FACULTY OF PSYCHOLOGY
AND EDUCATIONAL SCIENCES

PREDICTING EFFORT AVOIDANCE BASED ON DECODED NEURAL DIFFICULTY JUDGMENT

Word count: 10.011

Selin Kanatli

Student number: 01305291

Supervisor(s): Dr. Kobe Desender, Dr. Mehdi Senoussi, Prof. Dr. Tom Verguts

A dissertation submitted to Ghent University in partial fulfilment of the requirements for the degree of Master of Theoretical and Experimental Psychology

Academic year: 2019 – 2020

GHENT UNIVERSITY

Preamble concerning COVID-19

We had planned to start the EEG data collection in March 2019, specifically the pilot study. The rest of the data collection was planned between October and December 2019. Next, the meetings for the analysis and decoding were planned from March 2020 until May 2020.

The data collection was finished before COVID-19, so COVID-19 did not influence this part. However, due to COVID-19, we were not able to meet in person to discuss the final data and the next steps such as the analysis, decoding and writing the thesis.

Since the data was collected before COVID-19, the outbreak did not change the plan of the study that much. It was a little bit more challenging to execute the analysis and understand the decoding through online meetings, but in the end, it worked.

This preamble was drafted in consultation with the student and the promotor and was approved by both.

Abstract

Previous studies have shown that cognitive control is related to effort choices. People tend to minimise the effort they put into a task and often choose less effortful tasks. It has also been shown that metacognitive difficulty is related to cognitive control and avoiding cognitive control. The present studies tried to predict effort avoidance based on different neural correlates (i.e., N2 or objective difficulty and P3 or subjective difficulty). A paradigm was used where participants performed a masked priming task followed by them deciding the subjective difficulty for that trial (i.e., rating phase). In a subsequent choice phase, the participants had to choose whether they wanted to carry out this task in a low effort – low gain or high effort – high gain context. The ERP components showed modulation by objective (i.e., congruency) and subjective difficulty (i.e., rating). The decoding analysis, however, showed that we could not predict a participant's following choice based on the neural data of congruency and subjective difficulty ratings.

Corona preambule

We hadden gepland om in Maart 2019 te starten met de eeg-afnames, vooral de pilotstudies. De rest van de eeg-afnames zouden tussen oktober en december 2019 gebeuren. De meetings voor de analyse en de decoding waren gepland tussen maart en mei 2019.

De eeg-afnames waren afgerond voor COVID-19, dus dit had geen effect op de datacollectie. Maar door COVID-19 was het niet mogelijk om samen te komen met mijn begeleiders voor de volgende stappen zoals de analyse, decoding en het bespreken van de thesis zelf.

Aangezien de data verzameld was voor COVID-19, is de studie niet moeten aangepast worden. Het was wat moeilijker om de bespreking van de analyse en het uitleggen van de decoding via online meetings te doen, maar uiteindelijk ging dit wel.

Deze preambule werd in overleg tussen de student en de promotor opgesteld en door beide goedgekeurd.

Abstract

Studies hebben aangetoond dat cognitieve controle gerelateerd is aan moeite. Mensen gaan vaak de moeite die ze in een taak moeten steken minimaliseren en vaker de taken kiezen die minder moeite vereisen. Het is ook aangetoond dat metacognitieve moeilijkheid gerelateerd is aan cognitieve controle en het vermijden van cognitieve controle. Deze studie heeft geprobeerd om het vermijden van moeite te voorspellen gebaseerd op verschillende neurale correlaten (i.e., N2 of objectieve moeilijkheid en P3 of subjectieve moeilijkheid). In dit paradigma moesten de participanten een masked priming taak uitvoeren, gevolgd door het beslissen van de subjectieve moeilijkheid van de trial (i.e., rating fase). In de volgende keuze fase, moesten de participanten beslissen of ze dezelfde taak in een lage moeite – lage winst of hoge moeite – hoge winst context wouden uitvoeren. De ERP componenten toonden modulatie door objectieve (i.e., congruentie) en subjectieve moeilijkheid (i.e., rating). Echter, de decoding analyse toonde aan dat we een participants volgende keuze niet konden voorspellen gebaseerd op de neurale data van de congruentie of subjectieve moeilijkheidsratings.

Table of contents

Introduction

An interesting aspect of the human cognitive system is its ability to control its own cognitive processes. This is usually referred to as cognitive control. For example, when a person who grew up in Belgium walks down the street and wants to cross the street, they will automatically look to the left, because in Belgium the cars drive on the left lane. This is something that is taught as a child and thus has become an automatism. However, when that same person visits the United Kingdom, they have to take extra caution because there the cars drive on the other side of the road. So, that person now has to suppress the urge to look to the left and force themselves to look to the right instead. This is what we call cognitive control. The cognitive system needs to use cognitive control to suppress its natural tendency to look towards the left and instead look to the right. Cognitive control is usually required in situations where there is conflict. For example, when being in the United Kingdom, the conflict between wanting to look to the left in the United Kingdom, but having to look to the right. The cognitive control system is necessary for selecting and successfully monitoring behaviour that facilitates the achievement of a goal (Botvinick, Braver, Barch, Carter, & Cohen, 2001).

Conflict monitoring and cognitive control

To study cognitive control in the lab, researchers usually examine how participants deal with conflict. Like the conflict of having to look to the right but wanting to look to the left in the example. Researchers typically use cognitive conflict. The Stroop Task is a typical example of such a conflict task. In the Stroop task, participants are presented with colour words presented in different colours. The participants are instructed to say the ink colour that the word is presented in. These colours can be congruent ("BLUE" is presented in blue) or incongruent ("BLUE" is presented in red") with the semantic meaning of the word. Studies have shown that participants are significantly faster when the colours are congruent with the meaning of the word, compared to when the colours are incongruent (Cohen, Dunbar, & McClelland, 1990; Liotti, Woldorff, Perez, & Mayberg, 2000). The usual explanation for this effect is that reading can be regarded as an automatic process. Because humans are very much trained in reading words, when presented with a word, humans automatically read the meaning of the word without the need for active control. When both the ink colour and the meaning of the word are congruent, this is not a problem for naming the ink colour. However, on the incongruent trials, one needs to actively suppress the urge to read the word and say the colour of the word instead. So, we are faster when executing tasks with automatic processes (e.g., reading a word). Compared to tasks with non-automatic processes (e.g., saying the colour of the word) (Balota & Marsh, 2004; Schneider & Shiffrin, 1977).

An important question then is how the brain knows when it needs to exert cognitive control. How does the brain know when it has to look right, and not left? To say the colour, and not read the word? According to the conflict monitoring theory, there is a system in the brain that is sensitive to conflicts in information processing. This system in the brain is usually referred to as the conflict monitor (Botvinick et al., 2001). This theory further states that when the conflict monitor detects a conflict in the processing of information, it sends a signal to remote brain regions that implement increases in cognitive control (Gratton, Coles, & Donchin, 1992; Kerns et al., 2004; Stürmer, Leuthold, Soetens, Schröter, & Sommer, 2002). As a consequence, after the conflict monitor detects the need for control, the strength of control that is necessary for the situation or task is applied. Thus, by checking and monitoring our internal state, cognitive control allows us to shift our attention in an appropriate fashion (Botvinick, Cohen, & Carter, 2004). Previous research has proposed that there is a specific brain region involved in conflict monitoring, namely the dorsal anterior cingulate cortex (dACC) (Botvinick, 2007; Botvinick et al., 2001; Kerns et al., 2004; LaBerge, 1990; Shenhav, Botvinick, & Cohen, 2013). Although the involvement of the dACC in conflict monitoring is usually observed in research using fMRI, researchers have also managed to locate the conflict monitor in EEG signals. Specifically, conflict monitoring is supposed to be reflected in a frontocentral N2 component. The N2 component is generally observed around 200-300 ms post-stimulus. This is thought to reflect the activation to conflicting stimuli (Veen & Carter, 2002). A larger N2 is observed when we are observing conflict (Botvinick et al., 2004). The N2 occurs pre-response as a consequence of stimuli (Yeung, Botvinick, & Cohen, 2004) and it is larger when the stimuli are incongruent compared to when they are congruent

(Heil, Hennighausen, Osman, Wiegelmann, & Rolke, 2000; Liotti et al., 2000). The ACC also shows higher activation when there are more response possibilities, when the response is less determined (Barch, Braver, Sabb, & Noll, 2000). So, there is evidence that N2 is modulated by task-related conflict (Desender, Van Opstal, Hughes, & Van den Bussche, 2016).

Importantly, activating the ACC to perform cognitive control comes with a cost. This cost is the cognitive demand of manipulating our cognitive control (Botvinick, 2007). Botvinick and Rosen (2009) used a demand selection task to specify this cost. In their study, participants had to choose between two coloured decks of cards. One side of the card was orange or green, the other side of the card had a number on it, either in purple or in blue. If the number was purple, they had to make a parity judgement (i.e., they had to say whether the number was even or not). If the number was blue, they had to make a magnitude judgement (i.e., they had to say whether the number was larger or smaller than five). The difference between the desks was that in one deck (low demand) the colour of the number matched the previous one in 90% of the trials. In the other deck (high demand), this only matched in 10% of the trials. The participants were also free to choose when they wanted to switch between decks. Given that task-switching is known to require a lot of cognitive control, the researchers hypothesized that if cognitive control is indeed costly, participants would avoid the deck with a high frequency of task switches. As predicted, they found that the participants significantly preferred the low demand deck over the high demand deck. Based on findings like these, researchers have argued that cognitive control is not an unlimited resource. So, it's hard to maintain high levels of this control for a long time (Diamond, 2013; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004). Because of this cost and since cognitive control is not an unlimited resource, people tend to avoid cognitive control. They tend to minimize the effort they have to put in to do a task. Kool, McGuire, Rosen, and Botvinick (2010) used the same demand selection task as Botvinick and Rosen (2009). They found the same results, participants tend to choose the less effortful task, following the law of least effort. Hull (1943) clearly stated that when people can choose between two equal tasks, but one more effortful than the other, they will choose the less effortful one. This theory still stands in psychology, and even outside of psychology (Salamone, Correa, Farrar, & Mingote, 2007).

Building on these empirical findings, Shenhav, Botvinick, and Cohen (2013) introduced the normative model of the expected value of control (EVC). This theoretical framework integrates the previously described conflict-monitoring theory with the notion that cognitive control is costly. When allocating control to a task, this comes with a cost. This model allows us to represent the value of this allocation, which is supposed to be estimated by the ACC, which we know plays a role in cognitive control. Now, the ACC gathers information to estimate the EVC and the EVC uses task difficulty (state) and signal to estimate the probability of rewards. We can assume that conflict monitoring allows us to create an endogenous index of task difficulty. In sum, depending on the task difficulty and the potential reward, the ACC will arbitrate between investing in cognitive control or avoiding cognitive control.

The potential role of metacognition

However, one important aspect of human cognition is missing in the EVC-framework, namely metacognition. Metacognition is being aware of what you are thinking or feeling, it implies introspective access to your own cognitive processes (Desender et al., 2017). Another approach to induce conflict is to use priming. This experimental approach has the advantage that it allows to study metacognition. Priming is a method where a stimulus is preceded by a prime (e.g., word, picture, sentence, sound, etc.), which then causes facilitation or impediment of the task. For example, participants have to decide whether an arrow is pointing to the left or the right. This stimulus (e.g., arrow) is preceded by a prime (e.g., another arrow), which can point in the same or opposite direction as the target. When both arrows (i.e. prime and target) point in the same direction, the trial is congruent. When the arrows point in opposite directions, the trial is incongruent. Using this paradigm, Desender, Van Opstal, and Van den Bussche (2014) have found that participants are significantly faster when prime and target are congruent. Now, this prime can also be presented subliminally. This means it is presented for such a short time, too short for conscious perception (Loftus & Klinger, 1992). In the example above, the prime (i.e. first arrow) can be presented so short that it doesn't reach the threshold for consciousness, and therefore there is no conscious perception. Importantly, even without the participant being consciously aware of this prime, it nevertheless influences the response. Even though the prime is presented subliminally, participants are significantly slower when there is a conflict between the prime and the target (Desender, Calderon, Van Opstal, & Van den Bussche, 2017).

Interestingly, even though participants are unable to perceive the prime, Desender et al. (2014) have shown that they are nevertheless able to experience where a trial was difficult (i.e., incongruent) or easy (i.e., congruent). This was shown by asking participants after each trial whether they though that trial was rather easy or rather difficulty, this was used to study their metacognition. The fact that participants were able to introspect on the difficulty of each trial is remarkable because when the prime is not consciously perceived, congruent and incongruent trials are visually indistinguishable. Therefore, participants seemed to rely on non-visual information when making their judgment about the difficulty of that trial. Thus, it seemed that participants had good metacognitive awareness about the difficulty of responding to each trial. In the Desender et al. (2014) study, even though the participants reported not being aware of the primes, the incongruent trials were rated significantly more difficult than the congruent trials (Desender et al., 2017). Thus, even when participants are not aware of the primes, these still influence reaction times (Cohen et al., 1990), brain activity (Draine & Greenwald, 1998), and metacognitive experience associated with the performance. In other words, participants did not need to be aware of the primes to be able to measure their influence on performance.

As discussed before, there is evidence that the N2 component measured in the EEG (reflecting dACC activity) is modulated by task-related conflict. Interestingly, this activity appears unrelated to metacognitive appreciation of the difficulty of a trial (Desender et al., 2016). In other words, the N2 component seems to track the conflict on a given trial, but not the subjective evaluation of such conflict. Metacognitive difficulty did affect another component that takes place around 300-400 ms post-stimulus, the P3 component (Desender et al., 2016). The true role of this component is still uncertain, but there is some evidence that it plays a role in subjective perception (Del Cul, Baillet, & Dehaene, 2007), and it also seems to be modulated by conflict (Gratton et al., 1992). There is some more evidence that the P3 is modulated by difficulty in a masked priming paradigm (Desender et al., 2016) or with confidence in an auditory task (Zakrzewski, Wisniewski, Iyer, & Simpson, 2019). Desender et al. (2016) also showed that the size of the P3 was predictive for the congruency, but also the metacognitive experience (i.e. subjective difficulty).

From the above, it appears that there is evidence that conflict detection (i.e., dACC, N2) is related to investing cognitive control (Gratton et al., 1992) and avoiding cognitive control (Kool et al., 2010). However, there is also evidence that metacognition (i.e., P3) is related to investing cognitive control (Desender et al., 2014) and avoiding cognitive control (Desender et al., 2017). Given that both conflict detection and metacognition of conflict have been shown to be dissociable and have separable neural correlates, it remains an open question which of these two factors drives choices about cognitive control. According to the EVC model, participants will invest cognitive control in a task when the potential reward exceeds the costs of investing control. In the current study, we wanted to assess whether the cost of investing control in this equation is indexed by the detection of conflict on one hand, or by the metacognitive experience on the other hand.

To investigate this question, participants will first take part in a rating phase during which they are performing a masked priming task and additionally rate their subjective level of difficulty. During this task, we will measure EEG activity. This will allow us to look at which neural signals are associated with the detection of conflict, and which neural signals are associated with metacognitive experiences of subjective difficulty. Second, participants will take part in a choice phase. This will be the same masked priming task as before, but before each trial participants have to decide whether they want to perform a high-demand version of the task and have the potential to win a large reward (i.e., high effort – high gain), or perform a low-demand version of the task and have a potential to win a small reward (i.e., low effort – low gain). Also during this second part, we will measure EEG. Using the EEG data, we will then examine whether investment or avoidance of cognitive control during the choice phase is predicted by neural coding of response conflict (i.e., as predicted by the EVC theory) or by neural coding of metacognition (i.e., as predicted by metacognition research).

Materials and method

Participants

Twenty-six participants took part in this study (22 females, *M* age = 24.78 years, *SD* age = 3.86 years). They all participated in turn for monetary compensation, which was around 20 euros. They could also obtain an extra reward of two or ten euros based on their performance. All participants signed a written informed consent, reported normal or corrected vision and did not know the hypothesis of the study.

Stimuli and apparatus

The experiment was a masked priming task programmed in Python 2.7. The stimuli were presented on a grey background on an LCD screen with a 120Hz refresh rate. The stimuli consisted of white arrows that pointed to the left or the right. The prime and target were made up in such a way that the prime fits perfectly into the target, so the prime seemed invisible. They were also presented slightly above or below the fixation cross so the fixation cross could stay present during the trial. The rating scale was made using a gradient from green to red or red to green. This was presented in the centre of the screen. The ten- and two-euro boxes were pictures of a ten-euro bill and a two-euro coin. These were presented on four different positions (left, right, above or below the centre). The responses were made using a standard QWERTY keyboard and a mouse.

Procedure

Figure 1 shows an example of an incongruent trial in the rating phase. In this phase, a trial started with 500 ms of gaze-contingent fixation. Here, the participant had to continuously fixate on the fixation cross for 500 ms to continue, if they would break fixation, the 500 ms period would restart. This was followed by another gaze-contingent fixation of 500 ms where the trial would be aborted if they would break fixation. Next, the prime was presented for 23 ms, this was presented slightly above the fixation cross so the fixation cross could be presented during the whole trial. The prime could be an arrow pointing to the left or the right. Between the prime and the target, a blank screen with the fixation cross was presented for 23 ms. Next, the target was presented for 120 ms at the same position as the prime. The target could also point to the left or the right (i.e., congruent or incongruent with the prime). After the target, a blank screen with the fixation cross was presented for 1.5 seconds or until the participant responded. They had to respond by pressing the “s” or the “d” keys with their left middle finger and left index finger, respectively (“s” for the arrow pointing to the left and “d” for the arrow pointing to the right). Then, there was again a blank screen with the fixation cross for 750 ms. And lastly, the rating scale was presented until the participant responded. The rating scale was a gradient rectangle from green to red or red to green. The participants were told to rate the trial on their subjective difficulty using the mouse. The green part of the scale was associated with subjectively easy and the red part of the scale with subjectively difficult. On each trial, the orientation of the scale (red-to-green, or green-to-red) was randomly chosen. The entire trial was fixation contingent, meaning the participant could not look away from the fixation cross when it was presented, until the moment the rating scale appeared.

The choice phase was similar to the rating phase, except it entailed an additional choice aspect. Figure 2 shows the trial structure of an incongruent trial in the choice phase. Here, each trial started with the presentation of two boxes reflecting a € 10 bill or a € 2 coin which represented the high effort – high gain and the low effort – low gain conditions, respectively. This screen was presented until the participants chose a box, which they did by clicking with the mouse on one of the boxes. On each trial, the two boxes appeared randomly in two out of four conditions equally spaced around fixation. In the low effort – low gain condition, participants could potentially obtain € 2 if their subsequent response in responding to the target arrow was correct and faster than 1.5 seconds (hence, this condition was low effort - low gain). In the high effort – high gain condition participants could potentially obtain € 10 if their subsequent response in responding to the target arrow was correct and faster than a designed threshold. This threshold was determined on an individual level, and it was quantified as percentile 10 of the reaction times in the rating phase. Thus, participants could potentially obtain € 10 only if their subsequent response was correct and faster than 90% of responses in the rating phase (hence, this condition was high effort - high gain). In both conditions, when the participant made the wrong choice about the direction of the target arrow or they responded too slow, the participant could potentially win € 0. Participants were told that after completing the experiment, all of the € 0, € 2 and € 10 that they collected would be put together in one box, and the computer would randomly choose one of these values. Participants would then receive an extra monetary compensation equal to the value of the chosen number. Importantly, participants did not receive trial-by-trial feedback about the reward on a given trial. Thus, there was no direct way for participants to use the obtained reward to guide their effort choices. In order to provide participants with a sense of how well they were doing, every 10 trials we provided them with an overview of their performance in the last 10 trials. Specifically, the participants were presented with two boxes (i.e., € 2 rewards and € 10 rewards) where the number of times they were correct and fast enough in that condition was shown. After participants clicked on the € 2 or the € 10 box, the trial followed the same structure as in the rating phase, but without the presentation of the rating scale. Again, the whole trial was fixation contingent when the fixation cross was presented.

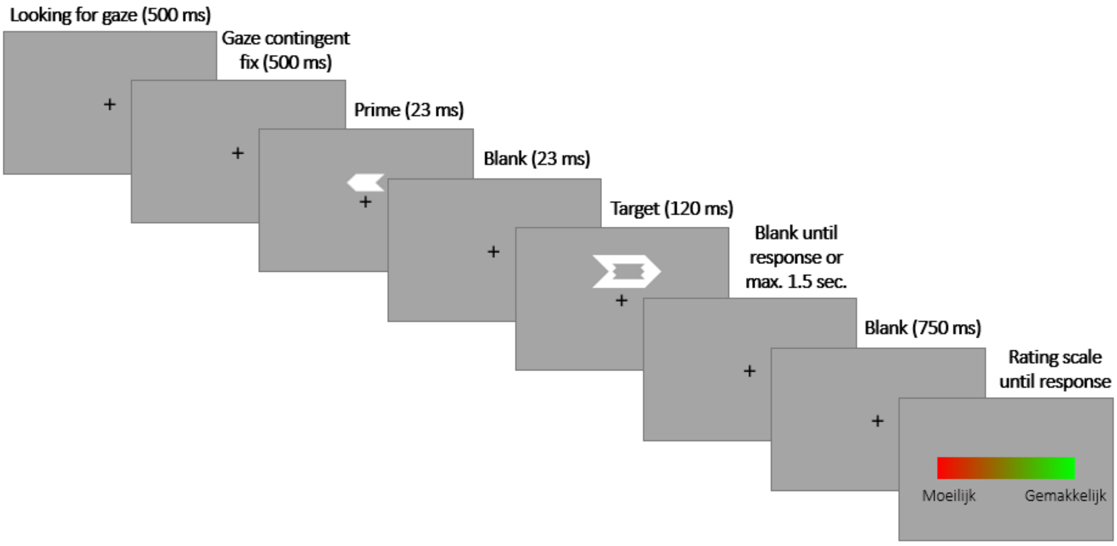


Figure 1. Timeline of an incongruent trial in the rating phase. A prime was presented for 23 ms and a target was presented for 120 ms, the participants had to react as fast as they could to the target by using the keyboard, reporting whether the target pointed to the left or the right. After reporting the target direction, the participants had to rate their subjective difficulty on the rating scale by using the mouse (stimuli larger for visualisation purposes).

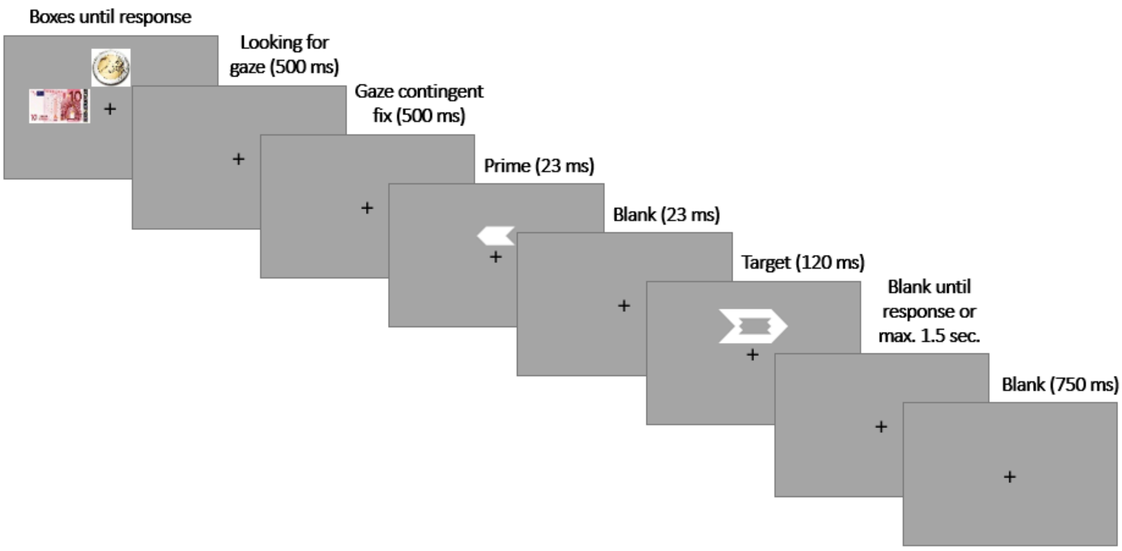


Figure 2. Timeline of an incongruent trial in the choice phase. Two boxes were presented that represented a low effort – low gain (i.e., no reaction time threshold) or a high effort – high gain (i.e., fast reaction time threshold) condition. The participants had to choose a

condition by clicking one of the two boxes. Next, a prime was presented for 23 ms and a target was presented for 120 ms, the participants had to react as fast as they could to the target by using the keyboard, reporting whether the target pointed to the left or the right (stimuli larger for visualisation purposes).

The main experiment consisted of six blocks of 60 trials for the rating phase and six blocks of 60 trials for the choice phase. Before each phase, there was also a practice block of 16 trials.

EEG recording and preprocessing

EEG was recorded using a fabric cap with 64 channels, referenced to channel Cz, while participants sat in a dimly lit room. Vertical and horizontal electrooculogram was measured from above and below the right eye. The stimulus-locked data was baselined at -250 to -50 ms before stimulus onset. To look at ERPs depending on the time of response, these data were then realigned to the onset of the response to the target (i.e., “s” or “d” keypress), while keeping the same pre-stimulus baseline.

For the preprocessing, trials without an associated reaction time (i.e., no reaction time data), trials that were slower or faster than three standard deviations from the individual mean, trials without an associated rating (i.e., no subjective difficulty rating) and trials where participants blinked or looked away were excluded. On the remaining data, a notch filter of 50 Hz was applied to remove line noise. Next, a low-pass filter of 30 Hz was applied. The epochs were then made using the trigger of the fixation cross. Then, the epochs were visually inspected and trials with large artefacts were manually removed. Bad or noisy channels were interpolated. The data was also down-sampled to 250 Hz. Lastly, for the independent component analysis (ICA), a high-pass filter of 1 Hz was applied to do the ICA decomposition on. Then the weights of the ICA were applied to the non-high-pass filtered data. This was used to remove eye-blinks from the data.

Statistical analysis

Behavioural

For the rating phase, linear mixed models were fit to the reaction times on correct trials and the z-scored difficulty ratings on correct trials to examine whether these variables were affected by congruency. We also fit generalised mixed models to accuracy to examine whether it was affected by congruency. All models included a random intercept, whereas random slopes were only added when this significantly improved the model fit. The model fit was assessed using model comparison. The degrees of freedom were estimated using the Satterthwaite approximation. Another linear mixed model analysis was used to examine the influence of congruency and reaction time on difficulty ratings. Again, only the correct trials were used. For the choice phase, a generalised mixed model analysis was fit on reaction times to examine whether they were affected by congruency, current context (i.e., two euro or ten euro) and their interaction. The same generalised mixed model was fit for accuracy and reward.

Event-related potentials (ERP)

To compute event-related potentials, we averaged the stimulus-locked EEG data and the response-locked EEG data separately for congruency (congruent vs. incongruent) and subjective difficulty (easy vs. difficult) in the rating phase, and separately for congruency (congruent vs. incongruent), and subsequent box choice (low effort – low gain vs. high effort – high gain) in the choice phase. To assess significance, we focused on average amplitude on electrode FCz and computed cluster-based t-tests for congruency and difficulty ratings in the rating phase, and for congruency and choice in the choice phase.

Decoding analyses: temporal generalisation

Then, a decoder was trained on congruency and rating in the rating phase, so we could confirm we could decode the information. This was also done for following choice in the choice phase. We started by doing a low-pass filter of 20 Hz on the already pre-processed data. The data was then epoched between .1 and 1 second. For the trial selection, the data was divided into congruent and incongruent trials. This was done in a way that for each congruent trial, there was a matched incongruent trial that differed by 5 points on subjective difficulty maximum. This was done to make sure the predictor predicted on congruency and not subjective difficulty. For congruency, we used the LDA classifier using a least-square solver with automatic shrinkage using the Ledoit-Wolf lemma. For ratings, we used a scorer made by Jean-Remi King. A scorer is a function that yields a performance score based on the relation between the predictions by the classifier and true labels. In this case, we used Spearman correlation to compare rating predictions from the Ridge regression and true ratings. For future choice, we took the number of trials for the class with the lowest number and used the same number for the other class. The training data was resampled 10 times and 10-fold cross-validation was used.

Results

Behavioural results

First, we focused on the behavioural results of the rating phase. A linear mixed model showed that the reaction times were faster on congruent trials (M = 375 ms) than incongruent trials (M = 443 ms), $F_{(1,25.96)} = 331.92$, $p < .001$. A generalized mixed model showed that accuracy was higher on congruent trials (M = 99.4%) compared to incongruent trials (M = 90.0%), $X_{(1)} = 86.09$, $p < .001$. Finally, a linear mixed model on normalized difficulty ratings showed participants rated incongruent trials as more difficult (M = .106) than congruent trials (M = -.34), $F_{(1,26.06)} = 42.00$, $p < .001$. At the individual level, the congruency effect was significant in reaction times for 27 out of 27 participants, in accuracy for 24 out of 27 participants, and in difficulty rating for 19 out of 27 participants (figure 3).

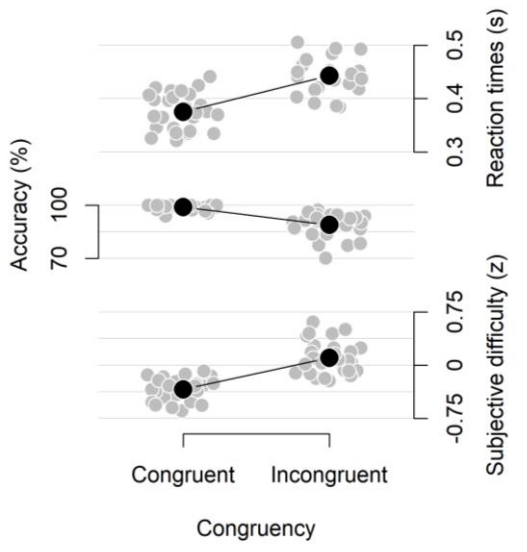


__Figure 3. Behavioural results rating phase.__ Reaction times, accuracy and subjective difficulty are all modulated by congruency. Each grey dot represents a participant and the black dot represents the mean.

Next, we examined whether difficulty ratings depended not just on the congruency itself, but also on the reaction time of a given trial. To examine this, we carried out a linear mixed model on normalized ratings with both congruency, reaction time and their interaction as predictors. This analysis showed significant main effects of congruency and reaction times ($p < .001$), demonstrating that both these variables have a unique influence on difficulty ratings. Moreover, there was also a significant interaction effect ($F_{(1,7323.9)} = 23.82$, $p < .001$), showing that the effect of congruency was mostly expressed at slow reaction times (figure 4).

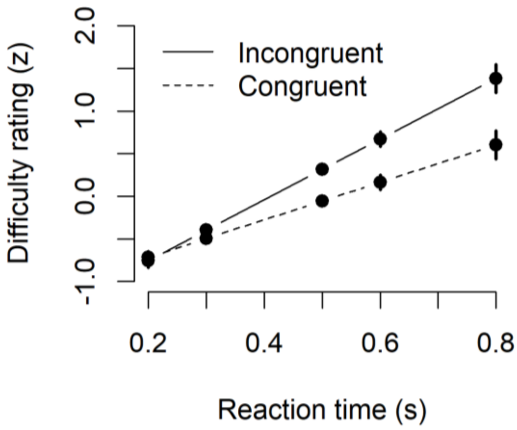


***Figure 4. Behavioural results rating phase**. The influence of congruency on subjective difficulty ratings was very large for slow reaction times, whereas it was much reduced for very fast reaction times.*

Next, we examined the data of the choice phase, and first examined whether we observed the same effects of congruency. A mixed regression model was fitted on reaction times with congruency, current context (low effort – low gain or high effort – high gain) and their interaction as predictors. This analysis showed a main effect of congruency on reaction times ($F_{(1,25.4)}$ = 240.34, p < .001), an effect of current context on reaction times ($F_{(6712.5)}$ = 998.68, p < .001), and a significant interaction between the two ($F_{(1,6173.8)}$ = 32.63, p < .001). In figure 5, you can see a large congruency effect in the low effort – low gain context (72 ms, z = 17.75, p < .001) and a slightly smaller, but still highly significant congruency effect in the high effort – high gain context (52 ms, z = 17.75, p < .001). The same analysis on the accuracy data gave highly similar results, namely a main effect of congruency ($X^2_{(1)}$ = 139.18, p < .001), current context ($X^2_{(1)}$ = 69.09, p < .001), and interaction between the two ($X^2_{(1)}$ = 3.96, p = .047). Similar to the reaction times, there was a very pronounced congruency effect in the high effort – high gain context (36.5%, z = 10.03, p < .001) and a slightly smaller but still significant congruency effect in the low effort – low gain context (14.8%, z = 9.46, p < .001).

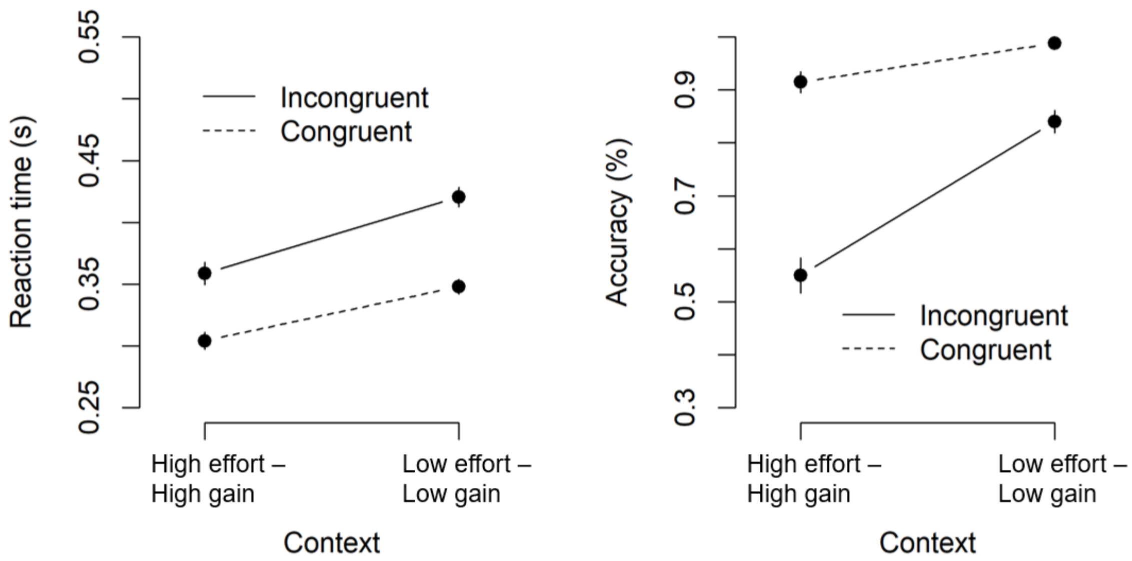


Figure 5. Behavioural results choice phase*. The congruency effect on reaction times and accuracy depending on the context.*

Next, to examine how these variables affected the probability of actually obtaining a reward (i.e., in the high effort – high gain context: being faster than 90% of reaction times of the rating phase and being correct; in the low effort – low gain context: being faster than 1.5 seconds and being correct). Therefore, we ran another generalized mixed model predicting whether or not a trial was rewarded, by congruency and current context. This analysis showed main effects of congruency ($X^2_{(1)}$ = 102.26, p < .001) and context ($X^2_{(1)}$ = 211.38, p < .001) on reward rate (i.e., accurate and fast enough), but there was no interaction between the two (p = .112). Figure 6 shows participants were more likely to obtain a reward in congruent trials (M = .73) compared to incongruent trials (M = .46). They were also more likely to obtain a reward in the low effort – low gain context (M = 91%) compared to the high effort – high gain context (M = 27%).

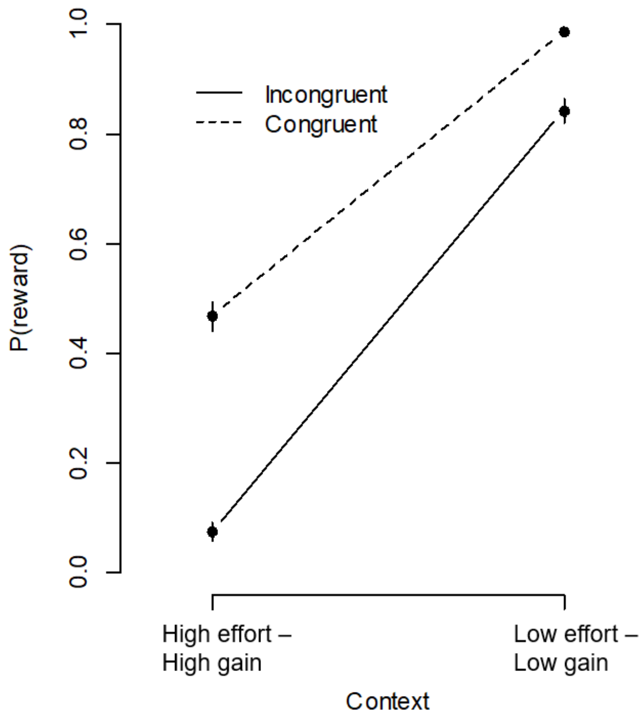


*Figure 6. **The probability of obtaining a reward in the choice phase.** The congruency effect and effects of context on reward rate.*

Next, we turned towards the context choices that participants made for the following trial. First, we fitted a generalised mixed model predicting which context participants chose to perform the next trial in, by current context, congruency, and their interaction. This analysis was performed on correct trials only. This analysis only showed a main effect of current context ($X^2_{(1)}$ = 13.082, p < .001), but no main effect of congruency (p = .989), nor interaction (p = .292). Thus, it appears that future context choices were not affected by whether or not there was a response conflict on the current trial.

Given the finding in the rating phase that the congruency effect was mostly pronounced at slow reaction times, we, therefore, considered the possibility that congruency only affected future context choices for slow reaction times. Therefore, we added reaction times (and all interactions with reaction times) to the previous model. Apart from the main effect of current context (p < .001), there now also was a main effect of reaction times ($X^2_{(1)}$ = 13.50, p < .001), and a few borderline significant effects (congruency: p = .056; congruency by reaction times: p = .080; choice by congruency by reaction times: p = .084). As can be seen in figure 7, the future choice did not depend on congruency for very fast reaction times, whereas for slower reaction times, participants tend to increasingly select the safe low effort – low gain condition when the trial they just performed was incongruent. This finding is consistent with the finding in the rating phase that the congruency effect was mostly pronounced at slow reaction times. Moreover, it is consistent with our hypothesis that people monitor congruency and rely on this source of evidence to guide effort choices. Note, however, that this conclusion should be taken with serious caution given that the three-way interaction did not reach significance (p = .080).

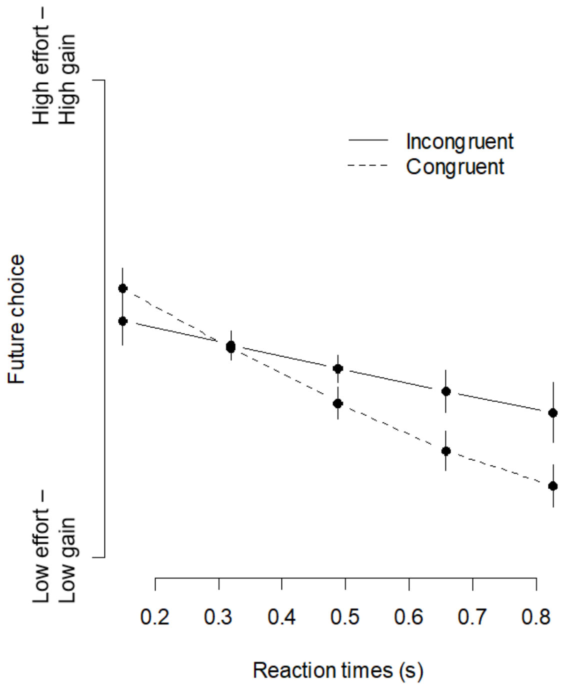


Figure 7. The congruency effect and effect of reaction time on future choice. *Participants increasingly selected the high effort – high gain context for future trials when they made a fast response to the current trial. It also appears that for slow responses, participants were more likely to select the high effort – high gain context after an incongruent trial, whereas this pattern was reversed for fast responses. Note, however, that this effect should be interpreted with caution is it did not reach statistical significance.*

In the final analyses, we directly examined the influence of obtaining a reward or not on the following choice (i.e., as an aggregate measure of accuracy and reaction times). Again, note that participants did not receive any trial-by-trial feedback so any effect of obtained reward is purely because participants internally monitored reward probability. To do so, we fitted a general mixed model featuring only current context, reward and their interaction as predictors. This analysis showed significant main effects of current context ($p < .001$), reward ($X^2_{(1)} = 12.46$, $p < .001$), as well as a significant interaction ($X^2_{(1)} = 6.56$, $p = .010$). This interaction reflected that people were more likely to select the high effort – high gain condition after a rewarded versus an unrewarded trial, an effect that was significant in the low effort – low gain context (12.9%, $z = 4.16$, $p < .001$), but not in the high effort – high gain context (1.8%, $z = 0.54$, $p = .586$) (figure 8).

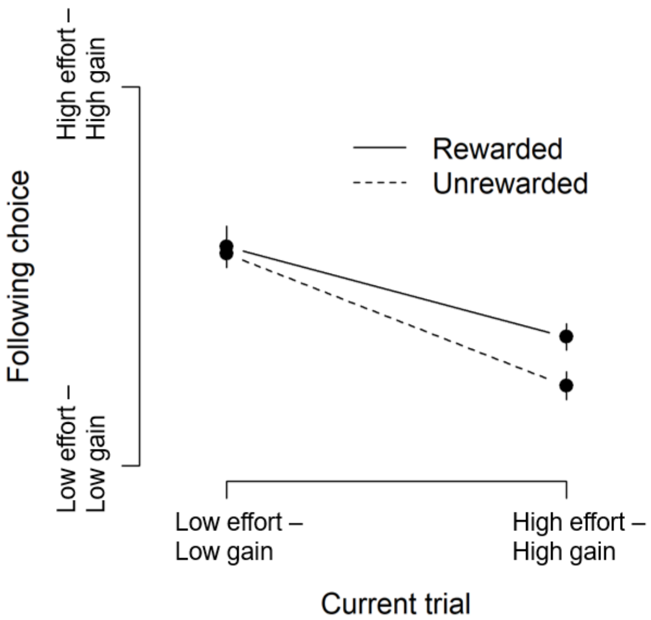


Figure 8. The effect of reward and current trial on following choice.

To summarise the above, the results from the choice phase demonstrate that participants do monitor their performance and reward outcome, and use this information to guide effort choices (although this was only shown for reaction times, accuracy and reward, not for congruency). Now, the critical question is whether we can also predict effort choices using the EEG data of the rating phase.

Event-related potentials

To examine how neural responses were modulated by congruency and subjective difficulty, we used the data of the rating phase and computed event-related brain potentials on correct trials only, and we did so separately for congruent and incongruent trials. These were then further split into subjectively easy trials and subjectively difficult trials using the median of each participant. As can be seen in figure 9 A, at electrode FCz, the stimulus-locked ERP showed significant modulation by congruency from 160 ms to 224 ms (p = .011), from 256 ms to 300 ms (p = .014) and from 320 ms to 416 ms (p = .003). The second significant time-window from 256 ms to 300 ms shows a frontocentral negativity on the topographical plot, which corresponds to the N2 component. The later significant time-window from 320 ms to 416 ms show a more frontal positivity on the topographical plot, which corresponds to the P3 component. Both these findings are well in line with the literature on conflict processing, where congruency usually affects the N2 and the P3 component. Unexpectedly, we did not observe a significant modulation of subjective difficulty. However, based on the previous behavioural analyses and the study by Desender et al. (2016), we decided to do a paired t-test on incongruent-difficult trials versus incongruent-easy trials, using the time-window between 360 ms and 460 ms post-stimulus (by Desender et al. (2016)). This showed a significant modulation (p < .001). So, even though we did not find any effects of subjective difficulty when using a cluster-based permutation test, when using a more specific time window chosen from a previous study on this topic, we did observe that subjective difficulty specifically modulated the P3 component. We also looked for this modulation in the response-locked ERPs. Figure 9 B shows the response-locked ERPs at electrode FCz for the rating phase. Here, the ERP showed nearly significant modulation by congruency during the time-window from -132 ms to 008 ms (p = .053). We also see a frontal positivity for congruency on the topographical plot for this cluster. For rating, there were no significant clusters.

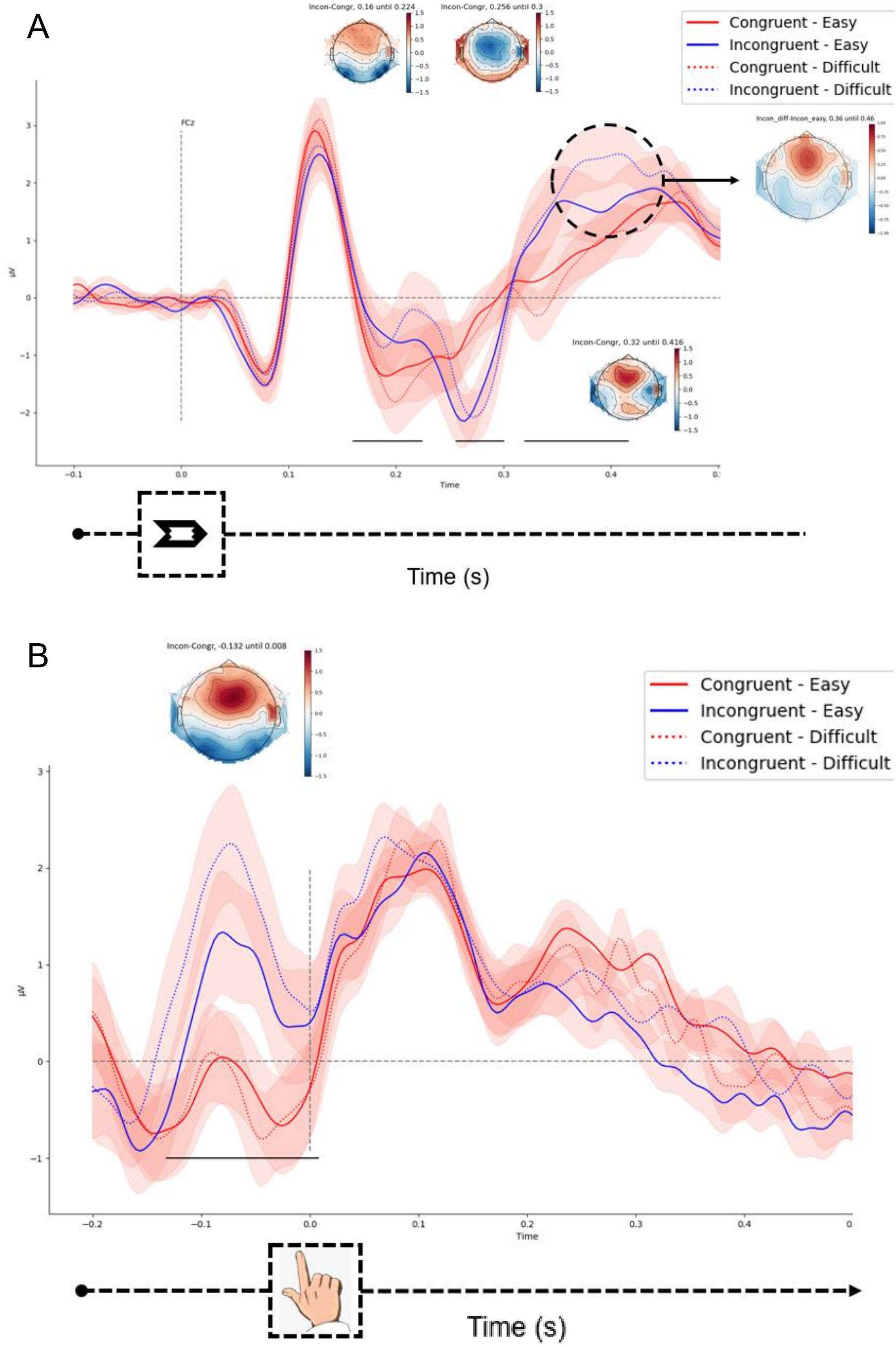

A
Incon-Congr, 0.16 until 0.224
Incon-Congr, 0.256 until 0.3
Congruent - Easy
Incongruent - Easy
Congruent - Difficult
Incongruent - Difficult
FCz
Incon_diff-Incon_easy, 0.36 until 0.46
Incon-Congr, 0.32 until 0.416
µV
Time
Time (s)
B
Incon-Congr, -0.132 until 0.008
Congruent - Easy
Incongruent - Easy
Congruent - Difficult
Incongruent - Difficult
µV
Time
Time (s)

Figure 9. Stimulus-locked ERPs at electrode FCz as a function of congruency and subjective difficulty for the rating phase (A). Response-locked ERPs at electrode FCz as a function of congruency and future choice for the choice phase (B).

For the choice phase, we again calculated the ERPs for the correct trials only. These were then split into congruent and incongruent trials and then also into the following choice (i.e., low effort – low gain or high effort – high gain). On figure 10 A, at electrode FCz, the stimulus-locked ERP shows a significant modulation by congruency during the time-window from 244 ms to 308 ms (p = .012) and during the time-window from 352 ms to 476 ms (p < .001). Both these effects correspond well to what was observed in the rating phase (i.e., figure 9), namely a modulation of the N2 and the P3 component by response conflict. Finally, there was no significant modulation by following choice. Figure 10 B shows the response-locked ERPs at electrode FCz for the choice phase. The ERP shows a significant modulation by congruency during the time-window from -128 ms to 4 ms (p < .001) and during the time-window from 80 ms to 232 ms (p = .002). The first significant time-window is accompanied by a large frontal positivity and the second significant time-window is accompanied by a smaller frontal positivity, which can both be seen on the topographical plots in figure 10 B. For following, choice, the ERP shows a significant modulation by following choice during the time-window from -200 ms to -140 ms (p = .047) and during the time-window from 80 ms to 568 ms (p < .001). The first significant time-window is accompanied by a posterior negativity, while the second significant time-window is accompanied by a central positivity, which can both be seen on the topographical plots in figure 10 B.

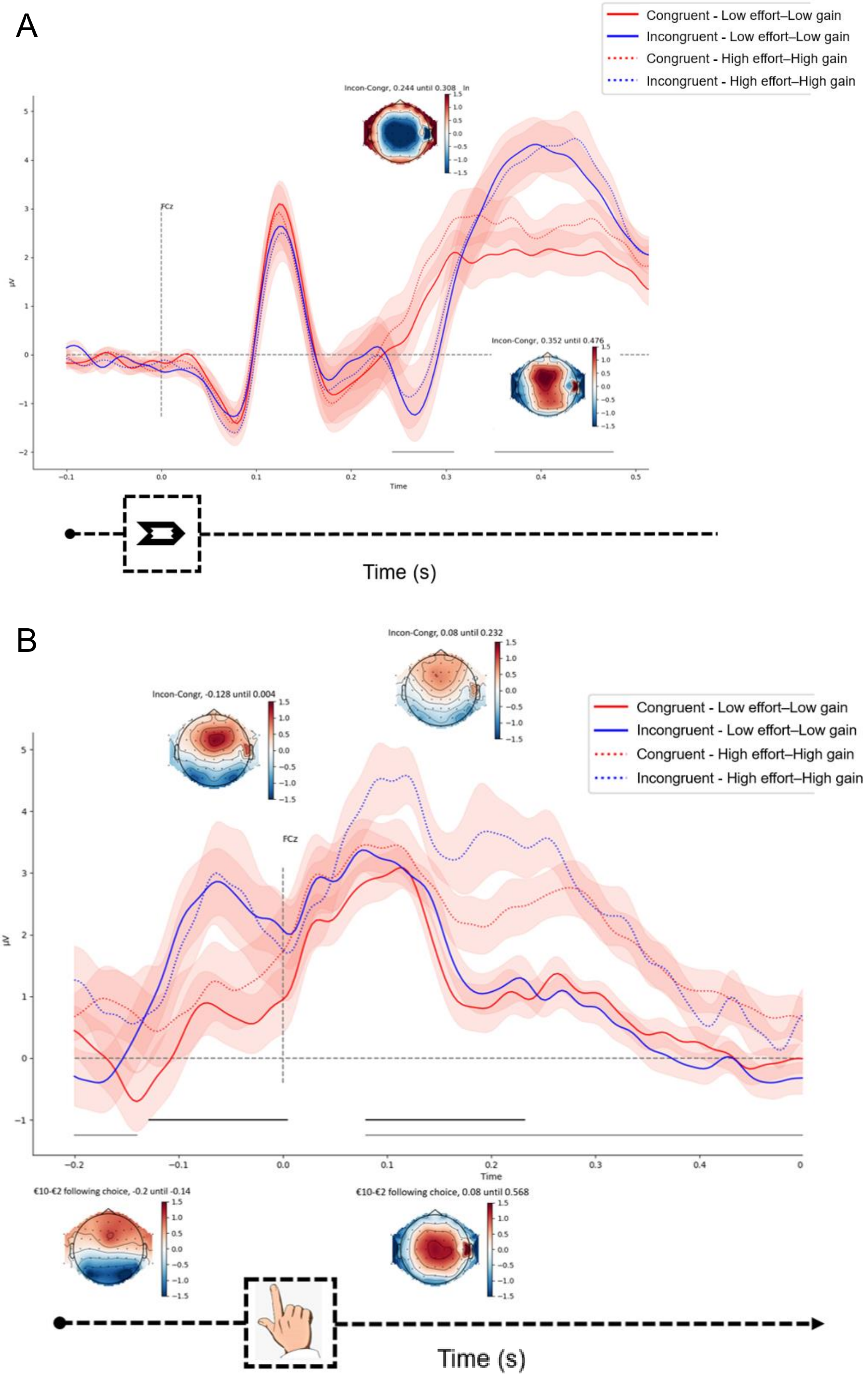

A
Congruent - Low effort–Low gain
Incongruent - Low effort–Low gain
Congruent - High effort–High gain
Incongruent - High effort–High gain
Incon-Congr, 0.244 until 0.308
Incon-Congr, 0.352 until 0.476
FCz
µV
Time
Time (s)
B
Incon-Congr, 0.08 until 0.232
Incon-Congr, -0.128 until 0.004
Congruent - Low effort–Low gain
Incongruent - Low effort–Low gain
Congruent - High effort–High gain
Incongruent - High effort–High gain
FCz
µV
Time
€10-€2 following choice, -0.2 until -0.14
€10-€2 following choice, 0.08 until 0.568
Time (s)

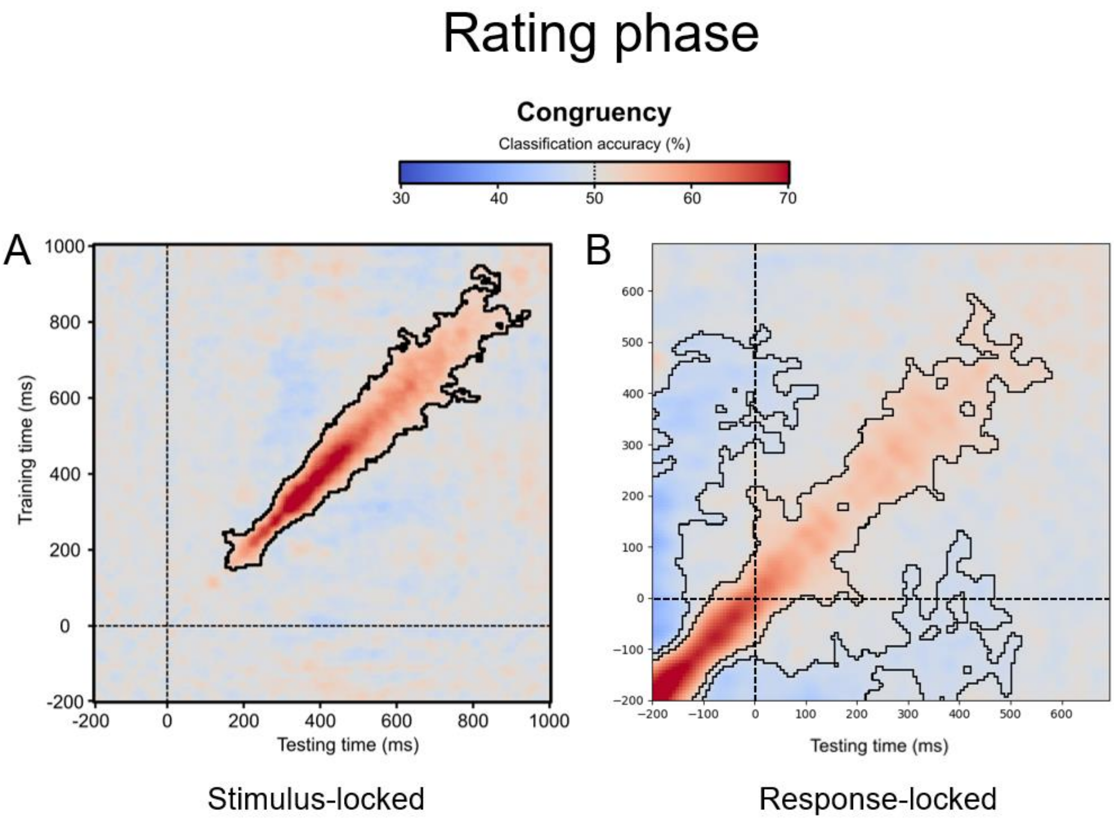
Figure 10. Stimulus-locked ERPs at electrode FCz as a function of congruency and future choice for the choice phase (A). Response-locked ERPs at electrode FCz as a function of congruency and future choice for the choice phase (B).

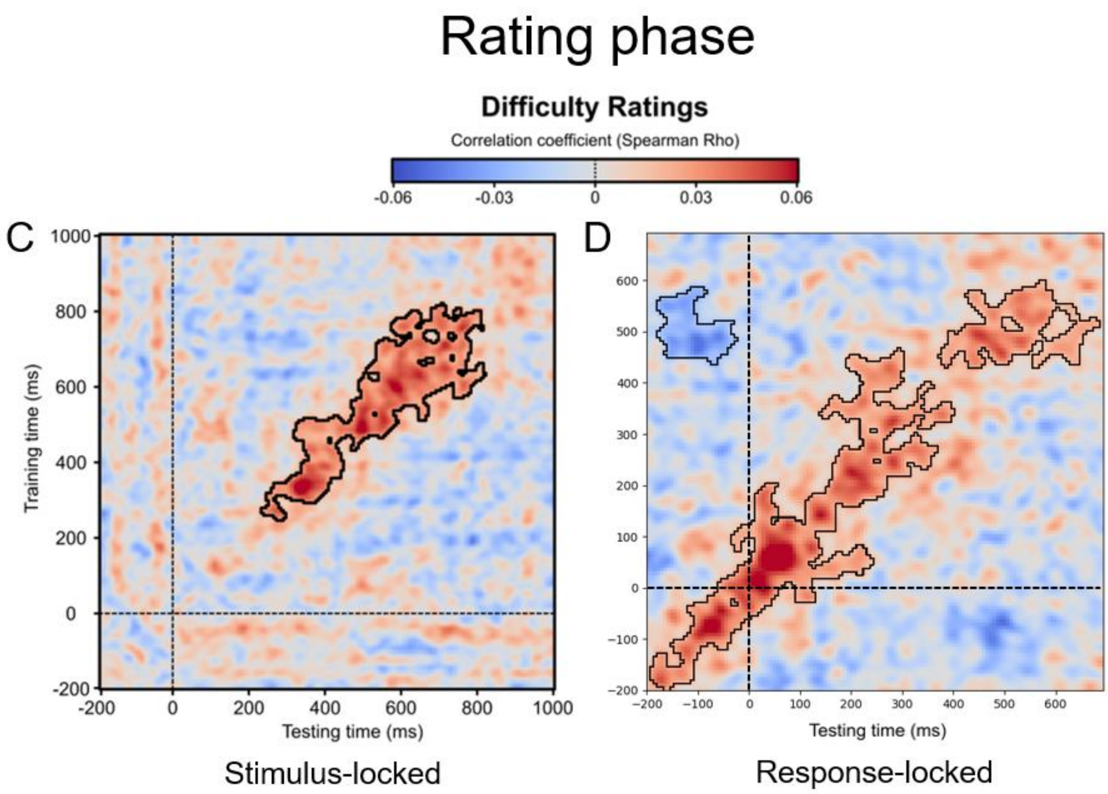
To summarise, the analyses on the ERPs shows us that there is a link between objective difficulty (i.e., congruency) and subjective difficulty (i.e., rating), given that they have similar neural markers.

Within-condition decoding

For the decoding analyses, we first wanted to see whether we could use a decoder to predict congruency, rating and following choice. For this part, only correct trials were used, to avoid that the decoder was instead decoding correctness versus errors (i.e. because congruent trials are more likely to be correct). Importantly, as explained before the trial selection for these decoding analyses was done orthogonal to the other variables (e.g., when decoding subjective difficulty it was reassured that each level of subjective difficulty had an equal level of congruent and incongruent trials). Therefore, if it is possible to decode these different variables this implies that these are associated with unique neural activity. For congruency, super-trials were made by averaging four trials together. It has been shown that averaging four trials is a good compromise to improve the signal to noise ratio (Grootswagers, Wardle, & Carlson, 2017). Figure 11 A and B show the temporal classification accuracy for congruency in the rating phase. This classifier was trained and tested on each point in time. As you can see, the classification accuracy goes as high as 70%. For the stimulus-locked data, you can see that significant decoding was possible from 200 ms to 900 ms (p < .0002) (figure 11 A). For the response-locked data, you can see the same significant clusters from -200 ms to 500 ms (p < .001) (figure 11 B). We can also see some significant clusters that have reversed polarity. This only appears pre-stimulus. Next, in figure 11 C and D you can see the decoded performance for subjective difficulty regression in the rating phase. The scale on this image shows how much the predicted subjective difficulty was correlated with the empirical subjective difficulty across the data and time. For stimulus-locked data, significant clusters were found from 200 ms to 800 ms (p < .0002) (figure 11 C). For response-locked data, significant clusters were found from -200 ms to 400 ms and from 400 ms to 700 ms ($p <$ .002) (figure 11 D). We again see significant clusters that have reversed polarity. Figure 11 E and F shows the decoded performance for future choice classification in the choice phase. This was to double-check whether we could decode this information, which was indeed possible. To overcome the imbalance between the low effort – low gain and high effort – high gain condition, we equated the number of trials for each condition. For the stimulus-locked data, multiple significant clusters were found from 250 ms to 800 ms (p < .006) (figure 11 E). For the response-locked data, we see the same significant clusters from -200 ms to 400 ms (p < .001) (figure 11 F). Some clusters have also reversed polarity. When we look at figure 11, we can see that the significant decoding always starts around 300 ms (stimulus-locked).

These analyses demonstrate that it is possible to decode each of these conditions based on EEG activity. Interestingly, the congruency and subjective difficulty decoding exhibited a significant decoding performance mostly on the diagonal, suggesting that these two cognitive processes were associated with distinct patterns of neural activity across time (King & Dehaene, 2014). Alternatively, if these conditions were decodable from a unique EEG pattern, the significant decoding performances would be spread in the horizontal and vertical axes (i.e., off-diagonal). This suggests that distinct informative neural activity about congruency and subjective difficulty unravels through time and that the neural representations of these two cognitive processes involve a succession of neural stages.


Rating phase
Congruency
Classification accuracy (%)
30
40
50
60
70
A
B
Training time (ms)
Testing time (ms)
Testing time (ms)
Stimulus-locked
Response-locked


Rating phase
Difficulty Ratings
Correlation coefficient (Spearman Rho)
-0.06
-0.03
0
0.03
0.06
C
D
Training time (ms)
Testing time (ms)
Testing time (ms)
Stimulus-locked
Response-locked

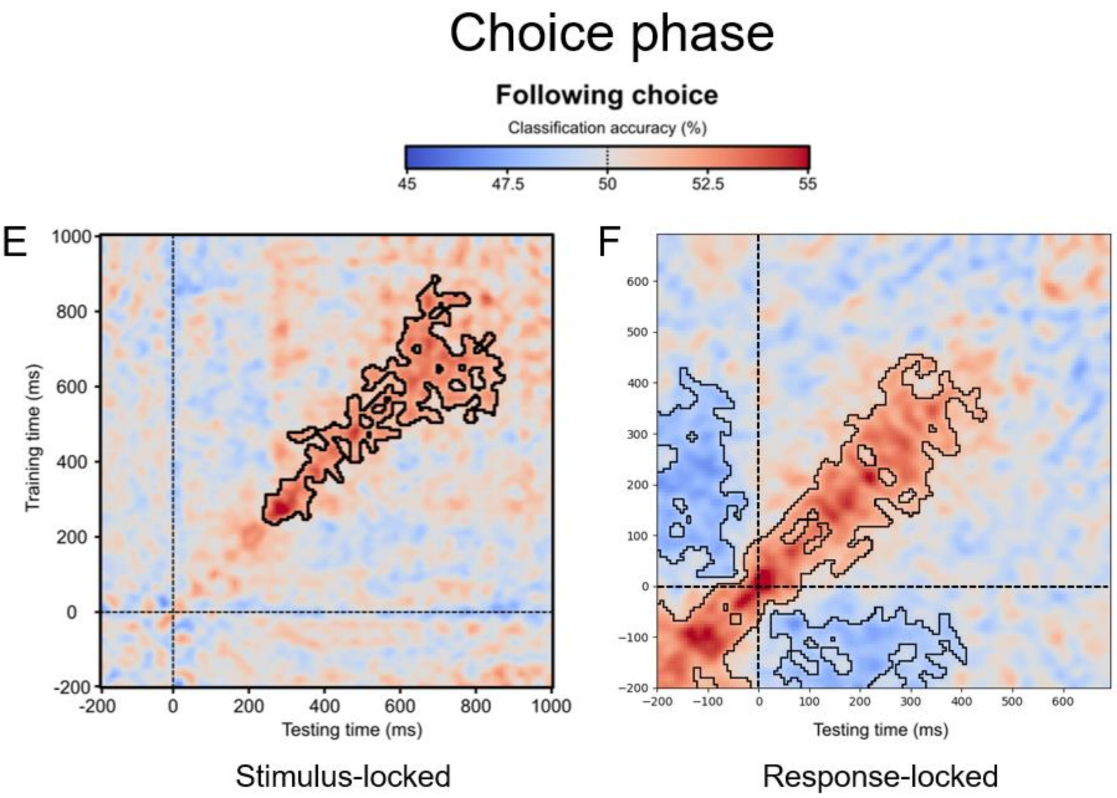


Figure 11. Decoding performances*. Decoding performances for congruency classification with stimulus-locked (A) and response-locked data (B), subjective difficulty regression with stimulus-locked (C) and response-locked data (D), future choice classification with stimulus-locked (E) and response-locked data (F).*

Across-condition decoding

Having shown that it is possible to decode congruency, subjective difficulty and future choices, we now turn towards the most critical research question. Namely, can we decode future context choices based on neural signatures of conflict or based on neural signatures of subjective difficulty? First, we examined whether a classifier that was trained to decode on congruency in the rating phase was able to predict future choices in the choice phase. When this analysis was done in the low effort – low gain context (figure 12 A and B), no significant clusters were found. When we repeated the same analysis on the data of high effort – high gain context (figure 12 C and D), also no significant clusters were found.

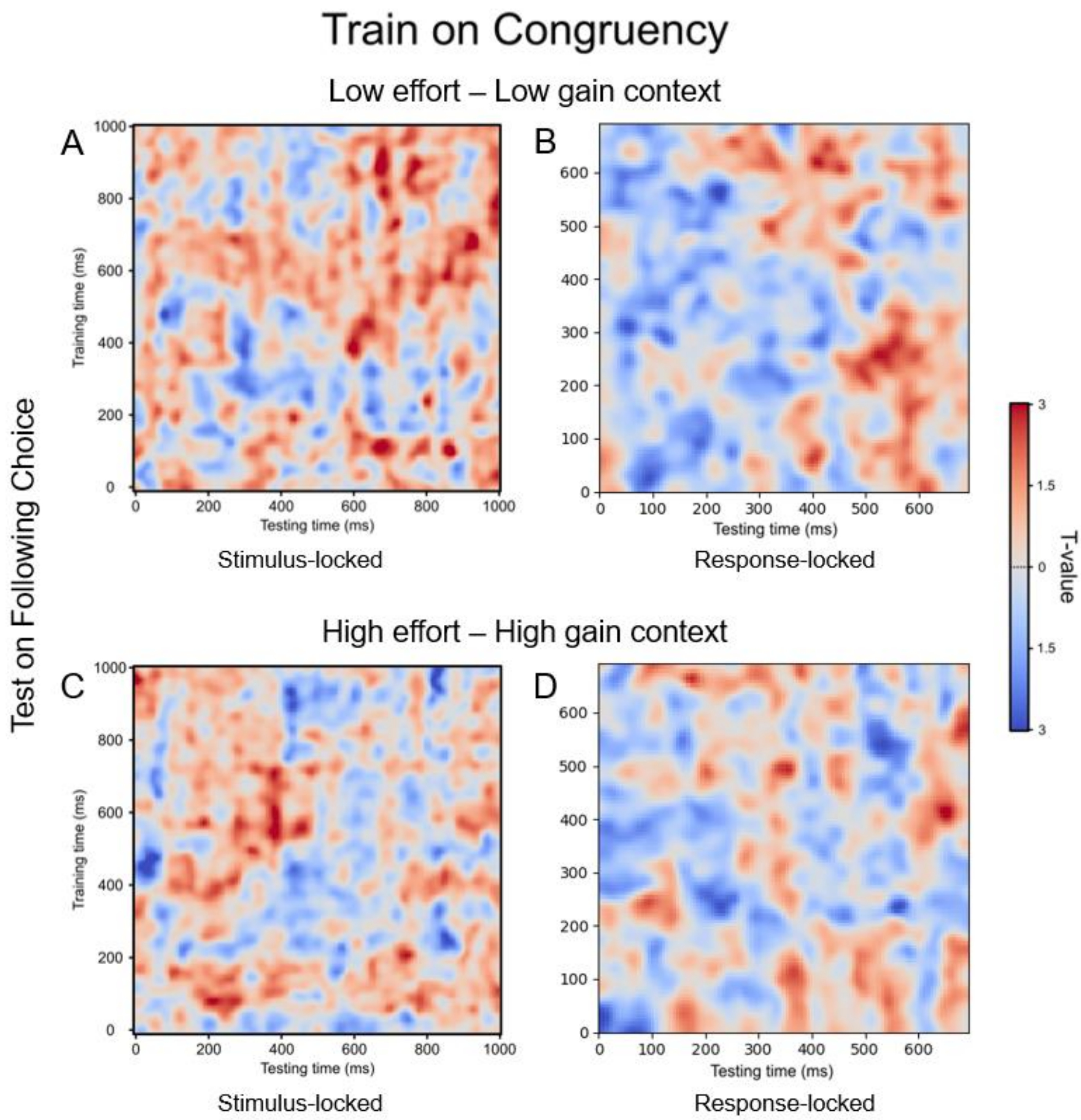


Figure 12. Cross-decoding accuracies for congruency in the rating phase. Training on congruency in the rating phase and testing on following choice in low effort – low gain context in the choice phase for stimulus-locked data (A), low effort – low gain context in the choice phase for response-locked data (B), high effort – high gain context in the choice phase for stimulus-locked data (C), and high effort – high gain context in the choice phase for response-locked data (D).

We then tried to decode by training on subjective difficulty rating in the rating phase and testing on future choice in the choice phase in the low effort – low gain context (figure 13 A and B). Again, no significant clusters were found. And lastly, we tried to decode by training on subjective difficulty rating in the rating phase and testing on future choice in the choice phase in the high effort – high gain context (figure 13 C and D). No significant clusters were found.

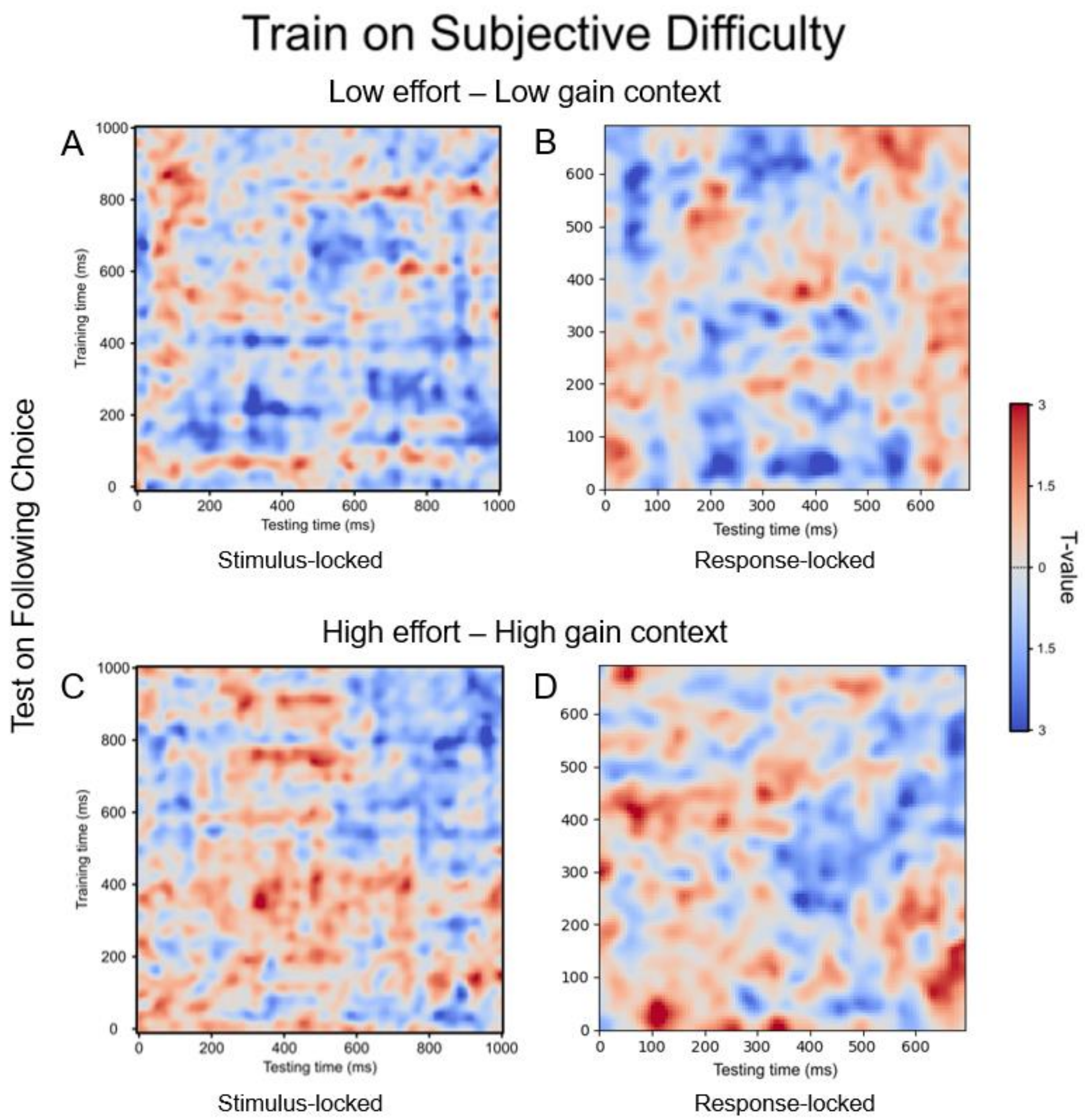


Figure 13. Cross-decoding accuracies for subjective difficulty in the rating phase. Training on congruency in the rating phase and testing on following choice in low effort – low gain context in the choice phase for stimulus-locked data (A), low effort – low gain context in the choice phase for response-locked data (B), high effort – high gain context in the choice phase for stimulus-locked data (C), and high effort – high gain context in the choice phase for response-locked data (D).

These results suggest that we cannot decode following choice based on the neural activity of congruency (i.e., objective difficulty) and rating (i.e., subjective difficulty).

Discussion

In this study, we examined whether neural markers of conflict detection or neural markers of subjective difficulty judgements could predict the investment of cognitive control in a demanding task. Which neural activity predicts our effort choices? Is it neural activity related to the feeling of subjectivity difficulty, or is it neural activity related to the detection of a conflict between two incompatible responses? To examine this, we recorded scalp EEG while participants performed a masked priming task where they had to decide whether the stimulus pointed to the left or to the right, followed by deciding the subjective difficulty for that trial (i.e., rating phase). In a subsequent choice phase, the participants had to choose whether they wanted to carry out this task in a low effort – low gain or high effort – high gain context.

The behavioural analysis for the rating phase showed that people were faster, more accurate and rated subjective difficulty lower for congruent trials compared to incongruent trials. The behavioural analysis for the choice phase showed that the participants being rewarded or not and the current context had an influence on their choice of the next condition. Here, when participants were in the high effort – high gain context (i.e., current trial) they were more likely to choose the low effort – low gain context (i.e., following trial) when they did not receive a reward. Importantly, this was observed even though participants did not receive trial-by-trial feedback, suggesting that participants internally monitor the probability of obtaining a reward. This observation is consistent with the EVC-theory (Shenhav et al., 2013) which states that the investment of cognitive control depends on the potential benefits that it brings (i.e., the reward). However, what these behavioural results do not show is whether this effect occurs because of subjective difficulty monitoring or conflict monitoring.

Next, the ERP analysis of the stimulus-locked data showed the same patterns as in the study by Desender et al. (2016). So, based on this study, there was an a priori hypothesis of how the ERP-signal would look like. We indeed were able to replicate a conflict-sensitive component at around 250 ms (i.e., N2) and a subjective difficulty monitoring component at around 360 ms (i.e., P3). Even though the experimental designs of both studies are not identical, we still found the same results. This suggests that these components are not solely found in this study. It should, however, be noted that the P3 component could not be found using a cluster-based permutation test, but only with an a priori time-window and specific electrodes. This a priori time-window and the specific electrode were decided on using the results from the previously mentioned paper (Desender et al., 2016). So, even though this effect was not sufficiently robust to be picked up by the cluster-based permutation test, it is still sufficiently robust when focusing on a priori time window. Topographical plots have also shown that the N2 component is more frontocentral while the P3 component is more frontal. So, the difference in time and topographies of the N2 and P3 components suggest that the brain has different processes for conflict and subjective difficulty. For the response-locked ERP analysis, we confirmed the presence of the P3 component right before the response. But there were no significant components after the response.

The temporal generalisation of congruency showed that we could predict congruency in the rating phase with an accuracy up to 70%. The temporal generalisation of difficulty ratings showed we could decode the subjective difficulty with an above chance level. Lastly, the temporal generalisation of the following choice showed that we could predict the following choice with and accuracy up to 55%. For the response-locked data, there were also significant clusters that reversed in polarisation. This isn't the most important part of the study, but should also not be ignored. The pattern of the temporal generalisation was located on the diagonal, which suggests that these are isolated/chain processes (King & Dehaene, 2014).

Lastly, the cross-temporal generalisation of congruency and rating on following choice showed no significant results. This suggests that we cannot decode a participant's following choice (i.e., low effort – low gain or high effort – high gain) based on the neural data of congruency and subjective difficulty ratings. This observation is unexpected, given that the EVC theory by Shenhav et al. (2013) predicts that effort investment depends on the detection of conflict, which is believed to be mediated by the anterior cingulate cortex (i.e., the N2 component). We can only speculate about why we were unable to find significant across-decoding results.

One reason why we could not predict a participant's following choice based on the neural data of congruency and subjective difficulty ratings could be that the EVC theory by Shenhav et al. (2013) is incorrect and conflict processing does not have an influence on effort choices. Or at least not sufficiently to alter the neural data of congruency and subjective difficulty in this study design. It is also possible that the € 10 and € 2 context were not perceived as high effort – high gain and low effort – low gain. Possibly, the difference between the two rewards was not enough. Lastly, by removing data to make sure the prediction was orthogonal, we may have lost too much data. While we had enough data overall, this may have lead to insufficient data to train and test the predictor.

Conclusion

Based on this study, we could not predict a participant's following choice based on the neural data of congruency and subjective difficulty ratings. This does not mean that this is impossible and thus requires further research.

References

Balota, D. A. & Marsh, E. J. (2004). Cognitive Psychology: Key Readings (1st Edition). Hove, United Kingdom: Psychology Press

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 Verb Generation. *Journal of Cognitive Neuroscience*, *12*(2), 298–309. https://doi.org/10.1162/089892900562110

Botvinick, M. M. (2007). Conflict monitoring and decision making: Reconciling two perspectives on anterior cingulate function. *Cognitive, Affective and Behavioral Neuroscience*, *7*(4), 356–366. https://doi.org/10.3758/CABN.7.4.356

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), 624–652. https://doi.org/10.1037//0033-295X.I08.3.624

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, *8*(12), 539–546. https://doi.org/10.1016/j.tics.2004.10.003

Botvinick, M. M., & Rosen, Z. B. (2009). Anticipation of cognitive demand during decision-making. *Psychological Research Psychologische Forschung*, *73*(6), 835–842. https://doi.org/10.1007/s00426-008-0197-8

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the stroop effect. *Psychological Review*, *97*(3), 332–361. https://doi.org/10.1037/0033-295X.97.3.332

Del Cul, A., Baillet, S., & Dehaene, S. (2007). Brain Dynamics Underlying the Nonlinear Threshold for Access to Consciousness. *PLoS Biology*, *5*(10), 2408–2423. https://doi.org/10.1371/journal.pbio.0050260

Desender, K., Calderon, C. B., 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*, *43*(7), 1397–1410. https://doi.org/10.1037/xhp0000391

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. https://doi.org/10.1016/j.neuropsychologia.2016.01.003

Desender, K., Van Opstal, F., & Van den Bussche, E. (2014). Feeling the Conflict: The Crucial Role of Conflict Experience in Adaptation. *Psychological Science*, *25*(3), 675–683. https://doi.org/10.1177/0956797613511468

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, *64*(1), 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General*, *127*(3), 286–303. https://doi.org/10.1037/0096-3445.127.3.286

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*(4), 480–506. https://doi.org/10.1037/0096-3445.121.4.480

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding Dynamic Brain Patterns from Evoked Responses: A Tutorial on Multivariate Pattern Analysis Applied to Time Series Neuroimaging Data. *Journal of Cognitive Neuroscience*, *29*(4), 677–697. https://doi.org/10.1162/jocn_a_01068

Heil, M., Hennighausen, E., Osman, A., Wiegelmann, J., & Rolke, B. (2000). N200 in the Eriksen-Task: Inhibitory Executive Processes? *Journal of Psychophysiology*, *14*(4), 4218–4225. https://doi.org/10.1027/0269-8803.14.4.218

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior Cingulate Conflict Monitoring and Adjustments in Control. *Science*, *303*(5660), 1023–1026. https://doi.org/10.1126/science.1089910

King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: the temporal generalization method. *Trends in Cognitive Sciences*, *18*(4), 203–210. https://doi.org/10.1016/j.tics.2014.01.002

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology. General*, *139*(4), 665–682. https://doi.org/10.1037/a0020198

LaBerge, D. (1990). Thalamic and cortical mechanisms of attention suggested by recent positron emission tomographic experiments. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn.1990.2.4.358

Liotti, M., Woldorff, M. G., Perez, R., & Mayberg, H. S. (2000). An ERP study of the temporal course of the Stroop color-word interference effect. *Neuropsychologia*, *38*(5), 701–711. https://doi.org/10.1016/S0028-3932(99)00106-2

Loftus, E. F., & Klinger, M. R. (1992). Is the Unconscious Smart or Dumb? *American Psychologist*, *47*(6), 761–765. https://doi.org/10.1037/0003-066X.47.6.761

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443–447. https://doi.org/10.1126/science.1100301

Salamone, J. D., Correa, M., Farrar, A., & Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology*, *191*(3), 461–482. https://doi.org/10.1007/s00213-006-0668-9

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, *84*(1), 1–66.

https://doi.org/10.1037/0033-295X.84.1.1

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. https://doi.org/10.1016/j.neuron.2013.07.007

Stürmer, B., Leuthold, H., Soetens, E., Schröter, H., & Sommer, W. (2002). Control over location-based response activation in the Simon task: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(6), 1345–1363. https://doi.org/10.1037/0096-1523.28.6.1345

Veen, V. van, & Carter, C. S. (2002). The Timing of Action-Monitoring Processes in the Anterior Cingulate Cortex. *Journal of Cognitive Neuroscience*, *14*(4), 593–602. https://doi.org/10.1162/08989290260045837

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), 931–959. https://doi.org/10.1037/0033-295X.111.4.931

Zakrzewski, A. C., Wisniewski, M. G., Iyer, N., & Simpson, B. D. (2019). Confidence tracks sensory- and decision-related ERP dynamics during auditory detection. *Brain and Cognition*, *129*(December 2018), 49–58. https://doi.org/10.1016/j.bandc.2018.10.007