Do performance-monitoring related cortical potentials mediate fluency and difficulty effects on decision confidence?

Alan Voodla¹, Andero Uusberg¹

¹ Institute of Psychology, University of Tartu, Näituse 2, Tartu 50411, Estonia

Author Note

This research was supported by the Estonian Research Council grants MOBTP69 and PSG525. We would like to express gratitude to Paula Proos for her help in carrying out the experiment.

Corresponding author: alanvoodla@gmail.com

Word count: 9276

Number of figures: 5

Abstract

It remains unclear to what extent the performance-monitoring (PM) processes involved in perceptual decisions are also involved in more complex reasoning decisions. To address this problem, we examined the role of PM-related cortical potentials in decision confidence. In simple perceptual decisions, the amplitude of these potentials has been shown to correlate with errors on the one hand and confidence ratings on the other. We therefore asked whether PM-related potentials mediate the effects of experimental manipulations of decision confidence on confidence ratings in an arithmetic reasoning task. In an EEG experiment, 49 participants had to quickly decide whether an equation (e.g. $9 * 7 = 65$) was correct or incorrect and then report their confidence in that decision. Task difficulty and response fluency were varied to manipulate confidence. Pooling responses across correct as well as incorrect trials, we analyzed amplitudes of Error-Related Negativity (pooled ERN) and Error Positivity (pooled Pe). The results indicated that although pooled ERN and pooled Pe correlated with task difficulty, neither potential mediated the effect of task difficulty on confidence. Response fluency had an effect on confidence only for simple problems and this effect was partially mediated by pooled ERN but not by pooled Pe. These results suggest that pooled ERN and pooled Pe may reflect implicit performance monitoring, whereas pooled ERN may mediate the effect of the fluency heuristic on decision confidence.

Keywords: decision confidence, mental arithmetic, ERN, Pe, fluency, task difficulty, ACC, performance-monitoring

Introduction

Decision confidence is a metarepresentation monitoring the unfolding of a decision process with respect to its expected outcome (e.g. how sure am I that I picked the right kind of clothes for today's weather under the assumption that I do not want to feel too cold nor too hot?). In this respect, decision confidence is like a nose - although always present in our day-to-day experiences we remain unaware of it unless it is explicitly referenced. And just like a nose has an important function within the respiratory system, decision confidence also has an important function regulating reasoning and adaptive behaviour (Luttrell, Briñol, Petty, Cunningham, & Díaz, 2013). Research into decision confidence dates back to the seminal years of experimental psychology (e.g. Peirce & Jastrow, 1884) and has been more recently re-established mainly within the study of perceptual decision-making (Yeung & Summerfield, 2012) and abstract reasoning (Griffin & Tversky, 1992; Shynkaruk & Thompson, 2006). Research within these two contexts has not always been well integrated (Cruz, Arango-Muñoz, & Volz, 2016). For instance, experimental studies of perceptual decision making (using paradigms such as the the Flanker task, Stroop task, dot count task) suggest that electroencephalographic event-related potentials associated with performance-monitoring are important neural correlates of confidence computations (Boldt & Yeung, 2015; Selimbeyoglu, Keskin-Ergen, & Demiralp, 2012). This suggests that PM-ERPs should also correlate with confidence in more complex tasks where decisions require reasoning, which have been mostly studied in the domain of abstract reasoning (e.g. deductive/inductive inference or numeric reasoning). Surprisingly, the prediction that PM-ERPs could underlie decision confidence has not yet been extended to more complex reasoning problems. The current study helps to synthesize the perceptual decision-making and abstract reasoning perspectives by reconceptualizing error-related EEG activity more widely as PM-ERPs and asking whether they mediate experimentally induced changes on decision confidence in a complex reasoning task.

Decision confidence

We begin with an overview of decision confidence, its functions and its relation to the neural correlates of performance monitoring. Decision confidence can be defined as a belief about the validity of our own thoughts, knowledge or performance that might be consciously accessible as a subjective feeling (Grimaldi, Lau, & Basso, 2015; Meyniel, Sigman, &

Mainen, 2015). This feeling has been termed feeling of rightness for high confidence (Thompson, Turner, & Pennycook, 2011) or feeling of error for low confidence (Gangemi, Bourgeois-Gironde, & Mancini, 2015). Decision confidence is often measured as a confidence rating on a discrete or continuous scale after a participant has made a decision in a perceptual judgement or reasoning task (Grimaldi et al., 2015).

The construct of decision confidence implies a distinction between object- and metalevel psychological processes (Flavell, 1979). Object-level processes refer to psychological activity that results in phenomena such as perceptions, decisions, emotions, and memories which generally represent bodily or environmental states. For example, thinking that the correct answer to the product of 15 and 7 is 105 results from object-level processing. Meta-level processes are a kind of re-representation - they represent states of object-level processes for purposes such as cognitive control and self-regulation (Botvinick, Braver, Barch, Carter, & Cohen, 2001). In the previous example, being confident that the provided answer of 105 is correct results from meta-level processing.

A key function of meta-representations may be to aid adaptive decision-making. For example, feeling of rightness is taken to be a starting or stopping rule for deliberate reasoning (Thompson et al., 2011) in the sense that if a reasoner feels that she is correct in her answer (high confidence), she will no longer try to re-evaluate her decision or find flaws in it. On the other hand, feeling of error and low confidence can lead to increased effort and persistence.

For decision confidence to be a useful signal, it needs to be sufficiently accurate. If decision confidence is not well calibrated, it can lead to suboptimal behaviors such as overconfidence in case of unjustified feeling of rightness and anxiety in case of unjustified feeling of error. An important question from this perspective is, are our confidence estimates a good guide to the quality of our decisions? For example - can we detect if we make errors in Flanker or similar forced-choice tasks? This relationship between object-level performance and its metacognitive evaluation has been termed metacognitive accuracy (Fleming & Dolan, 2012). The general finding seems to be that humans are reliable detectors of their own errors in many perceptual decision tasks and they often exhibit compensatory behaviour after a mistake has been committed (Yeung & Summerfield, 2012). However, this is not true for all tasks. For example, metamemory research shows that study participants can be at odds with reality when indicating their certainty about how well they performed in a memory test (Koriat & Bjork, 2005). Similarly, meta-cognitive accuracy can be low in reasoning tasks,

where participants are asked to evaluate their confidence after solving more complex problems, such as deductive reasoning problems (e.g. Premise 1: Some elephants are ill-mannered. Premise 2: Some ill-mannered beings are smart. Conclusion: Therefore some elephants are smart; Thompson & Shynkaruk, 2006)

The diverging meta-cognitive accuracies of simple and complex decisions raises the question whether the computation of confidence for these decisions might rely on different mechanisms. Although decision confidence might be accurate for simple perceptual judgments, in which participants can easily use heuristic rules to respond, it might be different in cases where simple heuristics are not available or applicable and giving an adequate response requires step-by-step algorithmic reasoning. For instance, Thompson and Shynkaruk (2006) observed that the confidence ratings were influenced by believability of the conclusion, perceived difficulty of the problem and the time taken to make a decision. More recently it has been shown that the ease with which the answer comes to mind (response fluency) also has an effect on decision confidence in more complex decision problems (Thompson et al., 2011). This indicates that participants rely on indirect heuristic information when reporting their confidence level and if this information is unavailable, they might have to rely on another mechanism. These results therefore raise a question about the differences in computational mechanisms of confidence in simple and complex decisions.

One approach to investigate confidence computation differences in simple and complex decisions is through electroencephalographic analysis, which provides a window to underlying neural mechanisms of these processes (Yeung & Summerfield, 2012). If the same mechanism underlies decision confidence for simple and difficult decision-making, then a similar pattern of neural correlates should be observed for their computation. However, if, as the behavioural level analysis suggests, decision confidence computation for complex decisions relies on different neural mechanisms than for simple decisions, a different neural pattern should be observed. Performance monitoring related EEG potentials provide a way to address this hypothesis by offering a window into the neural mechanisms that are known to be involved in driving decision confidence during simple perceptual decisions.

Performance monitoring related potentials

PM-ERPs have been primarily studied as correlates of error-monitoring in perceptual decision-making (Yeung, Botvinick, & Cohen, 2004). In a typical study, authors compare

participants' brain activity recorded during correct and incorrect responses in a speeded forced choice 2-alternative task (for example, evaluating which one of the two displayed batches has more dots in it). This approach has revealed two error-related EEG potentials - error-related negativity (ERN) and error-positivity (Pe) (Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993).

ERN is characterized as a short-lived (75 - 150 ms) negative spike in the EEG signal beginning at the time of, or slightly before committing an erroneous response (van Veen & Carter, 2006). The most likely neural origin for ERN is the anterior cingulate cortex (ACC; Yeung et al., 2004). The functional role of ACC activity reflected in ERN in cognition is of active research interest. According to three main theories, ERN a) reflects a mismatch between expected and observed representations; b) is a signature of two incompatible streams of information; or c) indicates a negative feedback signal for learning (van Veen & Carter, 2006). As the three theories are not necessarily incompatible, it is likely that the ERN can reflect contributions from different functional processes.

Following ERN, within 200-400 ms after a commission of an error, an error-related positive potential (Pe) has been identified (van Veen & Carter, 2006). Although its functional role is more ambiguous than in the case of ERN, it has been associated with conscious processing of error (as opposed to unconscious processing for ERN; Overbeek, Nieuwenhuis, & Ridderinkhof, 2005). In a recent account of ERN and Pe, it has been argued that ERN is most clearly associated with error-monitoring whereas Pe is an index of error correction (Yeung & Summerfield, 2012).

In addition to errors, both ERN and Pe have been found to correlate with decision confidence in perceptual judgement tasks. For instance, ERN amplitude was found to covary with perceived accuracy for a given response in a Flanker task (Scheffers & Coles, 2000). In a modified Flanker task and a visual discrimination task where task difficulty was manipulated, Pe, as well as ERN were observed to correlate with decision certainty (Selimbeyoglu et al. 2012). In another study, in which participants had to evaluate the number of dots and report their confidence on a 6 point scale in each trial, it was observed that both ERN and Pe were on average distinguishable for 6 levels of reported confidence - higher amplitude Pe was associated with higher confidence ratings (Boldt & Yeung, 2015). Interestingly, this association was independent of objective accuracy. This study suggests that

Pe, rather than ERN is the most robust neural index of subjective decision confidence, although for both a clear relationship should be observable.

Whereas the previous studies have mostly emphasized commission of errors as the target of analysis, our focus is more generally on confidence, regardless of whether the response given was objectively correct or incorrect. The idea of error-related potential implies that only trials in which an error is committed are analyzed to detect the effect. However if ERN and Pe really reflect the underlying process of performance monitoring which correlates with decision confidence, then a correlation between ERN, Pe and decision confidence should be observed for both correct and incorrect trials. This interpretation is also supported by the presence of a ERN-like negative deflection when a participant gives a correct answer (Correct Related Negativity or CRN; Gehring et. al, 2012). For the current study we therefore defined pooled ERN as an early frontocentral event-related potential for both correct and incorrect trials. We also defined pooled Pe as a late frontoparietal event-related potential in both correct and incorrect trials.

The present study

The existing literature suggests that for simple perceptual decision tasks, PM-ERPs such as ERN and Pe relate to decision confidence. However, it is unclear whether these cortical potentials, and the neural processes they reflect, play a role in more complex decisions that require reasoning. To fill these gaps, we tested whether pooled ERN and pooled Pe relate to decision confidence in a more complex mental arithmetic task that includes explicit manipulations of decision confidence through task difficulty and response fluency. We assume that if pooled ERN and pooled Pe truly represent neural responses that underlie decision confidence, they should mediate the effects of systematic manipulations of confidence on confidence ratings for different levels of task difficulty and response fluency. If, however evidence for mediation is not detected, then confidence in complex decisions can be said to rely on neural mechanisms not reflected in PM-related ERPs than confidence in simple decisions.

To test these predictions, we designed an EEG experiment in which participants had to quickly decide whether an equation (e.g. $9 * 7 = 65$) is correct or incorrect and then report their confidence in that decision. Task difficulty and response fluency were used as manipulations of decision confidence, which we expected to have a robust effects based on

previous literature (Cruz et al., 2016; Thompson et al., 2013) and a pilot study. Task difficulty was manipulated by presenting equations where one of two multiplicands was either a double-digit number (hard, e.g. $6 \times 12 = 72$) or single-digit number (easy, e.g. $6 \times 6 = 36$). Response fluency was manipulated by the correctness of the displayed equation. We assumed that displaying a correct answer (high fluency, e.g. $6 \times 6 = 36$) facilitates identifying it as the correct answer due to higher cognitive availability whereas displaying an incorrect answer (low fluency, e.g. $6 \times 6 = 38$) interferes with identifying the true correct answer. To further magnify the fluency effect, we presented the answer of the equation as a short prime before displaying the full equation. We hypothesized that the behavioural effects of difficulty and fluency on decision confidence should be mediated by pooled ERN and pooled Pe.

Method

The primary objective of the study design was to test whether the effects of difficulty and fluency on confidence are mediated by PM-ERPs. In order to do this, the experiment had a 2x2 within-subject factorial design with decision confidence and PM-ERPs as the dependent variables (DV). Task difficulty and response fluency were both treated as 2-level independent variables (IV). The main research questions and analyses were pre-registered in *Open Science Foundation* online-environment, the preregistration and other study-related files are retrievable from <https://osf.io/cj2rf/>. The analyses conducted in the results section are consistent with the ones pre-registered except for the last mediation model.

Participants

A priori power analysis for repeated-measures ANOVA with 2 factors indicated that 80% power for an alpha level of 0.05 for a large effect size ($\eta^2 = 0.138$) would be provided by a sample size of 52. Anticipating data loss, we invited 55 volunteers (ages between 18 and 65, 24 males) to register to the study through online invitations distributed in Facebook and mailing lists of University of Tartu. 6 participants had to be excluded from the main analyses due to technical issues with EEG recording. The final sample consisted of 49 participants (age range 18-65, 20 males). As compensation, participants were given personality feedback and an opportunity to win a 20-euro bookshop gift card. Students of psychology could also earn course credits. The project was approved by the Research Ethics Committee of the University of Tartu (application 286/T-3).

Design, task and stimuli

We designed a mental arithmetic task that would simultaneously satisfy the following criteria: be suitable for objective measurement of accuracy; require symbolic as opposed to perceptual processing; allow for simple manipulations of task difficulty and response fluency; and enable gathering data over many trials.

In the experimental task, a participant sitting behind a computer was required to quickly respond whether equations (e.g. $6 \times 8 = 42$) displayed for 2 seconds were correct or incorrect by clicking on a respective button on the upper right and left corners of the screen using a computer mouse (Figure 1). After giving a response, the participant was presented with a visual-analogue scale for 4 seconds to indicate her decision confidence at the time of responding. If a participant did not respond within the timeframe, her answer was counted as incorrect, the confidence question was omitted and the program displayed the next trial. Each trial was initiated when a participant clicked on a button at the center of the lower edge of the screen. The button placements were designed to elicit mouse trajectories which are not analyzed in this paper. Before the onset of the full equation, its' answer was displayed for 0.5 seconds as part of the response fluency manipulation. The task was implemented with Psychopy (Peirce et al., 2019).

As we did not counter-balance the positions of incorrect and correct response buttons, there is a risk that confidence ratings were influenced by which button was pressed prior to giving the rating. We used a linear mixed model regression of confidence ratings to estimate the size of this effect while controlling for other experimental effects. We found that the button placement induced a less than a 2-point shift on the 100 point VAS scale ($\beta = 1.76$, $p < 0.05$) which is substantially smaller than the effects of task difficulty (~ 20 points) and fluency under low difficulty (~ 8 points, see Results). This indicates that the button position cannot account for the majority of the experimental effects of fluency and difficulty.

Figure 1. Trial structure. Participants clicked on the white box to start the sequence. Next they saw the answer of the upcoming equation as part of the response fluency manipulation. Participants then had 2 seconds to indicate the correct response to the equation and finally 4 seconds to report their decision confidence. The choice options at the top of the “Correct or incorrect?” screen are “INCORRECT” on the left and “CORRECT” on the right. The question above the confidence scale translates as “At the time of responding, how certain were you in Your choice?”.

Decision confidence was measured on a visual-analogue scale from 0-100 with 4 labels at the points of 0, 33, 66 and 100 (Figure 1). The question above the scale asked, “At the time of responding, how certain were You in Your choice?”. The labels of the scale were named starting from 0: “Very unconfident”, “Rather unconfident”, “Rather confident” and “Very confident”. We also recorded whether the provided answers were correct and what was the response speed for each answer.

We used two independent variables - task difficulty and response fluency - to experimentally elicit variance in decision confidence. Task difficulty was manipulated by changing one multiplicand for half of the equations to a double-digit number (e.g. $16 * 7 =$

102 as opposed to $6 * 7 = 42$ for easy condition). The response fluency manipulation relied on the fluency heuristic whereby the ease with which a mental representation is activated is used as a cue to indicate the truth value of that representation (Alter & Oppenheimer, 2009). In the present context, we assumed that the common multiplication problems we presented can often be solved by recalling the correct answer from memory, rather than calculating it on the spot. We further assumed that in this process, people rely on the fluency heuristic to decide which of the several response options that come to mind is the correct one. In the experiment, the manipulation of fluency was designed to interfere with this natural use of the fluency heuristics. Specifically, we expected that presenting a response option on the screen (e.g. 36 at the end of $6 \times 6 = 36$) increases the fluency of this option relative to all other options that come to mind. In case the presented option is correct (e.g. 36), it should bias the heuristic decision process towards recognizing this as the correct answer. In case the presented option is incorrect (e.g. 34), the fluency manipulation is expected to bias the heuristic decision process away from identifying the actual correct option and toward erroneously inferring that the presented option is correct due to its higher fluency. To further amplify the impact of the presented response option, and to assure it can interfere with the decision process from its onset, we presented the response options also as primes 0.5 seconds before the problem onset.

Each participant responded to a total of 160 two-multiplicand equations that were divided into 4 blocks so that each block would have equal distributions of difficulty and fluency levels. For the easy condition, the equations were generated by picking numbers from 4, 6, 7, 8, and 9 for both multiplicands. For the hard condition, one multiplicand was selected from a number from 16 to 19 and the other was either 4, 6, 7, 8 or 9. For half of the trials, the displayed answers were correct and incorrect for the other half. The incorrect answers were computed by adding or subtracting 2 from the correct answer so that the incorrect answer would begin with the same number as the correct one. After each block, participants were given true feedback about the number of correct answers in the previous block and deceptive feedback about the percent of participants from a study conducted in the USA that their result exceeds. The deceptive feedback was designed to maintain task motivation by assuring that the result ranged between 50 and 90 %. Between the 4 blocks, participants also reported their emotional state and tiredness on visual-analogue scales.

Procedure

Prior to coming to the experiment, participants filled in online questionnaires measuring Big-5 personality (Konstabel et al., 2017) and Need for cognitive closure (NFCC; Redi, 2017; Webster & Kruglanski, 1994). The data from the questionnaires are not analyzed in this paper. The experimental procedure for each participant took on average about 1 hour and 15 minutes. The experiments took place in the Psychology Institute of University of Tartu. Participants arrived at the experiment room, were asked to sign an informed consent which had also been sent to them earlier, and were introduced to the general setup of the experiment and the basic idea of EEG. The EEG cap was then fit, electrode holders were filled with conductive gel and the electrodes were attached. After setting up the electrodes, each participant was positioned to sit 80 cm from a LCD monitor. They were then instructed to solve multiplication problems as fast and as accurately as possible. Participants were shown how to respond to the equations with a mouse click. The visual-analogue confidence evaluation scale was explained with a reminder that they should report the confidence they felt at the time of responding to the equation rather than at the time of the confidence rating. During the experiment, participants were instructed to take breaks between 4 experimental blocks and were left alone in the room. The experimentator checked in on the participant and the EEG set-up midway through the task. The experiment ended with participants filling in a short follow-up questionnaire about their opinion of the task, research question and response strategies. Participants were then briefed about the aims of the project and provided with a possibility to wash their hair.

EEG data acquisition

To capture PM-ERPs (pooled ERN, pooled Pe), we used BioSemi ActiveTwo system with 32 electrodes referenced to two electrodes placed on earlobes. To eliminate eye movement noise from the signal, four electrodes were placed bipolarly on each participant vertical and horizontal ocular locations. BioSemi Actiview software was used to record the EEG signal with DMS/DRL reference scheme and 512 Hz sampling rate.

EEG preprocessing was conducted in EEGLAB (Brunner, Delorme, & Makeig, 2013) and Matlab. All scalp channels were re-referenced to linked earlobes. Independent Component Analysis was used to computationally remove ocular artifacts. An ICA mixing matrix was obtained by running the Infomax algorithm on a training data consisting of

4-second epochs (-2000 to 2000 ms relative to the response) of 1 Hz high-pass filtered data cleaned of noisy channels and epochs using EEGLAB automatic algorithms. Components corresponding to blinks and eye movements were identified visually and removed from the mixing matrix before using it to reconstruct the unfiltered continuous data. The ICA-pruned data were then low-pass filtered at 30 Hz and cut into epochs from -400 to 1000 ms relative to the response. Segments with artifacts were automatically excluded based on the $\pm 100\mu\text{V}$ threshold criterion. If any single channel would have been exclusively responsible for removing more than 2% of trials, the channel was removed. All removed channels were later spherically interpolated. After preprocessing, on average 67.1% (SD = 15.1%) EEG data in each design cell was retained for each participant. The mean voltage between -400 and -200 was subtracted as a baseline. We used this interval as a baseline, because participants responded using mouse clicks as opposed to single button presses. As a consequence, we expect that decision formation time extended further into the time preceding the mouse button press than it would extend in a typical study using key presses. We followed the same rationale, together with existing literature and visual inspection of the current waveforms to define pooled ERN as the mean voltage at Fz between -70 and 60 ms for both correct and incorrect trials. The pooled Pe was defined as mean voltage at Cz between 200 and 300 ms after a response was given for both correct and incorrect trials (Figure 2).

Figure 2. Grand average response-locked ERPs for electrodes Cz and Fz and scalp plots of pooled ERN and pooled Pe. Panel A shows grand-average activity for electrodes Fz and Cz. We defined pooled ERN as frontocentral activity from -70 to 60 captured by electrode Fz (the blue line) and pooled Pe as frontoparietal activity captured by electrode Cz (the red line). On panel B the scalp plot depicts activity for all electrodes in the interval for pooled ERN (from -70 to 60 ms). Panel C depicts activity for all electrodes in the interval for pooled Pe (from 200 to 300ms).

Results

R environment for statistical computing (R Core Team, 2013) with RStudio interface (RStudio Team, 2015) was used to conduct all the statistical analyses. We used packages *psycho* (Makowski, 2018) and *lmerTest* (Kuznetsova et.al, 2015) for mixed models and package *mlm* for multilevel mediation models (Vuorre & Bolger, 2018).

How did difficulty and fluency influence decision confidence?

The aim of the first analyses was to make sure that fluency and difficulty had indeed influenced decision confidence, response times and accuracy as expected. In total, erroneous responses were given on 27% of the trials (1785 out of 6488), the average error rate for individual participants was 33.1 (sd = 12.8). A mixed-model logistic regression with fluency and difficulty as fixed and participant as a random factor indicated that fluency ($\beta = 0.21$), difficulty ($\beta = -0.83$) and their interaction ($\beta = 0.57$) had a statistically significant effect on response accuracy (all $p < 0.001$). The interaction effect arose because responses were more accurate on easy trials with high compared to low fluency ($p < 0.001$) whereas for hard trials, the fluency effect was reversed, so that low fluency condition resulted in more accurate responses (Figure 2A). Out of all the erroneous responses, most errors were committed in the hard difficulty and high fluency trials (777) and less in hard difficulty but low fluency trials (470). For easy condition 130 errors were committed in high and 408 in low fluency trials.

For response times, a linear mixed-model with fluency and difficulty as fixed and participant as a random factor showed a large main effect of difficulty ($\beta = 0.41$, SE = 0.01, 95% CI [0.39, 0.43], $t(6431) = 33.27$, $p < .001$) and a small main effect of fluency ($\beta = 0.20$, SE = 0.01, 95% CI [0.18, 0.22], $t(6431) = 16.42$, $p < .001$). A small interaction effect between fluency and difficulty was also detected ($\beta = -0.21$, SE = 0.02, 95% CI [-0.25, -0.18], $t(6431) = -11.84$, $p < .001$). Further inspection indicated that response times were slower for difficult trials and high fluency condition accelerated responding only for easy trials (Figure 2B).

In line with the previous results, the mixed model for confidence ratings showed a large main effect of difficulty ($\beta = -24.71$, SE = 1.82, 95% CI [-28.27, -21.16], $t(141) = -13.58$, $p < .001$, $\beta = -24.71$, std. SE = 1.82) and a small significant effect of fluency ($\beta = 8.08$, SE = 1.82, 95% CI [4.52, 11.63], $t(141) = 4.44$, $p < .001$, std. $\beta = 8.08$, std. SE =

1.82). We also found a large and statistically significant interaction effect between fluency and difficulty ($\beta = -8.89$, $SE = 2.57$, 95% CI $[-13.91, -3.87]$, $t(141) = -3.45$, $p < .001$, std. $\beta = -8.89$, std. $SE = 2.57$) which indicated that the manipulation of fluency was only effective in easy trials (Figure 2C). These findings indicate that behavioural level manipulations of task difficulty and response fluency generally showed expected effects on response accuracy, response time and confidence ratings. Unexpectedly we found a reversed effect of fluency on response accuracy in hard trials and no effect of fluency on response time and confidence ratings on hard trials.

Figure 3. Mean response accuracy (A), response times (B) and confidence ratings (C) on y-scale compared in two levels of trial difficulty (on x-scale) and fluency (colour of the lines). Confidence bars indicate 95% confidence intervals.

How did task difficulty and response fluency influence PM-related potentials?

Next, we analyzed whether manipulations of difficulty and fluency had an effect on pooled ERN and pooled Pe using 2 mixed-effect models. For pooled ERN, we found a small statistically significant main effect of difficulty ($\beta = 0.52$, $SE = 0.15$, 95% CI [0.22, 0.83], $t(5163) = 3.39$, $p < .001$; Figure 4, panel A) and a small significant main effect of fluency ($\beta = 0.38$, $SE = 0.15$, 95% CI [0.07, 0.68], $t(5162) = 2.44$, $p < .05$; Figure 4, panel B). There was also a small and statistically significant interaction effect between fluency and difficulty ($\beta = 0.48$, $SE = 0.15$, 95% CI [0.18, 0.79], $t(5161) = 3.13$, $p < .01$; Figure 3, panel A).

The second mixed model indicated that difficulty had a small statistically significant effect on pooled Pe ($\beta = -0.39$, $SE = 0.19$, 95% CI [-0.76, -0.02], $t(5160) = -2.07$, $p < .05$; Figure 4, panel D) but no statistically significant main effect of fluency was detected ($\beta = 0.20$, $SE = 0.19$, 95% CI [-0.17, 0.57], $t(5160) = 1.05$, $p > 0.1$; Figure 4, panel E). The interaction effect between fluency and difficulty was small and not statistically significant ($\beta = 0.35$, $SE = 0.19$, 95% CI [-0.02, 0.72], $t(5159) = 1.85$, $p = 0.06$; Figure 3, panel B). These findings resonate with behavioural results suggesting that our manipulations had expected effects on pooled ERN (for which the average response was least negative in easy and fluent trials) and to a lesser extent on pooled Pe.

Figure 3. Effects of difficulty and fluency on pooled ERN (panel A) and pooled Pe (panel B). Error bars around point estimates depict 95% confidence intervals.

Figure 4. Difficulty and fluency effects on pooled ERN (-60 to 70 ms, electrode Fz) on panels A,B,C and pooled Pe (200 to 300 ms, electrode Cz) panels D, E, F. Shaded area around average lines depicts standard error around the mean. Panel A visualises main effect of difficulty, panel B main effect of fluency and panel C interaction effect of fluency for easy trials on pooled ERN. Panel D depicts main effect of difficulty, panel E main effect of fluency and panel C interaction effect of fluency for easy trials.

Did PM-related potentials relate to decision confidence?

For estimating the total variation in confidence ratings which can be explained by PM-ERPs, we conducted repeated-measures correlation analyses separately for pooled ERN and pooled Pe with *rmcorr* R package (Bakdash & Marusich, 2017). Repeated-measures correlation coefficient (r_{rm}) is similar to Pearson r (it takes values from 0-1, with 0 representing no association and 1 perfect linear relationship), except that it is more suitable for repeated measures design, when there is a need to take into account between-subject variation. Results indicated that pooled ERN had a small significant positive association with confidence ratings ($r_{rm}(5158) = 0.064$, $p < 0.001$) whereas pooled Pe had a small nonsignificant negative association with confidence ratings ($r_{rm}(5158) = -0.086$, $p = 0.532$). The relationship of pooled ERN and confidence is consistent with our predictions, but it is nevertheless very small so in order to gain a better understanding of these associations, we used single trial mediation analysis.

Did PM-related potentials mediate the difficulty and fluency effects on decision confidence?

We ran four multilevel mediation models with R package *bmlm* (Vuurde & Bolger, 2018) to determine, whether the experimental manipulations of difficulty and fluency on confidence ratings would be mediated by PM-related potentials. A Bayesian multilevel mediation analysis determines whether the effect of an independent variable (IV) on a dependent variable (DV) can be explained by a third intervening variable M (Vuurde & Bolger, 2018) while accounting for the hierarchical structure of the data (e.g. trials are nested in participants). In the context of this study, this translates to whether the effects of fluency (IV) and difficulty (IV) on confidence (DV) are mediated by pooled ERN or pooled Pe (M) on a trial-by-trial level. Each mediation model was estimated with 4 Markov Chain Monte Carlo chains with 10 000 iterations for each chain. In each model, confidence ratings were treated as the dependent variable, difficulty (easy/ hard) or fluency (high/ low) as the independent variables and pooled ERN or pooled Pe as the mediators. To quantify the effects, we use non-standardized path coefficients (me for indirect effect and c' for the direct effect of the independent variable on the dependent variable). An effect is detected if the 95% credible intervals for the coefficient do not contain 0. We also use pme (proportion effect mediated) which indicates the proportion of total effect explained by the mediating variable ranging

from 0 to 1 (Vuorre & Bolger, 2018).

We first used a pair of mediation models to ask whether the effect of difficulty on decision confidence was mediated by its effect on pooled ERN or pooled Pe. We found that pooled ERN did not mediate the effect of difficulty on confidence (indirect effect = -0.08, 95% CI = [-0.18, 0.01], proportion mediated effect = 0.00, 95% CI = [0.00, 0.01]). . Likewise, pooled Pe also did not mediate the effect of difficulty on confidence ratings (indirect effect = -0.00, 95% CI = [-0.05, 0.05], proportion mediated effect = 0.00, 95% CI = [0.00, 0.00]). These findings indicate that although the difficulty manipulation had an effect on both confidence ratings as well as PM-related potentials, these effects were largely unrelated. In other words, difficult effect on confidence was not mediated through difficulty effects on pooled ERN or pooled Pe.

The next analysis asked whether the effects of fluency on decision confidence were mediated by PM-related potentials. As the prior analyses indicated that fluency did not have an effect on decision confidence for the hard trials and that fluency did not have an effect on pooled Pe, we conducted this mediation analysis only for the easy trials and pooled ERN. This analysis indicated that the effect of fluency on confidence was partially mediated by pooled ERN (indirect effect = -0.19, 95% CI = [-0.40, -0.03], proportion mediated effect = 0.02, CI = [0.00, 0.05]). The direct effect remained significant after taking pooled ERN into account (c' = -8.03, 95% CI = [-10.32, -5.74]) (Figure 5). This model demonstrates that the fluency manipulation had an effect on pooled ERN (Figure 5), which in turn had an effect on decision confidence (b = 0.11, 95% CI = [0.015, 0.2]). Although the effect was small, this result corresponds to the idea that the fluency effect on decision confidence is partially mediated by pooled ERN.

Figure 5. Path diagram for the mediation model for pooled ERN for easy trials. Point estimates (a, b, c') describe direct effects with associated 95% credible intervals. “SD” shows the effect’s standard deviation which indicates the degree of variation between participants.

Discussion

The aim of the current work was to better understand how performance monitoring related potentials relate to decision confidence. To do this, we set up an arithmetic reasoning experiment with simultaneous EEG recording in which we could systematically manipulate two known determinants of decision confidence: task difficulty and response fluency. Our goal was to see whether the effects of these manipulations on confidence ratings would be mediated by PM-ERPs (pooled ERN, pooled Pe).

We first confirmed that manipulation of task difficulty indeed had an expected effect on decision confidence, so that confidence ratings were on average lower in the hard condition and higher in the easy condition. The effect of task difficulty was further validated

by analyses of reaction times and response accuracy - easy equations were solved on average more accurately and with less time than hard equations. For fluency, an effect on confidence was detected only in low but not in hard condition. As expected, confidence ratings on easy trials were higher in high fluency condition and lower in low fluency condition. This dynamic was also supported by response times and accuracy - in fluent and easy trials participants tended to give faster and more accurate responses.

We also found a reversed effect of fluency on accuracy for hard trials, so that low fluency resulted in higher accuracy. This was an unexpected finding that might relate to the heuristic rules that participants tended to use. According to the post-task questionnaire, participants approached hard trials (e.g. $16 * 7 = 114$) by testing whether the product of the second digit in the first number and the second multiplicand (e.g. $6 * 7 = 42$) ends with the same number as the displayed solution of the equation (114). In case of a mismatch, the equation is very often incorrect. However, if the numbers match (e.g. the answer would be 112), the outcome would still often be a wrong answer. This means that the heuristic is only useful for incorrect (low fluency trials) and therefore should produce more correct responses for low fluency condition.

The unexpected interaction between fluency and difficulty notwithstanding, the behavioral data indicated that our experimental manipulations were successful in systematically altering decision confidence levels. We therefore proceeded to the main question of this study asking whether the effects our manipulations had on confidence were mediated by pooled ERN and/or pooled Pe. We set up 4 hypotheses to quantify the extent to which pooled ERN and pooled Pe mediate the effect of fluency and difficulty on confidence ratings. Our first pair of hypotheses were that if pooled ERN and pooled Pe underlie the computation of decision confidence, a manipulation of task difficulty on confidence should be mediated by pooled ERN and pooled Pe. We found no support for these hypotheses.

The second pair of hypotheses asked whether the fluency effect on confidence would be mediated by pooled ERN and pooled Pe. As our behavioral analyses indicated no effect of fluency in hard task condition, we tested this hypothesis only on easy trials. We found evidence for partial mediation of fluency by pooled ERN on confidence. For pooled Pe, however, there was no evidence for a mediated relationship.

The finding that response fluency but not task difficulty effects on confidence were mediated by PM-ERPs has interesting implications for understanding the origins of

metacognitive representations such as confidence across different types of decisions. PM-related potentials have been associated with the anterior cingulate cortex (ACC) which is thought to signal when mental resources (e.g. working memory) should be allocated for more flexible behaviour (Botvinick et al., 2001; Thompson et al., 2011; Shackman et al., 2011). Past studies of perceptual decisions had suggested that PM-related potentials may indeed reflect a performance monitoring system that detects conflicts and errors which manifests as lowered sense of confidence (Yeung & Summerfield, 2012). We were interested in whether the portions of ACC reflected in PM-related potentials perform a similar function in more complex reasoning tasks. More specifically, if pooled ERN and pooled Pe reflected monitoring for cognitive control, we should have observed a pattern whereby both difficulty and fluency effects on confidence were mediated by pooled ERN and/or pooled Pe. The data from this study, however, indicates that this is not the case. Instead, we found that PM-related potentials may contribute to confidence only for easy decisions that can be addressed using type-1 processes such as heuristics rather than type-2 processes such as reasoning (Evans & Stanovich, 2013).

A potential explanation for this starts from the behavioural level analysis which indicated that fluency only had an effect on decision confidence in easy trials. A similar difference in the effectiveness of fluency manipulation on confidence has also been previously observed in a mathematical reasoning task (Cruz et al., 2016). The authors argue that fluency only influences feelings of error up to a point where the effort to solve the problem disrupts the link between fluency and confidence. In the present context, fluency may have similarly ceased to be a usable signal for the performance monitoring systems for equations with double-digit multiplicands that can rarely be solved by recalling a stored engram of the equation from memory. This hypothesis is in line with the evidence from the mediation models with fluency as the independent variable. The model indicated that the effect of fluency on decision confidence was partially mediated by pooled ERN. That is, pooled ERN amplitude was more negative in low fluency condition which in turn was manifested in lower confidence ratings. This is consistent with the idea that for simple decisions, confidence is an outcome of a system that monitors pattern-matching correspondence and is subject to fluency effects, whereas for more complex decisions this mechanism is no longer solely responsible for confidence computations, because heuristic

responding strategy is unavailable. Although pooled Pe shows a similar pattern in two conditions of fluency (Figure 4, panel F), this effect was not found to be significant.

Despite not observing that PM-related potentials would mediate task difficulty effects on confidence, we did observe that task difficulty influenced pooled Pe, which has also been associated with error-awareness (Yeung & Summerfield, 2012). The contrasts between these findings generates a puzzle: on the one hand, difficulty has an effect on both pooled Pe and decision confidence, but these effects seem to be unrelated in the sense that although pooled Pe indexes difficulty, it does not in turn contribute to confidence ratings. This indicates that although the effect of difficulty on confidence is not mediated by pooled ERN nor pooled Pe, it is possible that pooled Pe and pooled ERN could be associated with implicit monitoring in the sense that they do not contribute to the feeling of confidence. These processes may thus be inaccessible for verbal reports (Shea et al., 2014), but could nevertheless have an adaptive regulatory function over behavioural responses. A remarkable example of this is a neurological patient, with damaged ACC, who did not report increased conscious effort with increased task difficulty, although she clearly allocated more resources to the task when it was harder to solve (Naccache et al., 2005).

To sum up, our findings do not support a simplistic model whereby pooled ERN and pooled Pe would reflect performance monitoring that becomes manifest in confidence ratings across all types of decisions. Instead, the findings point toward a more nuanced model in which pooled Pe and ERN may implicitly monitor task difficulty in a way that does not become accessible for verbal reports unless the decision relies on heuristic selection between mental representations.

Contributions, limitations and further questions

Elucidating the neural mechanisms for decision confidence is important for understanding metacognition and cognitive control. A relevant practical example is a study in which brain-computer interface which provided participants feedback about ERN and Pe resulted in 21% more accurate responses in a visual discrimination task (Parra, Spence, Gerson, & Sajda, 2003). The current study in which we tried to untangle the behavioural level effects on confidence by theoretically relevant neural activity patterns contributes toward understanding our reasoning and decision-making in multiple ways.

First, the results suggest that at the current state of knowledge, the implicit

measurement of decision confidence through PM-ERPs is more tractable for relatively simple decisions for which error is clearly definable as a representational conflict. This does not necessarily imply that confidence in more difficult decisions cannot be implicitly measured. Rather the concept of difficulty warrants more rigorous behavioural level analysis, which could be then adequately mapped onto relevant neural mechanisms (Krakauer, Ghazanfar, Gomez-Marin, MacIver, & Poeppel, 2017). In the current task, it could be argued that for difficult problems, participants had a hard time using a matching heuristic in hard task condition (Does my intuitive answer to $5 * 5$ correspond to presented answer e.g 27?). Currently, difficulty can be thought of as an emergent feature in reasoning process, which could arise for various reasons (too much information, not enough time, too many necessary operations, strategy selection, etc). This notion is therefore highly variable for different tasks and the question “What makes a problem difficult?” warrants a detailed analysis in each setting, in order to capture this on the neurobiological level. It is possible that a similar study design would benefit from a more specific manipulation of confidence since the current results indicate that the partial mediation effect of pooled ERN is relatively small compared to the total effects of difficulty and fluency.

It could be argued that the inferences drawn from the study are limited because the response buttons “Correct” and “Incorrect” at the bottom of the screen were not counterbalanced between trials and therefore when participants gave

Another question inspired from these results concerns the relationship between negative affect and confidence. It has been shown that negative affect is associated with greater ERN and Pe (Hajcak, McDonald, & Simons, 2004; Shackman et al., 2011). Accordingly, ERN and Pe could be interpreted as signals that convey information about uncertainty which in the case of high probability of error is experienced as a negative feeling and should result in low ratings of confidence. However, as the data from the current study suggest, even if negative affect is manifested in confidence ratings, it is not always mediated by pooled ERN or pooled Pe. Therefore in future studies, it would be interesting to see if Pe or ERN would mediate difficulty effects on negative affect. This avenue of research should be further investigated, as it is often taken for granted that negative epistemic feelings are associated with low confidence ratings.

The main contributions of the current study are the discovery that fluency effect on confidence is partially mediated by pooled ERN for simple decisions and that ERN and pooled Pe track task difficulty but do not mediate difficulty effects on confidence ratings.

These results set constraints on the development of metacognitive theory of decision-making and reasoning, by arguing that performance-monitoring in the brain works in different ways for easy and difficult problems. The effect of fluency on pooled ERN also provides a window into manipulating ERN for various further research questions. For example, it would be important to replicate this result in a simple perceptual decision-making task, to see whether the effect of fluency would be more pronounced.

The study also extends the understanding of the relationship between decision confidence and PM-ERPs in arithmetic reasoning. This is important, because it shows the validity of the claim, that pooled ERN and pooled Pe are related to confidence computation in general and not only for specific decision-making tasks. To further extend these claims, it would be beneficial to address the ecological validity of the decision-making task, use various reasoning tasks and consider more nuanced manipulations of decision confidence in future studies.

In summary, current study provides evidence that ACC activity reflected in performance-monitoring related cortical potentials mediates confidence in simple pattern-matching decisions, but not in decisions which require more complex reasoning. Pooled ERN was found to index decision confidence for simple decisions and both, pooled ERN and pooled Pe can be thought of as as implicit monitors for cognitive control. As these are *post-hoc* explanations, further experiments should be designed to test them.

References

- Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and social psychology review*, 13(3), 219-235.
- Bakdash, J. Z., & Marusich, L. R. (2017). Repeated measures correlation. *Frontiers in Psychology*, 8, 456.
- 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(3), 624.
- Brunner, C., Delorme, A., & Makeig, S. (2013). Eeglab—an open source matlab toolbox for electrophysiological research. *Biomedical Engineering/Biomedizinische Technik*.
- Cruz, A. L. F., Arango-Muñoz, S., & Volz, K. G. (2016). Oops, scratch that! Monitoring one’s own errors during mental calculation. *Cognition*, 146, 110–120.
- 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.
- Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, 78(6), 447–455.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American Psychologist*, 34(10), 906.
- Fleming, S. M., & Dolan, R. J. (2012). The neural basis of metacognitive ability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1338–1349.
- Gangemi, A., Bourgeois-Gironde, S., & Mancini, F. (2015). Feelings of error in reasoning—in search of a phenomenon. *Thinking & Reasoning*, 21(4), 383–396.
- Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, 4(6), 385–390.
- Gehring, W. J., Liu, Y., Orr, J. M., & Carp, J. (2012). The error-related negativity (ERN/Ne). *Oxford handbook of event-related potential components*, 231–291.
- Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology*, 24(3), 411–435.
- Grimaldi, P., Lau, H., & Basso, M. A. (2015). There are things that we know that we know, and there are things that we do not know we do not know: Confidence in decision-making. *Neuroscience & Biobehavioral Reviews*, 55, 88–97.
- Hajcak, G., McDonald, N., & Simons, R. F. (2004). Error-related psychophysiology and negative affect. *Brain and Cognition*, 56(2), 189–197.
- Konstabel, K., Lönnqvist, J.-E., Leikas, S., Velázquez, R. G., Qin, H., Verkasalo, M., & Walkowitz, G. (2017). Measuring single constructs by single items: Constructing an even shorter version of the “Short Five” personality inventory. *PloS One*, 12(8), e0182714.

- Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(2), 187.
- Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience Needs Behavior: Correcting a Reductionist Bias. *Neuron*, 93(3), 480–490. <https://doi.org/10.1016/j.neuron.2016.12.041>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package 'lmerTest'. R package version, 2(0).
- Luttrell, A., Briñol, P., Petty, R. E., Cunningham, W., & Díaz, D. (2013). Metacognitive confidence: A neuroscience approach. *Revista de Psicología Social*, 28(3), 317–332.
- Makowski, D. (2018). The psycho package: An efficient and publishing-oriented workflow for psychological science. *Journal of Open Source Software*, 3(22), 470.
- Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian probability: From neural origins to behavior. *Neuron*, 88(1), 78–92.
- 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(9), 1318–1328.
- Overbeek, T. J., Nieuwenhuis, S., & Ridderinkhof, K. R. (2005). Dissociable components of error processing: On the functional significance of the Pe vis-à-vis the ERN/Ne. *Journal of Psychophysiology*, 19(4), 319–329.
- Parra, L. C., Spence, C. D., Gerson, A. D., & Sajda, P. (2003). Response error correction—a demonstration of improved human-machine performance using real-time EEG monitoring. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2), 173–177.
- Peirce, C. S., & Jastrow, J. (1884). *On small differences in sensation*.
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., ... Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 1–9.
- Redi, K. (2017). *Vajadus kognitiivse selguse järele küsimustiku adapteerimine ja seosed sotsiaalse soovitususega* (Thesis, Tartu Ülikool). Retrieved from <https://dspace.ut.ee/handle/10062/57298>
- R Core Team. (2013). *R: A language and environment for statistical computing*.

- RStudio Team. (2015). RStudio: integrated development environment for R. RStudio. *Inc., Boston, MA, 14.*
- Scheffers, M. K., & Coles, M. G. (2000). Performance monitoring in a confusing world: error-related brain activity, judgments of response accuracy, and types of errors. *Journal of Experimental Psychology: Human Perception and Performance*, 26(1), 141.
- Selimbeyoglu, A., Keskin-Ergen, Y., & Demiralp, T. (2012). What if you are not sure? Electroencephalographic correlates of subjective confidence level about a decision. *Clinical Neurophysiology*, 123(6), 1158–1167.
- Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, 12(3), 154.
- Shea, N., Boldt, A., Bang, D., Yeung, N., Heyes, C., & Frith, C. D. (2014). Supra-personal cognitive control and metacognition. *Trends in Cognitive Sciences*, 18(4), 186–193.
- Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, 79(2), 217–240.
- Shynkaruk, J. M., & Thompson, V. A. (2006). Confidence and accuracy in deductive reasoning. *Memory & Cognition*, 34(3), 619–632.
- Thompson, V. A., Turner, J. A. P., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology*, 63(3), 107–140.
- Thompson, V. A., Turner, J. A. P., Pennycook, G., Ball, L. J., Brack, H., Ophir, Y., & Ackerman, R. (2013). The role of answer fluency and perceptual fluency as metacognitive cues for initiating analytic thinking. *Cognition*, 128(2), 237–251.
- van Veen, V., & Carter, C. S. (2006). Error detection, correction, and prevention in the brain: a brief review of data and theories. *Clinical EEG and Neuroscience*, 37(4), 330–335.
- Vuorre, M., & Bolger, N. (2018). Within-subject mediation analysis for experimental data in cognitive psychology and neuroscience. *Behavior Research Methods*, 50(5), 2125–2143.
- Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, 67(6), 1049.
- 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(4),

931.

Yeung, N., & Summerfield, C. (2012). Metacognition in human decision-making: confidence and error monitoring. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1594), 1310–1321.