixed model was fit in which activity in the anterior insula was predicted by activity in the RCZ and congruency. An outcome mixed model was fit in which ratings (after regressing out RTs) were predicted by activity in the anterior insula, the RCZ and congruency. A mediation analysis was then performed on these two models (using the mediation package; Tingley, Yamamoto, Hirose, Keele, & Imai, 2014). This method partitions the total effect on ratings into an indirect effect (i.e. the effect of RCZ on ratings that is mediated by anterior insula) and a direct effect (i.e., correlation between RCZ and ratings that is not explained by anterior insula). If this indirect effect is significant, this is evidence for a significant mediation effect. Note that this latter observation is equivalent to showing that the influence of a direct path decreases when a mediation path is added to the model. Second, we tested the reversed hypothesis that the influence of anterior insula on difficulty ratings was mediated by the anterior insula. For this, the same mediation analysis was run but after exchanging RCZ and anterior insula in all models.

3. Results

3.1 Behavioral Results

3.1.1 Main Task

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Trials where participants did not respond within the 1500 ms response window were removed from the data (0.6% of the trials). For the remaining data, mean RTs on correct trials, mean error rates and mean difficulty judgments on correct trials were submitted to a repeated-measures ANOVA with prime-response compatibility (prime-response compatible vs. incompatible vs. neutral) as factor. For RTs (table 1), this analysis yielded a significant effect of prime-response compatibility, ($F(2, 28) = 39.24$, $p < .001$, $\eta p^2 = .737$). Prime-compatible responses ($M = 426.8$ ms) were significantly faster than prime-incompatible responses ($M = 453.9$ ms; incompatible – compatible = 27 ms; $t(29) = 7.94$, $p < .001$, $d = 0.68$). Prime-compatible responses were not faster than prime-neutral responses ($M = 430.6$ ms; neutral – compatible= 7ms; $t(29) = -1.27$, $p = .22$, d = 0.11), meaning that directional primes did not lead to a significant facilitation effect. There was, however, a significant interference effect, meaning that prime-incompatible responses were slower than responses to neutral primes (incompatible – neutral = 23ms; $t(29) = 7.97$, $p < .001$, $d = 0.62$).

The error rates showed a similar effect of prime-response compatibility, ($F(2, 28) = 12.53$, $p < .001$, $\eta p^2 = .472$). Participants made significantly more errors on prime-incompatible trials ($M = 7.93\%$) than on prime-compatible trials ($M = 2.92\%$; $t(29) = 5.1$, $p < 0.001$, $d = 0.93$) and on neutral trials ($M = 3.84\%$; $t(29) = 4.3$, $p < 0.001$, $d = 0.73$). Error rates were also slightly higher on neutral-prime trials than on prime-compatible trials, but this difference was not significant ($t(29) = -1.7$, $p = 0.1$, $d = 0.28$).

Table 1. Reaction times, percentage of errors and difficulty ratings as a function of prime-action compatibility.

	Reaction time (ms)	Errors (%)	Subjective difficulty rating (°)
Compatible	426.8 (8.1)	2.92 (0.6)	248.7 (9.5)
Incompatible	453.9 (6.8)	7.93 (1.3)	238.4 (9.5)
Neutral	430.6 (6.7)	3.84 (0.6)	249.2 (9.4)

Note: numbers in parentheses show standard errors of the means across participants. The subjective difficulty ratings are reported in degrees ranging from 0 – 359.

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

For the subjective difficulty ratings we also observed a main effect of prime-response compatibility ($F(2, 28) = 9.60$, $p < .001$, $\eta p^2 = .407$). Due to the circular nature of the scale, ratings lie between 0 (difficult) and 360 degrees (easy). Participants rated prime-incompatible trials ($M = 238.4$) as significantly more difficult than prime-compatible trials ($M = 248.7$; $t(29) = 4.1$, $p < 0.001$, $d = 0.20$) and more difficult than neutral trials ($M = 249.2$; $t(29) = -4.3$, $p < 0.001$, $d = 0.21$). Ratings for neutral-prime trials did not differ from ratings for prime-compatible trials ($t(29) = -.29$, $p = 0.77$, $d = 0.01$).

3.1.2 Prime visibility

Based on the data of the prime visibility task, a *d′* value was computed for each participant as an index of prime visibility. The *d′* values were not significantly different from chance level performance (i.e., zero; mean *d′* = 0.077, *sd* = 0.37; one-sample t-test, $t(29) = 1.13$, $p = 0.27$). Thus, it can be concluded that participants show no reliable sign of awareness of the direction of the prime stimuli. Furthermore, when correlating the compatibility effect in the subjective ratings with the individual *d′* values, we found no significant correlation ($r(28) = .12$, $p = .54$), indicating that the subjective ratings were not influenced by prime visibility.

3.2 fMRI

3.2.1 Whole-Brain Analysis Results

In the whole brain analysis, we first looked for brain regions that showed significant activity that was correlated with subjective ratings (Table 2). There were two significant clusters of FWE corrected activation in the RCZ [MNI -3 20 46 and MNI 0 17 52]. This indicates that increased activity in the RCZ is related to increased subjective difficulty.

Although this analysis is an important first step, it cannot unravel whether activity in the RCZ reflects actual variation in difficulty ratings, or whether both are driven by another variable (such as prime-response compatibility or reaction time; Desender, Van Opstal, & Van den Bussche, 2017).

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Therefore, in a second analysis, we looked for brain regions that showed activity uniquely correlated with subjective ratings and for brain regions uniquely related to reaction times. As explained before, we used the residuals from a regression analysis to obtain measures of reaction times and subjective ratings that were independent from each other. This way we could use the subjective rating data while controlling for reaction times and compatibility and vice versa. Looking at the results of the model were reaction times were added first and subjective difficulty ratings were added second (Table 2), on the whole brain level, the parameter of subjective rating residuals revealed significant clusters of FWE corrected activation in the RCZ [MNI 0 17 52] and the left anterior insula [MNI -39 20 1]. The left anterior insula cluster is closely located to the anterior insula [MNI -36 20 -2] that we observed in our previous study (Teuchies et al., 2016). These results indicate that the RCZ and the anterior insula showed increased activity with increased subjective difficulty, independent from prime-response compatibility or reaction times. Looking at the results for the model in which we first added subjective difficulty ratings and reaction times second, the parameter of reaction time residuals revealed one significant FWE corrected cluster of activation in the left inferior parietal cortex [MNI -51 -34 52].

Table 2.

Region	Peak Coordinates (MNI)	*z*-score	Extent
Subjective Ratings			
rostral cingulate zone	-3 20 46	6.21	3*
	0 17 52	6.20	3
Subjective Ratings			
(corrected for compatibility and RTs)			
rostral cingulate zone	0 17 52	5.86	17
Left anterior insula	-39 20 1	4.62	1**

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Reaction Times			
(corrected for compatibility and ratings)			
left postcentral gyrus	-51 -34 52	4.84	8

Active regions on whole-brain level revealed using a 2nd level flexible factorial design, using one-sample t-tests to look at the effects of reaction times and subjective ratings.
** this RCZ cluster consists of 87 voxels when we looked at it using a more liberal threshold of p < .001 uncorrected).
(all regions cluster-level threshold *P* < .05, family-wise error correction, initial voxel level threshold T = 5.63, p < .001).
** the anterior insula cluster consists of 104 voxels when we looked at it using a more liberal threshold of p < .001 uncorrected).

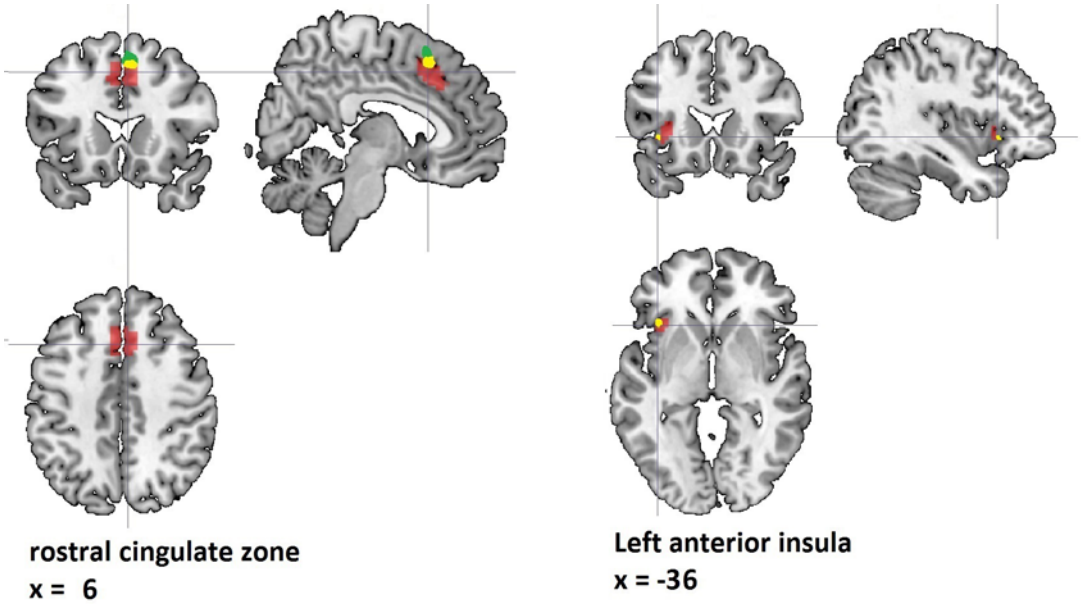

Figure 2. Main areas of interest showing higher activation when the sense of subjective difficulty increased, independent from prime-response compatibility and after regressing out the effect of reaction times. A) Rostral Cingulate Zone. RED = peak coordinate [MNI 6 20 43] obtained from a previous study by Teuchies et al. (2016); GREEN = peak coordinate [MNI 0 17 52] obtained on the whole brain level when looking at one-sample T maps for the parameter of the subjective difficulty ratings; YELLOW = overlap **B) Left Anterior insula.** RED = peak coordinate [MNI -36 20 -2] obtained from a previous study by Teuchies et al. (2016); Yellow = overlap with Anterior insula [-39 20 1], obtained on the whole brain level when looking at one-sample T maps for the parameter of the subjective difficulty ratings. The subjective difficulty ratings were not added as raw data, but we used the residuals from a regression analysis in which we regressed the subjective difficulty ratings on the reaction times in order to filter out the influence of reaction times.

3.2.2 Region of Interest (ROI) Analysis Results

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To avoid double dipping, for the RCZ and the anterior insula we looked at ROI's (Figure 2) based on our previous study (Teuchies et al., 2016). During the whole brain analysis, we found that activity in the RCZ increased when the subjective difficulty residuals increased (Figure 3), independent from prime-response compatibility and reaction times, as the parameter for subjective rating residuals was orthogonal to both these factors. Having found that subjective ratings increased activity independent of prime-response compatibility, we then used the GLM model based on the binned subjective rating residuals (consisting of 6 regressors compatible easy; compatible difficult; incompatible easy – incompatible difficult; neutral easy – neutral difficult) to look at the unique effects of subjective difficulty and prime-compatibility.

Rostral Cingulate Zone

As expected based on the whole-brain analysis we observed a main effect of subjective difficulty residuals in the RCZ (Figure 2, Figure 3) (MNI [6 20 43]; $F(1,29) = 9.974$, $p = 0.004$, $\eta^2 = .256$). Trials that were rated as more difficult led to higher activation than trials that were rated as easy ($t(29) = 3.16$, $p < .01$, $d = .30$). We observed no significant interaction between compatibility and subjective rating residuals ($F(2,28) = 0.81$, $p = 0.46$, $\eta^2 = .055$). Replicating the findings of our previous study we also observed a main-effect of prime-response compatibility, ($F(2,28) = 4.53$, $p < 0.05$, $\eta^2 = .245$), showing that the RCZ was more strongly activated in the incompatible prime-response condition compared to the compatible condition ($t(29) = -2.64$, $p = .013$, $d = .28$) and compared to the neutral condition ($t(29) = 2.17$, $p = .04$, $d = .24$). The compatible and the neutral conditions did not differ in activation ($t(29) = -0.46$, $p = .65$, $d = .04$).

Anterior Insula

In the left anterior insula ROI [MNI -36 20 -1] (Figure 2, Figure 3) we observed a main effect of subjective rating residuals ($F(1,29) = 5.28$, $p = 0.029$, $\eta^2 = .154$). Trials that were rated as more difficult led to higher activation than trials that were rated as easy ($t(29) = 2.27$, $p = .031$, $d = .22$). We observed no significant interaction between compatibility and subjective rating residuals ($F(2,28) =$

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

0.686, p = 0.51, η^2 = .047). Replicating our previous findings we also observed a main effect of prime-response compatibility, (F(2,28) = 4.25, p = 0.024, η^2 = .233). This region was more strongly activated in the incompatible prime-response condition compared to the compatible condition (t(29) = -2.43, p = .021, d = .18) and the neutral condition (t(29) = 2.38, p = .024, d = .29). The compatible and the neutral conditions did not differ in activation (t(29) = -1.69, p = .87, d = .03).

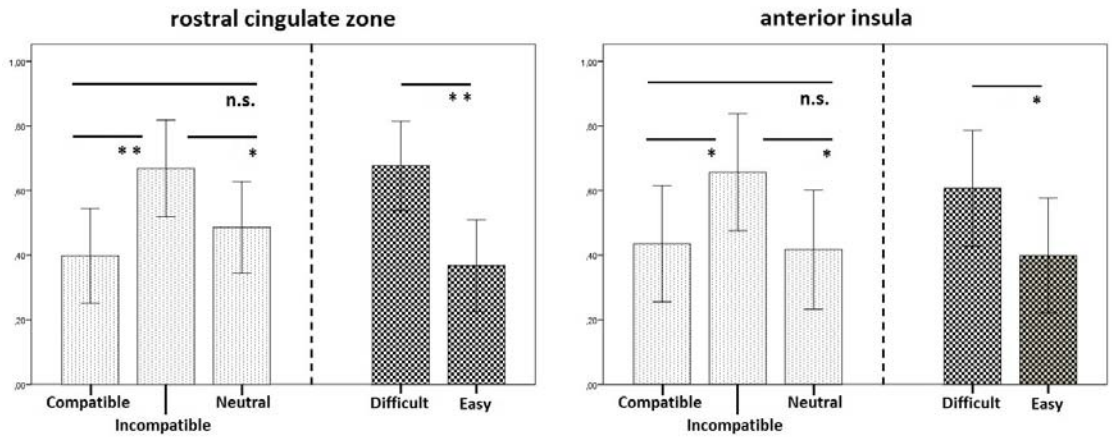

Figure 3. Main effects of compatibility and subjective difficulty in the RCZ MNI [6 20 43] (***left panel***) and the anterior insula [MNI -36 20 -1] (***right panel***). * p < .05; ** p < .01

3.2.3. Mediation analysis

To shed light on the directionality between both brain regions, we performed causal mediation analysis. First, we tested the hypothesis that the influence of the RCZ on ratings is mediated by the anterior insula. A prerequisite for mediation analysis is that all three paths are significant. Indeed, in the mediator model, the RCZ predicted activity in the anterior insula (F(1,28.6) = 287.60, p < .001), and in the outcome model subjective rating residuals were significantly predicted by the anterior insula (F(1,5436) = 8.96, p = .003), and RCZ (F(1,33.3) = 5.25, p = .028). Results of the mediation analyses showed that there was a significant part of the influence of the RCZ on subjective rating residuals that was mediated by activity in the AI (β = -.010, p = .004), whereas there was only a

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

very weak direct effect of the RCZ on subjective rating residuals (β = -.019, p = .050). Because mediation analysis is a correlational technique, we also tested the reversed effect, namely that the influence of the AI on ratings is mediated by the RCZ. This analysis showed a highly significant direct effect of AI on ratings (β = -.015, p = .001), and only a very weak mediation effect by the RCZ (β = -.008, p = .048). In sum, results from the mediation analyses suggest that the influence of the RCZ on ratings is mediated by activity in the AI.

4. Discussion

In the current study, we set out to test which brain regions are involved in self-monitoring difficulty in action selection. To accomplish this, a subliminal response priming task was used that influences the difficulty of action selection. The benefit of this task is that participants remained unaware of the conflict-inducing stimulus itself, thereby eliminating the possibility that participants rated their sense of subjective difficulty based on perceiving the conflict-inducing stimuli or based on how they believe they should respond. In the current study we observed that participants reported increased subjective difficulty on trials in which subliminal response conflict was induced. In a first analysis step, a peak in the RCZ was found that was related to the raw subjective difficulty ratings. In a second step, we then wanted to unravel activity specific to subjective ratings, which was not explained by prime-response compatibility or reaction times. This analysis showed that both the RCZ and the anterior insula were related to unique variation in subjective ratings. Further, replicating previous work, we found that both these regions were increased in the presence of subliminal response conflict. In the remainder of this section, we discuss how these findings are compatible with a grounding of metacognitive experiences of difficulty within interoceptive signals.

4.1. Metacognitive computations of subjective difficulty

The RCZ is part of the anterior cingulate cortex, which is a central hub in the cerebral cortex. This brain region plays a key role in cognitive control processes, as it is argued to be involved in conflict processing (Botvinick, Braver, Barch, Carter, & Cohen, 2001), computing the expected value

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of control (Shenhav, Botvinick, & Cohen, 2013), detecting violations from predictions (Alexander & Brown, 2011), and in effort processing (Walton, Bannerman, & Rushworth, 2002). These different functions have been integrated by positing that the ACC controls the degree of effort invested in a certain task (Holroyd & McClure, 2015). Indeed, focal damage to rat medial frontal cortex decreased the frequency of high effort responses to obtain a reward. Within this framework, the sensitivity of the anterior cingulate to response conflict results from an increased need for effort in difficult trials (i.e., conflict trials). Confirming this prediction and replicating previous work, the current study observed higher activation in the RCZ for incompatible trials compared to compatible and neutral trials. Given that subjective difficulty judgments track response conflict, a relation between subjective difficulty and RCZ was expected. Critically, however, we were able to demonstrate that the relation between difficulty ratings and RCZ was present, even after controlling for the influence of prime-response compatibility and reaction times. This shows that metacognitive computations of subjective difficulty do not merely track experimental manipulations, rather, they are based on brain regions, such as the RCZ, that code for the required degree of effort, over and above that induced by the experimental manipulation.

Whether or not metacognitive computations of subjective difficulty are directly related to RCZ activity or only indirectly remains an open question. One interesting possibility is that the RCZ only codes for the required degree of effort, and this is subsequently implemented by other brain regions. In this regard, it is interesting that apart from the RCZ we also observed that the anterior insula was sensitive to subjective difficulty ratings, over and above the effect of response conflict and reaction times. Note that the anterior insula was not observed in the analysis where we only looked at subjective ratings (i.e., uncontrolled for compatibility and RTs). One explanation for this might be that by factoring out variance from RTs and compatibility from subjective ratings, the remaining variance reflects a more veridical measure of subjective difficulty (e.g., uncontaminated by RTs) and therefore we have more power to detect effects. The anterior insula is a key brain region involved in interoceptive awareness (Craig, 2009; Grupe & Nitschke, 2013; Gu, Hof, Friston, & Fan, 2013).

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Interoception can be described as the sense of the physiological condition of the body, or the perception of sensory events occurring within one's body (Craig, 2002; 2003; Grupe & Nitschke, 2013). The anterior insula is thought to monitor and control internal, embodied states, such as the degree of arousal. Thus, when the RCZ detects the need for increased effort allocation, this might subsequently be implemented by the anterior insula that increases arousal, via interactions with the sympathetic nervous system. Given that humans are typically unaware of their own brain activity (Prinz, 1992), this raises the interesting possibility that metacognitive evaluations of difficulty are based on bodily signals in response to required effort. Thus, when judging whether a trial was easy or difficult, participants might integrate, among other things, their autonomous bodily reactions towards subliminal response conflict (i.e. cardiac acceleration, increased skin conductance; Allen et al., 2016; Hauser et al., 2017) in order to come to a single judgment of difficulty. Indeed, a recent study demonstrated that participants relied on motor activity in their response effectors (i.e., in the thumbs of both hands) when judging the difficulty of a trial (Questienne, Atas, Burle, & Gevers, 2017).

4.3. Domain-general versus domain-specific metacognition

In recent years, the metacognitive evaluation of performance has been tackled from different angles. This has raised the question whether this metacognitive evaluation of performance is supported by a set of domain-general mechanisms, or whether there is domain-specificity. To tackle this question, McCurdy and colleagues (2013) compared metacognition about visual performance with metacognition about memory performance. Although metacognitive performance was correlated across both domains (see also Faivre, Filevich, Solovey, Kühn, & Blanke, 2017; Song et al., 2011), different neural structures were involved in each. Whereas metacognitive performance about visual decisions was related to volume in anterior prefrontal cortex (see also Fleming, Weil, Nagy, Dolan, & Rees, 2010), metacognitive performance about memory decisions was related to the precuneus. In a subsequent study, metacognitive performance about visual decisions was linked to

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

white matter microstructure in the ACC whereas metacognitive performance about memory was linked to white matter microstructure in the inferior parietal lobule (Baird, Cieslak, Smallwood, Grafton, & Schooler, 2014). The current work adds to this debate, by demonstrating that in a different type of self-evaluation, subjective difficulty in decision making, both RCZ and anterior insula are involved. The latter is particularly interesting, because although the anterior insula is critically involved in self-referential processes such as self-awareness (Craig, 2009), previous studies did not, to our knowledge, implicate the anterior insula in the metacognitive evaluation of performance. As such, the current findings lend further support for the domain-specific view of metacognition.

4.4. The influence of reaction times

It has been suggested that reaction times might serve as a proxy for task difficulty (Kiani, Corthell, & Shadlen, 2014; Bryce & Bratzke, 2014). This would imply that participants (partially) base their subjective difficulty ratings on how long they thought it took them to respond. In the current study, no overlap in brain activity was observed when looking at the unique effect of reaction times compared to the unique effect of the subjective ratings. Reaction times, when controlling for prime-response compatibility and subjective difficulty, elicited a small cluster of activation in the postcentral gyrus (Brodmann area 2) which is part of the somatosensory cortex that mostly represents the fingers (see e.g. Kurth et al., 2000; McGlone, Kelly, Trulsson, Francis, Westling, & Bowtell, 2002; Nelson & Chen, 2008). While a clear explanation for this pattern of activation is lacking, one interesting explanation is that perception of time to respond is related to proprioception. This finding is compatible with the idea that motor control is guided by action effects (Greenwald, 1972). Hence we best monitor our behavior by attending to the sensory consequences of the behavior. A study looking into the neural correlates of conscious conflict reported that subjective difficulty ratings correlated with activity in the post-central gyrus as well (Gray et al., 2013). The authors argued that activation in this region correlated with perceptual processing of events that are associated with somatosensory regions of the brain and that urges from conflicting

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

action intentions often intrude into consciousness. However, a major drawback of that study was that the conflict-inducing stimulus was clearly visible to participants, raising the possibility that participants simply based their judgement on the visual conflict.

5. Conclusion

In the current work, we observed that the subjective sense of difficulty is represented in the RCZ and the anterior insula, two regions that are functionally closely connected. Importantly, this was observed when controlling for prime-response compatibility and reaction times. Since the RCZ and the anterior insula typically activate in unison, future research is needed to test our hypothesis that the RCZ codes for the required level of effort and the anterior insula implements this by increasing arousal, which is what participants become aware of.

6. Acknowledgements

The research reported in this paper was funded by the Interuniversity Attraction Poles Program initiated by the Belgian Science Policy Office (IUAPVII/33). K.D. is an FWO [PEGASUS]^2 Marie Skłodowska- Curie fellow (grant number 12T9717N).

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience, 14*, 1338 – 1344.

Allen, M., Frank, D., Schwarzkopf, D. S., Fardo, F., Winston, J. S., Hauser, T. U., Rees, G. (2016). Unexpected arousal modulates the influence of sensory noise on confidence. *eLife, 5*, e18103.

Barch, D. M., Braver, T. S., Sabb, F. W., & Noll, D. C. (2000). Anterior cingulate and the monitoring of response conflict: Evidence from an fMRI study of overt word generation. *Journal of Cognitive Neuroscience, 12,* 298 – 309.

Boldt, A., & Yeung, N. (2015). Shared Neural Markers of Decision Confidence and Error Detection. Journal of Neuroscience, 35(8), 3478–3484.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624 – 652.

Botvinick, M. M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selectionfor-action in anterior cingulate cortex. *Nature, 402*, 179 – 181.

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), *Advances in instructional psychology* (pp. 367 – 406). Hillsdale, NJ: Erlbaum.

Bryce, D., & Bratzke, D. (2014). Introspective reports of reaction times in dual-tasks reflect experienced difficulty rather than timing of cognitive processes. *Consciousness and Cognition, 27,* 254 – 67.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D. C., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science, 280*, 747 – 749.

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Chambon, V., & Haggard, P. (2012). Sense of control depends on fluency of action selection, not motor performance. *Cognition, 125,* 441 – 451.

Cole, M. W. & Schneider, W. (2007). The cognitive control network: Integrated cortical regions with dissociable functions. *Neuroimage, 37,* 343 – 360.

Coles, M. G. H., Gratton, G., Bashore, T. R., Eriksen, C. W., & Donchin, E. (1985). A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 529 – 561.

Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience, 3,* 655 – 666.

Craig A. D. (2003). Interoception: the sense of the physiological condition of the body. *Current Opinion in Neurobiology, 13,* 500 – 505.

Craig, A. D. (2009). How Do You Feel Now? The Anterior Insula and Human Awareness. *Nature Reviews Neuroscience, 10*, 59 – 70.

Deary, I. J., Simonotto, E., Meyer, M., Marshall, A., Marshall, I., Goddard, N., & Wardlaw, J. M. (2004). The functional anatomy of inspection time: an event-related fMRI study. Neuroimage, 22, 1466 – 1479.

Demanet, J., De Baene, W., Arrington, C. M., & Brass, M. (2013). Biasing free choices: the role of the rostral cingulate zone in intentional control. *Neuroimage, 72,* 207 – 213.

Desender, K., Buc Calderon, C., Van Opstal, F., & Van den Bussche, E. (2017). Avoiding the conflict: Metacognitive awareness drives the selection of low-demand contexts. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 1397 – 1411.

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Desender, K., Van Lierde, E., & Van den Bussche, E. (2013). Comparing Conscious and Unconscious Conflict Adaptation. *PLoS ONE, 8,* Article e55976.

Desender, K., Van Opstal, F., Hughes, G., & Van den Bussche, E. (2016). The temporal dynamics of metacognition: Dissociating task-related activity from later metacognitive processes. *Neuropsychologia, 82,* 54 – 64.

Desender, K., Van Opstal, F., & Van den Bussche, E. (2014). Feeling the conflict. *Psychological Science, 25,* 675 – 683.

Devinsky, O., Morrell, M. J., & Vogt, B. A. (1995). Contributions of anterior cingulate to behaviour. *Brain, 118,* 279 – 306.

Dhar, M., Wiersema, J. R., & Pourtois, G. (2011). Cascade of neural events leading from error commission to subsequent awareness revealed using EEG source imaging. *PLoS One* 6:e19578.

Dosenbach, N. U., Fair, D. A., Miezin, F. M., Cohen, A. L., Wenger, K. K., Dosenbach, R. A., et al. (2007). Distinct brain networks for adaptive and stable task control in humans. *Proceedings of the National Academy of Sciences of the United States of America, 104,* 11073 – 11078.

Dunn, B. D., Galton, H. C., Morgan, R., Evans, D., Oliver, C., et al. (2010). Listening to your heart: how interoception shapes emotion experience and intuitive decision making. *Psychological Science, 21*, 1835 – 1844.

Eimer, M., & Schlaghecken, F. (2003). Response facilitation and inhibition in subliminal priming. *Biological Psychology, 64,* 7 – 26.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L (1990) Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In: Brunia CHM, Gaillard AWK, Kok A (eds) Psychophysiological brain research. Tilburg University Press, pp 192–195.

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Feinstein, J. S., Stein, M. B., & Paulus, M. P. (2006). Anterior insula reactivity during certain decisions is associated with neuroticism. *Social Cognitive and Affective Neuroscience, 1,* 136 – 142.

Furstenberg, A., Breska, A., Sompolinsky, H., & Deouell, L. Y. (2015). Evidence of Change of Intention in Picking Situations. *Journal of Cognitive Neuroscience, 27,* 2133 – 2146.

Garfinkel, S. N., & Critchley, H. D. (2013). Interoception, emotion and brain: new insights link internal physiology to social behaviour. Commentary on: Anterior insular cortex mediates bodily sensibility and social anxiety by Terasawa et al. (2012). *SCAN, 8*, 231 – 234.

Gray, J. R., Bargh, J. A., & Morsella, E. (2013). Neural correlates of the essence of conscious conflict: fMRI of sustaining incompatible intentions. *Experimental Brain Research, 229*, 453 – 465.

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience, 4,* 488 – 501.

Gu, X., Hof, P. R., Friston, K. J., & Fan, J. (2013). Anterior Insular Cortex and Emotional Awareness. *Journal of Comparative Neurology, 521*, 3371 – 3388.

Hauser, T. U., Allen, M., Purg, N., Moutoussis, M., Rees, G., & Dolan, R. J. (2017). Noradrenaline blockade specifically enhances metacognitive performance. *eLife, 6*, e24901.

Hester, R., Nestor, L., & Garavan, H. (2009). Impaired error awareness and anterior cingulate cortex hypoactivity in chronic cannabis users. *Neuropsychopharmacology, 34,* 2450 – 2458.

Huettel, S. A., Song, A. W., & McCarthy, G. (2005). Decisions under uncertainty: probabilistic context influences activation of prefrontal and parietal cortices. *Journal of Neuroscience, 25*, 3304 – 3311.

Jahn, A., Nee, D. E., Alexander, W. H., & Brown, J. W. (2014). Distinct regions of anterior cingulate cortex signal prediction and outcome evaluation. *Neuroimage, 95,* 80 – 89.

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Jaskowski, P. (2008). Conscious contributions to subliminal priming. *Consciousness and Cognition, 17,* 72 – 83.

Jiang, J., Zhang, Q., & van Gaal, S. (2015). Conflict awareness dissociates theta-band neural dynamics of the medial frontal and lateral frontal cortex during trial-by-trial Cognitive control. *NeuroImage, 116*, 102 – 111.

Johansson, P., Hall, L., Sikström, S. & Olsson, A. (2005) Failure to detect mismatches between intention and outcome in a simple decision task. *Science, 310,* 116 – 119.

Kahnt, T., Heinzle, J., Park, S. Q., & Haynes, J. D. (2011). Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. *Neuroimage, 56*, 709 – 715.

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron, 84,* 1329 – 1342.

Kiehl, K. A., Liddle, P. F., & Hopfinger, J. B. (2000). Error processing and the rostral anterior cingulate: An event-related fMRI study. *Psychophysiology, 33*, 282 – 294.

Kiesel, A., Wagener, A., Kunde, W., Hoffmann, J., Fallgatter, A. J., & Stöcker, C. (2006). Unconscious manipulation of free choice in humans. *Consciousness and Cognition, 15*, 397 – 408.

Klein, T. A., Endrass, T., Kathmann, N., Neumann, J., Von Cramon, D. Y., & Ullsperger, M. (2007). Neural correlates of error awareness. *Neuroimage, 34*, 1774 – 1781.

Klein, A. T., Ullsperger, M., & Danielmeier, C. (2013). Error awareness and the insula: links to neurological and psychiatric diseases. *Frontiers in Human Neuroscience, 7, article 14*.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review. *Philosophical Transactions of the Royal Society of London B, 362*, 857 – 875.

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Kuipers, M., Richter, M., Scheepers, D., Immink, M. A., Sjak-Shie, E., & van Steenbergen, H. (In press). How effortful is cognitive control? Insights from a novel method measuring single-trial evoked beta-adrenergic cardiac reactivity. *International Journal of Psychophysiology*.

Kurth, R. et al. (2000). fMRI shows multiple somatotopic digit representations in human primary somatosensory cortex. *Neuroreport, 11*, 1487 – 1491.

McGlone, F., Kelly, E. F., Trulsson, M., Francis, S. T., Westling, G., & Bowtell, R. (2002). Functional neuroimaging studies of human somatosensory cortex. Behavioral Brain Research, 135, 147 – 158.

Metcalfe, J., & Shimamura, A. P. (1994). *Metacognition: knowing about knowing*. Cambridge, MA: Massachusetts Institute of Technology.

Naccache, L., Dehaene, S., Cohen, L., Habert, M.-O., Guichart-Gomez, E., Galanaud, D., & Willer, J.-C. (2005). Effortless control: executive attention and conscious feeling of mental effort are dissociable. *Neuropsychologia, 43,* 1318 – 1328.

Nelson, A. J. & Chen, R. (2008). Digit somatotopy within cortical areas of the postcentral gyrus in humans. *Cerebral Cortex, 18*, 2341 – 2351.

Nisbett, R. E., & Wilson, T. D. (1977) Telling more than we can know: verbal reports on mental processes. *Psychological Review, 84*, 231 – 259.

Questienne, L., van Dijck, J.-P., & Gevers, W. (2017). Introspection of Subjective Feelings Is Sensitive and Specific. *Journal of Experimental Psychology: Human Perception and Performance.* Advance online publication.

Orr, J. M., & Banich, M. T. (2014). The neural mechanisms underlying internally and externally guided task selection. *Neuroimage, 84,* 191–205.

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Purmann, S., Badde, S., Luna-Rodriguez, A., & Wendt, M. (2011). Adaptation to frequent conflict in the eriksen flanker task. *Journal of Psychophysiology, 25*, 50 – 59.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron, 79,* 217 – 240.

Silvetti, M., Seurinck, R., & Verguts, T. (2011). Value and prediction error in medial frontal cortex: integrating the single-unit and systems levels of analysis. *Frontiers in Human Neuroscience, 5,* 75.

Singer, T., Critchley, H. D., & Preuschoff, K. (2009). A common role of insula in feelings, empathy and uncertainty. *Trends in Cognitive Sciences, 13,* 334 – 340.

Teuchies, M., Demanet, J., Sidarus, N., Haggard, P., Stevens, M. A., & Brass, M. (2016). Influences of Unconscious Priming on Voluntary actions: role of the Rostral Cingulate zone. *Neuroimage*, *135*, 243 – 252.

Thielscher, A. & Pessoa, L. (2007). Neural correlates of perceptual choice and decision making during fear-disgust discrimination. *Journal of Neuroscience, 27*, 2908 – 2917.

Ullsperger, M., Harsay, H. a, Wessel, J. R., & Ridderinkhof, K. R. (2010). Conscious perception of errors and its relation to the anterior insula. Brain Structure & Function, 214(5–6), 629–43.

Van den Bussche, E., Van den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: A meta-analysis. *Psychological Bulletin, 135*, 452 – 477.

Van Gaal, S., Lamme, V. A. F., & Ridderinkhof, K. R. (2010). Unconsciously triggered conflict adaptation. PLoS ONE, 5, Article e11508.

Van Veen, V., & Carter, C. S., (2002). The timing of action-monitoring processes in the anterior cingulate cortex. *Journal of Cognitive Neuroscience, 14,* 593 – 602.

bioRxiv preprint doi: https://doi.org/10.1101/641340; this version posted May 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Vorberg, D., Mattler, U., Heinecke, A., Schmidt, T., & Schwarzbach, J. (2003). Different time courses for visual perception and action priming. *Proceedings of the national academy of sciences of the United States of America, 100,* 6275–6280.

Wenke, D., Fleming, S. M., & Haggard, P. (2010). Subliminal priming of actions influences sense of control over effects of action. *Cognition, 115*, 26–38.

Weissman, D. H., Roberts, K. C., Visscher, K. M. & Woldorff, M. G. (2006). The neural bases of momentary lapses in attention. Nature Neuroscience, 9, 971 – 978.

Werner, N. S., Peres, I., Duschek, S., & Schandry, R. (2010). Implicit memory for emotional words is modulated by cardiac perception. *Biological Psychology, 85*, 370 –376.

Wessel, J. R. (2012). Error awareness and the error-related negativity: evaluating the first decade of evidence. *Frontiers in Human Neuroscience*, *6*, 88.

Wessel, J. R., Danielmeier, C., Morton, J. B., & Ullsperger, M. (2012). Surprise and error: common neuronal architecture for the processing of errors and novelty. *Journal of Neuroscience, 32,* 7528 – 7537.

Wilson, T. D., & Dunn, E. W. (2004) Self-knowledge: its limits, value, and potential for improvement. *Annual Review of Psycholog, 55*, 493 – 518.

Wisniewsky, D., Goschke, T., & Haynes, J. -D. (2016). Similar coding of freely chosen and externally cued intentions in a fronto-parietal network. *Neuroimage, 134*, 450 – 458.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: conflict monitoring and the error-related negativity. *Psychological Review, 111*, 931 – 959.