Electrophysiological indices of anterior cingulate cortex function reveal changing levels of cognitive effort and reward valuation that sustain task performance

Akina Umemoto[a,b,*], Michael Inzlicht[c,d], Clay B. Holroyd[a]

[a] Department of Psychology, University of Victoria, Canada
[b] Department of Psychology, University of Toyama, Japan
[c] Department of Psychology, University of Toronto, Canada
[d] Rotman School of Management, University of Toronto, Canada

ARTICLE INFO

ABSTRACT

Successful execution of goal-directed behaviors often requires the deployment of cognitive control, which is thought to require cognitive effort. Recent theories have proposed that anterior cingulate cortex (ACC) regulates control levels by weighing the reward-related benefits of control against its effort-related costs. However, given that the sensations of cognitive effort and reward valuation are available only to introspection, this hypothesis is difficult to investigate empirically. We have proposed that two electrophysiological indices of ACC function, frontal midline theta and the reward positivity (RewP), provide objective measures of these functions. To explore this issue, we recorded the electroencephalogram (EEG) from participants engaged in an extended, cognitively-demanding task. Participants performed a time estimation task for 2 h in which they received reward and error feedback according to their task performance. We observed that the amplitude of the RewP, a feedback-locked component of the event related brain potential associated with reward processing, decreased with time-on-task. Conversely, frontal midline theta power, which consists of 4–8 Hz EEG oscillations associated with cognitive effort, increased with time-on-task. We also explored how these phenomena changed over time by conducting within-participant multi-level modeling analyses. Our results suggest that extended execution of a cognitively-demanding task is characterized by an early phase in which high control levels foster rapid improvements in task performance, and a later phase in which high control levels were necessary to maintain stable task performance, perhaps counteracting waning reward valuation.

1. Introduction

Goal-directed behavior often requires cognitive control in order to facilitate the execution of non-automatic behaviors (Norman and Shallice, 1986). It is believed that the application of control feels effortful (e.g., Botvinick and Braver, 2015; Shenhav et al., 2017) and that cognitive effort is inherently aversive, such that people tend to avoid cognitively demanding tasks (Kool et al., 2010; Inzlicht et al., 2015; McGuire and Botvinick, 2010; Westbrook and Braver, 2015). This process is thought to recruit a mechanism that weighs the benefits of applying control against the effort-related costs in doing so (Botvinick and Braver, 2015; Shenhav et al., 2017; Westbrook and Braver, 2015). For example, although prolonged cognitive effort normally results in mental fatigue, which disrupts task performance, these performance decrements can be counteracted if subjects are offered motivational incentives (Boksem et al., 2006; Hockey, 2011; Hopstaken et al., 2015; Lorist et al., 2005; Boksem and Tops, 2008, for review). Yet despite an upsurge of interest in this topic in recent years, the neurocognitive mechanisms that sustain cognitive effort are not well-understood.

2. Anterior cingulate cortex function in reward processing and effortful control

Recently, several theories and computational models have proposed that anterior cingulate cortex (ACC) may provide such a mechanism (e.g., Botvinick and Braver, 2015; Holroyd and McClure, 2015; Holroyd and Umemoto, 2016; Holroyd and Yeung, 2012; Shenhav et al., 2013, 2017; Vassena et al., 2017; Verguts et al., 2015). The function of ACC is famously controversial (Ebitz and Hayden, 2016). Nevertheless, accumulating evidence suggests that the caudal subdivision of ACC, which is

formally known as "anterior midcingulate cortex" (Vogt, 2009; see also Shackman et al., 2011), may serve as a computational hub that integrates cognitive processes related to motivation and control. For example, Holroyd and Yeung (2012) proposed that ACC is responsible for motivating and selecting extended action sequences based on learned task values. According to this view, instead of concerning itself with the minutia of moment-to-moment behaviors (such as typing each letter in a manuscript), ACC regulates behavior at a higher level of abstraction (such as whether or not to write the manuscript at all), and determines how well the task should be performed at a global level (cf. Botvinick et al., 2009; Botvinick, 2012). Holroyd and McClure (2015) later implemented these ideas in a computational model of rodent behavior. Although these theories differ in their specifics, they hold in common the idea that ACC regulates the control levels it invests in a task according to the rewards received for doing so.

A challenge in studying these neurocognitive processes, however, is that the sensations of reward valuation and effort expenditure are available only to introspection, rendering them difficult to assess empirically. Task performance is an imperfect proxy for these processes because participants can perform a task well either because it is easy or because they apply enough effort to make it appear to be easy. Therefore, objective measures of reward valuation and cognitive effort, were they available, would provide insight into how the control system self-regulates (Botvinick and Braver, 2015). Here we propose that electrophysiological correlates of cognitive effort and of reward valuation can fulfill this purpose, which we explore by recording the electroencephalogram (EEG) from subjects as they perform an extended, mentally-fatiguing task.

3. Electrophysiological correlates of reward valuation and cognitive effort

We explored this question by utilizing two electrophysiological correlates of ACC activity: the reward positivity (RewP) and frontal midline theta (FMT) oscillations (Holroyd and Umemoto, 2016). More commonly known as the feedback error-related negativity, the RewP is an event-related potential (ERP) component that is differentially modulated by feedback stimuli with negative vs. positive valence (Miltner et al., 1997); the component has recently been renamed the RewP because it appears to be relatively more sensitive to reward outcomes than to error outcomes (Holroyd and Umemoto, 2016). Holroyd and Coles (2002) proposed that the RewP is produced by fast, phasic midbrain dopamine reward prediction error signals modulating ACC activity. Consistent with the theory, numerous studies have confirmed that the RewP indexes a reward prediction error signal (Sambrook and Goslin, 2015; Walsh and Anderson, 2012; but see Janssen et al., 2016). Given the inverse problem, the neural source of the RewP is less clear, but the balance of evidence suggests that it is produced in anterior midcingulate cortex (e.g., Miltner et al., 1997; Becker et al., 2014; but see Proudfit, 2015). Of particular relevance to this study, RewP amplitude correlates positively with individual differences in reward sensitivity (Bress and Hajcak, 2013; Umemoto and Holroyd, 2017; see also Cooper et al., 2014; Liu et al., 2014; Parvaz et al., 2016), and negatively with individual differences in depression levels (e.g., Umemoto and Holroyd, 2017; Proudfit, 2015), which suggests that the RewP may index subjective levels of reward valuation (see Holroyd and Umemoto, 2016 for review).

FMT consists of 4–8 Hz EEG oscillations distributed over frontal-central regions of the human scalp. Source localization studies of FMT point to ACC as the neural generator (e.g., Asada et al., 1999; Cavanagh and Frank, 2014; Ishii et al., 1999; Luu and Tucker, 2001), and FMT power appears to reflect an effortful control process (e.g., Cavanagh and Frank, 2014; Holroyd and Umemoto, 2016). Notably, cognitive tasks elicit two kinds of related FMT signals: phasic and ongoing. Brief, phasic changes in FMT power are observed following receipt of negative feedback stimuli (indicating, e.g., no-reward or punishment) (e.g.,

Bernat et al., 2015; Foti et al., 2015). A similar burst of FMT power is also observed immediately following error commission and during periods of response conflict (e.g., Cavanagh and Frank, 2014; Luu and Tucker, 2001; see also Lin et al., 2017), which according to one study decreases with time-on-task (Wascher et al., 2014). By contrast, ongoing FMT power is observed over extended periods of task execution (e.g., Asada et al., 1999; Ishii et al., 1999), as well as during the off-task, resting state (Scheeringa et al., 2008). It is associated with sustained mental effort, as observed, for example, when participants perform arithmetic calculations (Hsieh and Ranganath, 2014; Mitchell et al., 2008). Furthermore, FMT power rises with time on task (e.g., Barwick et al., 2012; Paus et al., 1997; Wascher et al., 2014), suggesting that ACC contributes to sustaining effortful behavior in the face of growing mental fatigue.

3.1. Present study

We have previously proposed that FMT reflects the control signal applied by ACC over task performance, and that RewP amplitude reflects the reward signal propagated to the ACC for the purpose of regulating the control level (Holroyd and Umemoto, 2016). In this way, ACC is well-positioned to motivate task performance (Holroyd and McClure, 2015; Holroyd and Yeung, 2012). Although previous studies have utilized these signals to study reward valuation and cognitive effort in isolation, to our knowledge none have examined how these processes are involved in sustaining behavior within the same, mentally fatiguing task. For this purpose, we recorded the EEG from participants while they performed a standard time estimation task for 2 h (Miltner et al., 1997). We predicted that FMT power would increase (Barwick et al., 2012; Paus et al., 1997; Wascher et al., 2014) and RewP amplitude would decrease with time-on-task due to increasing levels of fatigue. We further predicted that RewP amplitude and FMT power would interact, reflecting the neural processes that sustain task performance. However, most current theories of ACC function do not provide detailed predictions about the specifics of this interaction (e.g., Botvinick and Braver, 2015, Holroyd and Umemoto, 2016; Holroyd and Yeung, 2012; Shenhav et al., 2013, 2017; Vassena et al., 2017). That said, one computational model of rodent ACC suggests a complex process by which control levels increase in response to a decreasing reward rate (in order to maintain a high average reward), until the average reward itself decreases, which results in falling levels of control (because the elevated control is unlikely to yield more reward; Holroyd and McClure, 2015; see also Verguts et al., 2015). We therefore explored whether the relationship between FMT power and RewP amplitude would reflect similar dynamics. However, to foreshadow our results, we failed to find such a relationship.

Further, because P300 amplitude is believed to reflect attentional resource allocation and the motivational significance of a task, we also measured P300 amplitude as an index of task engagement (Murphy et al., 2011; Polich, 2007). Consistent with past fatigue studies, we expected P300 amplitude to decrease with time on task (e.g., Hopstaken et al., 2015, 2016). Lastly, we explored whether ACC function is expressed as a dimension of personality related to reward sensitivity and persistence, as we have recently proposed (Holroyd and Umemoto, 2016; see also Umemoto and Holroyd, 2016, 2017). For this purpose, participants answered several personality questionnaires related to motivation, reward sensitivity, and levels of depression, which we explored in relation to their behavior and to these electrophysiological signatures of ACC function.

4. Materials and methods

4.1. Participants

Sixty five undergraduate students were recruited from the University of Victoria. The sample size was determined based on past

studies examining individual differences related to the RewP (e.g., Bress and Hajcak, 2013; Nelson et al., 2016). A sensitivity analysis using the G*power software program (Erdfelder et al., 1996) indicated sufficient statistical power to detect small effects for the main within-subject analyses.[1] Participants were recruited from the Department of Psychology subject pool to fulfill a course requirement or earn bonus credits. All subjects (16 males, 10 left-handed, age range = 17–23 years, mean age = 19.3 +/− 1.5 years) had normal or corrected-to-normal vision. Each also received a monetary bonus in addition to the credits, the amount of which depended on their task performance (see below). All of the subjects provided informed consent as approved by the local research ethics committee. The experiment was conducted in accordance with the ethical standards prescribed in the 1964 Declaration of Helsinki.

4.2. Task design

We conducted a time estimation task that required participants to estimate 1 s on every trial (Miltner et al., 1997). Each trial started with a presentation of a visual cue in the form of a white central cross (1° by 1° square visual angle) on the computer screen. Participants were instructed to press a left button on a computer mouse using their right hand when they believed that 1 s had elapsed following the cue onset. After 600 ms following the response, they were presented with a feedback stimulus indicating whether their response was on-time or not on-time. The response was considered on-time if it occurred within a narrow window of time centered around 1 s, the size of which was adjusted from trial to trial according to a staircase procedure, as follows. At the start of the experiment the size of the window was initialized at 200 ms, such that responses occurring between 900 and 1100 ms were considered correct. The size of the time window (i.e., time-window size) was then adjusted depending on the participant's performance: error responses caused the time window to increase by 10 ms (making the task easier), and correct responses caused the time-window to decrease by 10 ms (making the task harder). This manipulation ensured that participants received reward feedback on approximately 50% of the trials (Miltner et al., 1997). The feedback stimuli were represented by abstract symbols presented at fixation (3.3° by 3.3° square visual angle). For half of the participants, a white circle indicated that they earned 1 cent for that trial, and a white diamond indicated that they did not earn any reward for that trial. This mapping was reversed for the other half of the participants. After a variable inter-trial interval of 1400 ms, 1500 ms, or 1600 ms, determined at random, the next trial began with a presentation of the visual cue.

Further, we offered an additional motivational incentive to enhance individual differences in task performance by providing all of the participants an opportunity to participate in a lottery to win a CAN $100 Amazon.com gift card. Every participant earned at least 1 "ticket" regardless of their performance, which ensured that all subjects had at least a small chance of receiving the prize. In addition, the participants were told (truthfully) that the better performers would earn extra tickets. For this, when all of the participants' data were collected, we calculated each participant's mean time-window size across all trials and compared them against the grand mean time-window size averaged across all participants. Participants received an extra ticket for every 5 ms decrement with respect to the grand mean time-window size. For example, an individual average time-window size that was 15 ms less than the grand mean time-window size would earn that participant 4 tickets total (3 additional tickets plus the baseline 1 ticket). Upon completing the experiment, all of the tickets were entered into a metaphorical lottery box, from which we randomly selected 2 winning

participants. The two winners were then contacted and received the award via email afterword.

4.3. Task procedure

Participants were seated comfortably in front of a computer monitor (1024 by 1280 pixels) at a distance of about 60 cm in a dimly lit room. The task was programmed in Matlab (MathWorks, Natick, MA, USA) using the Psychophysics Toolbox extension (Brainard, 1997; Pelli, 1997). Before the experiment was described to them, all of the participants read a form that explained the opportunity to win a CAN $100 Amazon gift card. Participants were told that if they decided to consent to participate in this opportunity, at least one ticket would be entered in their name into the "lottery", and that additional tickets would be entered if their overall performance was better than the average performance across all of the participants. Consent was indicated by participants providing a contact email address on the form; only two participants did not provide consent.

Participants were then provided with both written and verbal instructions about the time estimation task. They practiced the task for 20 trials before starting the task proper, which consisted of 16 blocks of 95 trials each. Participants were not informed about the exact number of trials or blocks to complete, but instead were told that they would perform the task for about 2 h. Self-paced between-block rest periods were provided, and after about 1 h participants relaxed during a longer rest period while the experimenter checked the electrode impedances. Participants were told that the reward they accumulated across trials would be paid out to them at the end of the experiment, and that they should estimate 1 s on each trial as accurately as possible in order to maximize their reward earnings. Participants were informed about their winnings only at the end of the 2-h experiment (no feedback about their performance or winnings was provided during the experiment). Upon completing the experiment they answered several personality questionnaires (see below). These were followed by a post-experiment paper-and-pen questionnaire that asked about the participant's overall experience of the experiment, the strategies they employed (if any), and their level of task engagement on a scale of 1–5, with 1 indicating not at all engaged and 5 indicating very engaged.

4.4. Questionnaires

Participants completed a total of six personality questionnaires related to motivation, reward sensitivity, and depression symptoms, administered through LimeSurvey (https://www.limesurvey.org/) on the same computer where the task was performed. The personality questionnaires administered included 1) the Persistence Scale (Cloninger et al., 1993; Gusnard et al., 2003), which assesses the tendency to overcome daily challenges on a scale of 1 (definitely false) to 5 (definitely true). 2) The 8-item Reward Responsiveness (RR) Scale (Van den Berg et al., 2010), which is a self-report measure of reward-related behavior on a scale of 1 (strong disagreement) to 4 (strong agreement). (3) The 18-item Temporal Experience of Pleasure Scale (TEPS), which assesses two trait measures related to hedonic capacity, namely consummatory pleasure (TEPS-C: i.e., "liking" or in-the-moment experience of pleasure) and anticipatory pleasure (TEPS-A: i.e., "wanting"), on a scale of 1 ("very false for me") to 6 ("very true for me") (Gard et al., 2006). 4) The 18-item Apathy Evaluation Scale (AES; Marin et al., 1991), which measures lack of motivation regarding the behavioral, cognitive, and emotional aspects of goal-directed behavior on a scale of 1 (very characteristic) to 4 (not at all characteristic). 5) The 22-item Ruminative Responses Scale (RRS: Treynor et al., 2003), which measures the propensity to ruminate in response to depressed mood on a scale of 1 (almost never) to 4 (almost always). 6) The 21-item short-form of the Depression Anxiety Stress Scale (DASS-21) (Lovibond and Lovibond, 1995), which measures severity of depression, anxiety, and stress on a scale from 0 ("did not apply to me at all") to 3 ("applied to

---

[1] The results of these analyses are provided in the Supplementary materials. Note that this sample size only provides adequate power to detect medium-to-large effects related to individual differences in personality.

me very much, or most of the time"). However, the rumination scale and the anxiety and stress subscales of the DASS-21 were not included in the analyses as they tended to strongly correlate with other variables (for example, with the depression scores), and because they were not the primary focus of the study. Summed total scores were used for each of the questionnaires such that high scores indicated, respectively, high persistence, high reward responsiveness, high hedonic capacity (or re-duced anhedonia), high levels of apathy, and high levels of depression.

4.5. EEG data acquisition and pre-processing

The EEG was recorded using a montage of 30 electrode sites in accordance to the extended international 10–20 system (Jasper, 1958). Signals were acquired using Ag/AgCl ring electrodes mounted in a nylon electrode cap with an abrasive, conductive gel (EASYCAP GmbH, Herrsching-Breitbrunn, Germany). Signals were amplified by low-noise electrode differential amplifiers with a frequency response high cut-off at 50 Hz (90 dB–octave roll off) and digitized at a rate of 250 samples per second. Digitized signals were recorded to disk using Brain Vision Recorder software (Brain Products GmbH, Munich, Germany). Inter-electrode impedances were maintained below 20 kΩ. Two electrodes were also placed on the left and right mastoids. The EEG was recorded using the average reference. The electrooculogram (EOG) was re-corded for the purpose of artifact correction; horizontal EOG was re-corded from the external canthi of both eyes, and vertical EOG was recorded from the suborbit of the right eye and electrode channel Fp2.

4.6. Data analysis

4.6.1. Behavior

To recap, participants' estimation of 1 s was considered on-time if it occurred within a narrow window of time centered around 1 s, the size of which was initialized at 200 ms, such that responses occurring be-tween 900 and 1100 ms were considered correct at the start of the experiment. The time-window size was adjusted from trial to trial ac-cording to a staircase procedure that ensured 50% probabilities of re-ceiving positive and negative feedback (see Task Design). The mean time-window size averaged across trials for each block was calculated for each participant. Therefore, smaller and larger average time-window sizes indicate, respectively, better and worse performance.

4.6.2. Electrophysiology

Post-processing and data visualization were performed using Brain Vision Analyzer software (Brain Products GmbH). The digitized signals were filtered using a fourth-order digital Butterworth filter with a passband of 0.10–30 Hz. The data were segmented for an 800 ms epoch extending from 200 ms prior to 600 ms following presentation of re-ward and no-reward feedback. Ocular artifacts were corrected using an eye movement correction algorithm (Gratton et al., 1983). The EEG data were re-referenced to averaged mastoids electrodes. Data were baseline corrected by subtracting from each sample for each channel the mean voltage associated with that electrode during the 200 ms in-terval preceding feedback onset. Trials with muscular and other arti-facts were excluded according to a 150 μV Max-Min voltage difference, a ± 150 μV level threshold, a ± 35 μV step threshold, and a 0.1 μV lowest-allowed activity level as rejection criteria. ERPs were then cre-ated for each electrode and participant by averaging the single-trial EEG according to the reward and no-reward feedback type.

The RewP was measured at channel FCz, where it reached max-imum amplitude, utilizing a difference wave approach that isolated the RewP from overlapping ERP components such as the P300 (Holroyd and Krigolson, 2007; Sambrook and Goslin, 2015); a difference wave was created for each participant by subtracting the ERP to no-reward feedback stimuli from the ERP to reward feedback stimuli. RewP am-plitude was then determined by averaging the mean amplitude in the difference wave from 200 to 300 ms following feedback onset

(determined based on a visual inspection of the grand-average of the difference wave across all participants, which ensured that the grand-average RewP in this time range was characterized by a positive peak with a frontal-central voltage distribution). P300 amplitude was mea-sured at channel Pz, where it typically reaches maximum amplitude across the scalp (Donchin and Coles, 1988). Individual subject ERPs were averaged across feedback conditions, and P300 amplitude was defined by finding the maximum positive deflection from 280 to 420 ms following feedback onset (as determined based on a visual inspection of the grand-average ERP across all trials and all participants).[2] In order to assess FMT power, the continuous EEG data were segmented into consecutive 4 s epochs with 200 ms overlap between segments (starting from the beginning of the experiment and continuing to the end) averaged across feedback types.[3] Artifact rejection and ocular correc-tion were conducted on these EEG epochs as for the ERP data, and then submitted to a power spectral analysis using a Fast Fourier Transform (FFT) (Hanning Window, 10% length). FMT was assessed for each trial by averaging power between 4 and 8 Hz at each channel.

We employed a multi-level modeling analysis using the MIXED function in SPSS (IBM SPSS 24) in order to examine the effect of within-person changes in electrophysiological measures (i.e., RewP, FMT) on behavior (i.e., time-window size) across the 16 blocks. To assess how these measures fluctuated across time, we calculated person-centered scores for the predictor variables (i.e., RewP, FMT) such that mean performance averaged across all blocks of trials of a given participant was subtracted from the mean performance averaged across trials for each block (e.g., adjusted RewP amplitude for block 1 = average RewP amplitude for block 1 – the average RewP amplitude across all 16 blocks) (e.g., Saunders et al., 2015). By doing this, each participant's score for a given predictor was calculated relative to their own average score, with positive and negative values indicating higher and lower values than average for that participant, respectively. Note that cen-tering facilitates interpretation of the model's interaction terms (Aiken and West, 1991), which relate to the mean of each individual subject (within-subjects) rather than to the mean across subjects. That is, the interaction effects are interpreted in terms of how much each predictor variable (e.g., FMT power) is above or below the mean for each parti-cipant (see below). Our primary analysis of interest investigated how within-person fluctuations in RewP amplitude and FMT power related to changes in time-window size with time on task, with the time (block) variable coded both as a linear slope from 0 to 15 (reflecting the 16 blocks in total) and as a quadratic slope, as calculated as the square of the linear-time variable (i.e., 0^2 to 15^2).[4] For completeness, we also explored how the behavioral and electrophysiological measures related to each other by separately testing RewP amplitude and FMT power as a function of the other two variables and both the linear and quadratic time variables. For example, a model testing the time-course of RewP amplitude included time-window size, FMT power, linear-time, and quadratic-time as its predictors. All of the analyses were conducted with a random intercept for each participant,[5] unstructured covariance type, and maximum likelihood estimation. Both time (block) variables were treated as random slopes. Effect size, *r*, is reported for each model

---

[2] Note that the scalp distribution for P300 was not maximal centrally but rather over lateral scalp areas (e.g., electrodes P3 and P4). However, given that the P300 was not the focus of our analysis, P300 amplitude was measured at channel Pz in keeping with pre-vious studies, and because P300 amplitude was maximal there when comparing only the midline channels.

[3] When we segmented the EEG data with a 50% overlap between segments for the frontal midline theta analyses, there was no significant difference in the overall power (p = 0.415), and all the results remained the same.

[4] Mean-centering the block/time variable (i.e., coding it from −8 to +8 for the linear slope and from -8^2 to $+8^2$ for the quadratic slope) did not materially change the results. Therefore, we based both the linear and the quadratic effects on block/time 1 to make interpretation of the results easier.

[5] Random intercepts are used on dependent variables (DVs) so that the data for each person is expressed as a value that is relative to their own mean.

effect.

We also applied multi-level modeling analyses to examine between-subjects relationships (rather than within-subject relationships, as described above) among time-window size, RewP amplitude, and FMT power.[6] Each participant's average RewP amplitude and FMT power values were calculated separately across all the trials. Each score was then grand mean-centered for each participant by subtracting the grand average (across all participants) from each measure (e.g., the grand mean-centered RewP amplitude for a participant = the mean RewP amplitude averaged across all trials for this participant – the grand mean RewP amplitude averaged across all participants). We tested time-window size by RewP amplitude and FMT power, using a random intercept for each participant, unstructured covariance type, and maximum likelihood estimation. For completeness, we also tested models, separately, for RewP amplitude and FMT power as a function of the other two variables (e.g., predicting FMT power by time-window size and RewP amplitude).

We also explored the effects of personality traits using multi-level modeling analyses, the results of which are reported in the Supplementary materials.

5. Results

Three participants discontinued the task.[7] The data of three other participants were excluded from analysis due to self-reported neurological or psychiatric disorders. Further, during the first few blocks of trials one participant produced extreme time estimates (i.e., by responding about 7 s after cue onset); performance for this participant improved thereafter, but the probability of reward was strongly biased across blocks of trials due to the staircase procedure (mostly no reward in the first few blocks, followed by mostly reward in the following few blocks). The data of two participants were also excluded due to artifacts associated with excessive head movement. Further, we excluded data from three participants with average time-window size or FMT power that was 3 standard deviations (SD) above the group means. In total these exclusions resulted in the data of 53 participants used for all of the analyses.

5.1. Behavior

5.1.1. Time-window size

Fig. 1a depicts the time-course of the mean time-window size averaged across participants. The multi-level modeling analysis predicting the time-window size with RewP, FMT power, linear-time, and quadratic-time revealed significant main effects of linear-time and quadratic-time (Table 1, lines 1 & 2). These results confirm the impression of Fig. 1a that participants' performance improved with time but worsened towards the end of the experiment. We also found significant interactions of linear- and quadratic-time with FMT power on time-window size (Table 1, lines 5 & 7).[8] Fig. 2 reveals the nature of these effects. The relation between FMT power and time-window size

---

[6] Note that this results in a single/participant level remaining. Hence when this analysis was conducted using the standard linear regression by predicting each dependent variable by the remaining variables entered together in a model, the (null) results stayed the same, as expected (all p > 0.7).

[7] One participant indicated signs of claustrophobia, one participant left feeling unwell, and one participant discontinued after reporting that s/he had earned enough money.

[8] Note that when we performed the same analysis while retaining the data of the outliers (that were removed based on the 3 SD criterion), we no longer observed reliable quadratic-time effects (r = 0.22 including outliers, and r = 0.42 excluding outliers) nor the interaction effects with linear-time/quadratic-time and FMT (r = 0.04/0.04 including outliers, and r = 0.09/0.07 excluding outliers). These interaction effects are therefore less robust when the data of the participants who produced extreme scores were included in the analyses. This is not surprising given that the effect sizes for the interaction terms were small to being with, and the inclusion of extreme scores increased the variability to explain, which necessarily reduced the effect sizes.
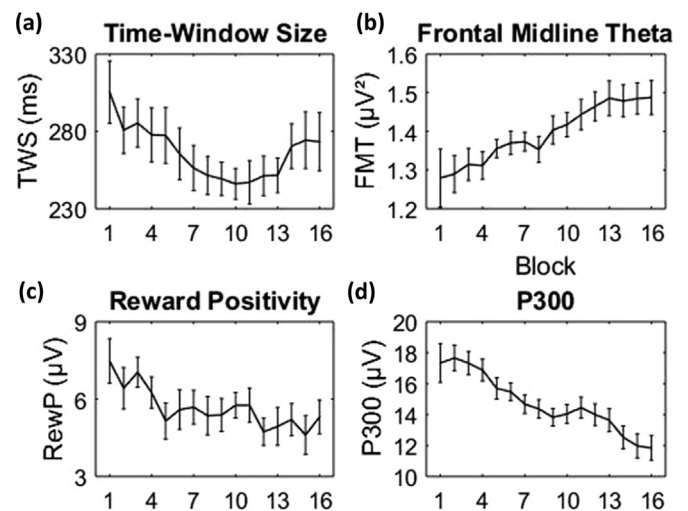
Fig. 1. Block by block performance and electrophysiological measures, averaged across subjects for (a) time-window size, (b) frontal midline theta power, (c) reward positivity amplitude, and (d) P300 amplitude. Error bars indicate within-subject 95% confidence intervals (Cousineau, 2005).

differed across time, such that participants performed the task better (i.e., time-window size was smaller) at the start of the task when they produced more FMT power relative to their average FMT power across the task, as compared to when they produced less FMT power relative to their average FMT power across the task. Note that this observation holds even though the participants generally performed the task worse at the start of the experiment compared to at the end (Fig. 1a). Further, their performance was more stable later in the task (exhibiting relatively shallow changes with time) when they produced greater FMT power relative to their average FMT power across the task as compared to when they produced less FMT power relative to their average FMT power across the task. A significant main effect of FMT power (Table 1, line 3) signifies that the difference in time-window size by FMT power in the beginning was statistically significant.

Additionally, we conducted several analyses that explored whether FMT power and RewP amplitude might interact to modulate performance. First, we tested the model predicting time-window size by adding to the original model terms capturing the interaction of FMT power with RewP amplitude, and the 3-way interaction of these predictors with linear- and quadratic-time (i.e., FMT, RewP, FMT-RewP interaction, time-linear/quadratic effects, and FMT-RewP-time – both linear and quadratic – interactions as predictors of time-window size). However, this analysis yielded only significant main effects of time (both linear and quadratic, both p < 0.001), and the interaction terms did not significantly predict time-window size (all p > 0.17). Second, in order to evaluate whether the failure to observe interactions with RewP and FMT were because they were both highly correlated with time on task, we tested the model predicting time-window size by including the interaction of FMT and RewP as predictors while excluding the time variable (i.e., FMT, RewP, and FMT-RewP interaction as predictors of time-window size). Nevertheless, although this analysis revealed that increased FMT power predicted smaller (better) time-window size (p = 0.039), the interaction term failed to predict time-window size (p = 0.433). Third, we explored possible mediation effects of FMT and/or RewP on time-window size by conducting a piecewise structural equation model (Lefcheck, 2016). This model tested whether time (both linear and quadratic effects) predicted time-window size via RewP amplitude and FMT power (i.e., time→RewP→FMT→time-window size). However, the model fit was poor (i.e., the model significantly differed from the actual data with p = 0), indicating that RewP amplitude did not mediate the effect of FMT on behavior (see Supplementary Fig. 5). These results suggest that RewP amplitude and

Table 1
Multi-level modeling results on behavioral and electrophysiological measures over time.

DV	Predictors	b	SE	95% Cis	r	pVal	Line
TWS	**Block-L**	**−10.28**	**2.41**	**[−15.1, −5.47]**	**0.478**	**< 0.001**	1
	Block-Q	**0.6**	**0.16**	**[0.27, 0.93]**	**0.424**	**0.001**	2
	FMT	**−78.3**	**30.22**	**[−137.9, −18.71]**	**0.181**	**0.01**	3
	RewP	2.04	1.82	[−1.53, 5.62]	0.043	0.263	4
	Block-L*FMT	**22.64**	**8.89**	**[5.2, 40.09]**	**0.093**	**0.011**	5
	Block-L*RewP	−0.9	0.57	[−2.02, 0.22]	0.057	0.115	6
	Block-Q*FMT	**−1.25**	**0.61**	**[−2.45, −0.05]**	**0.074**	**0.041**	7
	Block-Q*RewP	0.06	0.04	[−0.01, 0.14]	0.061	0.093	8
FMT	**Block-L**	**0.01**	**0.003**	**[0.01, 0.02]**	**0.532**	**< 0.001**	9
	TWS	**−0.0003**	**0.0001**	**[−0.0006, −0.0001]**	**0.084**	**0.019**	10
	RewP	0.004	0.003	[−0.001, 0.01]	0.053	0.142	11
	Block-L*TWS	**0.00004**	**0.00002**	**[0.00001, 0.00007]**	**0.091**	**0.012**	12
	Block-L*RewP	−0.0004	0.0004	[−0.001, 0.0003]	0.037	0.304	13
RewP	**Block-L**	**−0.35**	**0.09**	**[−0.53, −0.17]**	**0.453**	**< 0.001**	14
	Block-Q	**0.01**	**0.01**	**[0.004, 0.02]**	**0.326**	**0.009**	15
	TWS	0.007	0.004	[−0.001 0.01]	0.070	0.087	16
	FMT	1.27	1.46	[−1.61, 4.16]	0.064	0.385	17
	Block-L*TWS	**−0.003**	**0.001**	**[−0.006, −0.0006]**	**0.087**	**0.015**	18
	Block-L*FMT	0.05	0.45	[−0.83, 0.94]	0.004	0.909	19
	Block-Q*TWS	**0.0002**	**0.00008**	**[0.00004, 0.0004]**	**0.087**	**0.014**	20
	Block-Q*FMT	−0.01	0.03	[−0.07, 0.05]	0.013	0.724	21
P300	**Block-L**	**−0.38**	**0.05**	**[−0.47, −0.29]**	**0.755**	**< 0.001**	22

Multi-level modeling results on time-window size (TWS), frontal midline theta (FMT) power, reward positivity (RewP) amplitude, and P300 amplitude with time-on-task. Line indicates the row number in which each effect is reported (for the purpose of indexing in the text). Block-L = block (time) coded as a linear slope from 0 to 15. Block-Q = block (time) coded as a quadratic slope (as the power of linear slope) from 0^2 to 15^2. Significant results (p < 0.05) are highlighted in bold.
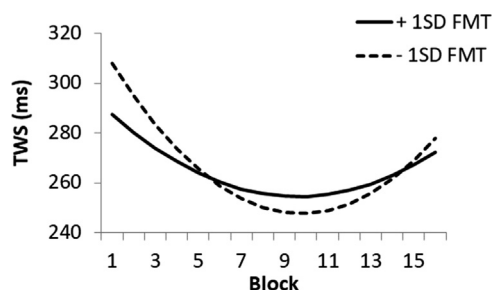

Fig. 2. The within-subjects effect depicting the interaction of frontal midline theta (FMT) power with linear-time (block) and quadratic-time on time-window size (TWS). The solid and dotted lines denote TWS across blocks when participants produced FMT power that was above or below one standard deviation relative to their own average FMT, respectively (+/− 1 SD FMT). Note that this figure is derived from the model predicting TWS based on time and FMT power (i.e., excluding reward positivity amplitude).

FMT power appear to be independent measures of task performance.

5.2. Electrophysiology

5.2.1. Frontal midline theta

Fig. 1b depicts the time-course of FMT power over time averaged across participants. As the inclusion of a quadratic time variable in the model did not reveal a significant effect, the multi-level modeling analysis here tested a model excluding this variable. The model predicting FMT power with time-window size, RewP amplitude, and linear-time revealed a significant main effect of linear-time, such that FMT power increased over time (Table 1, line 9). There was also a significant interaction of linear-time and time-window size on FMT power (Table 1, line 12; see Supplementary Fig. 1).[9] The pattern of interaction indicates that better task performance was associated with larger FMT power at the beginning (as shown by the significant main effect of time-window size; Table 1, line 10) and less of an increase in
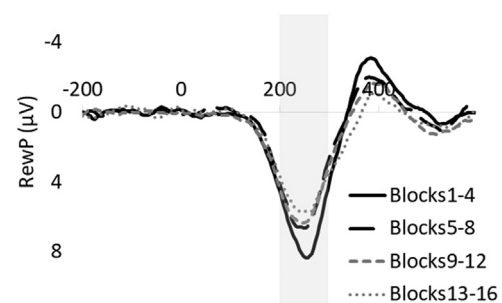
Fig. 3. The reward positivity (RewP) derived as a difference wave (reward feedback minus no-reward feedback) to event-related potentials averaged across participants and across blocks, separately for every consecutive four blocks. Blocks 1–4 = RewP difference wave for block 1–4 (i.e., 1st 30 min). Blocks 5–8 = RewP difference wave for block 5–8 (2nd 30 min). Blocks 9–12 = RewP difference wave for block 9–12 (3rd 30 min). Blocks 13–16 = RewP difference wave for block 13–16 (4th 30 min). RewP amplitude was measured between 200 and 300 ms following feedback onset at time 0 (as highlighted in gray). Negative is plotted up by convention.

FMT power with time on task (Supplementary Fig. 1).

5.2.2. RewP

Fig. 1c depicts the time-course of RewP amplitude averaged across participants, and Fig. 3 depicts the RewP waveforms averaged across participants for every half-hour of task performance. The multi-level modeling analysis predicting RewP amplitude with time-window size, FMT power, linear-time, and quadratic-time revealed statistically significant linear (Table 1, line 14) and quadratic effects (Table 1, line 15) of time on RewP amplitude. This result indicates a nonlinear reduction in RewP amplitude (more negative) with time (Fig. 1c). There were also significant interactions of linear-time and time-window size (Table 1, line 18)[10] and of quadratic-time and time-window size (Table 1, line

---

[9] Note that all the results remained unchanged when we performed the same analysis while retaining the data of the outliers (that were removed based on the 3SD criterion).

[10] Note that when we performed the same analysis while retaining the data of the outliers (that were removed based on the 3 SD criterion), we no longer observed reliable quadratic-time effects (the model with the quadratic-time variable failed to run when outliers were included) nor the interaction effect with linear-time and time-window size

20) on RewP amplitude, such that smaller (better) time-window size (relative to each participant's own average) was associated with a relatively smaller, linear decline in RewP amplitude (Supplementary Fig. 3).

5.2.3. P300

Fig. 1d depicts the time-course of P300 amplitude over time averaged across participants, and Supplementary Fig. 4 depicts the P300 waveforms averaged across participants for every half-hour of task performance, separately for the reward and no-reward trials. As the inclusion of a quadratic time variable in the model did not reveal a significant effect, the multi-level modeling analysis here excluded this variable. A multi-level modeling analysis predicting P300 amplitude with linear-time revealed a significant main effect of linear-time (Table 1, line 22), indicating that P300 amplitude decreased linearly with time on task.

5.3. Relationship between behavioral and electrophysiological measures

The multi-level modeling analysis examining the between-subjects relationships among time-window size, RewP, and FMT power did not reveal any statistically significant relationships (p > 0.40). This indicates that the overall sizes of FMT power and RewP amplitude were not related to task performance nor to each other. Note that because the statistical power for the between-participants analyses is low, these results should be interpreted with caution (as is also the case for the analyses related to personality, below).

5.4. Personality questionnaires

A summary of each questionnaire score is provided in Table 2, and zero-order correlations among questionnaires are provided in Table 3. We conducted a simple correlation analysis among time-window size, RewP, FMT, personality scores, and the task engagement level assessed by the post-experiment paper-and-pencil questionnaires (see Task Procedure).[11] Higher levels of task engagement were correlated with higher reward responsiveness scores (Pearson r = 0.310, p = 0.03), higher persistence scores (r = 0.316, p = 0.027), as well as better (smaller) time-window size (r = −0.381, p = 0.007). Exploratory analyses examining the impact of personality traits on time-window size, RewP amplitude, and FMT power did not reveal any significant effects of personality on these measures (see Supplementary Table 1).

6. Discussion

Goal-directed behaviors often require the deployment of cognitive control for their successful execution. Yet, because cognitive control is commonly perceived to be effortful, people tend to avoid applying it. Theories of control have therefore proposed the existence of a neurocognitive mechanism that regulates control levels by weighing the benefits (or rewards) of control against its associated effort-related costs (Botvinick and Braver, 2015; Holroyd and Yeung, 2012; Holroyd and McClure, 2015; Kurzban et al., 2013; Inzlicht et al., 2014; Shenhav et al., 2013, 2017; Westbrook and Braver, 2015). How this occurs is not fully understood, but ACC is believed to be partly responsible for it

---

(*footnote continued*)
(r = 0.01 including outliers, and r = 0.09 excluding outliers). These interaction effects are therefore less robust when the data of the participants who produced extreme scores were included in the analyses. Similar to the model predicting time-window size (see footnote #7), this is not surprising given that the effect sizes for the interaction terms were small to being with, and the inclusion of extreme scores increased the variability to explain, which necessarily reduced the effect sizes.

[11] Four of the participants' engagement levels were not obtained due to an error, hence the number of participants included in this analysis was 49.

Table 2
Summary of participant questionnaire scores. Temporal Experience of Pleasure-C = consummatory pleasure subscale of the temporal experience of pleasure scale. Temporal Experience of Pleasure-A = anticipatory pleasure subscale of the temporal experience of pleasure scale. Apathy is based on the apathy evaluation scale. Depression is based on the depression subscale of the depression, anxiety, and stress scale (DASS-21).

	Mean	SD	Range	Chronbach's α
Reward Responsiveness	26.4	2.57	19–30	0.71
Temporal Experience of Pleasure-C	37.3	4.44	26–46	0.46
Temporal Experience of Pleasure-A	44.62	6.24	32–60	0.74
Persistence	67.91	10.49	42–87	0.9
Apathy	31.83	5.41	23–43	0.77
Depression	3.96	4.34	0–18	0.89

Table 3
Zero-order correlations among questionnaire scores.

	1	2	3	4	5
1. Reward Responsiveness					
2. Temporal Experience of Pleasure-C	.03				
3. Temporal Experience of Pleasure-A	.44**	.20			
4. Persistence	.49**	−0.04	.26		
5. Apathy	−0.62**	−0.17	−0.54**	−0.58**	
6. Depression	−0.33*	.04	−0.28*	−0.57**	0.40**

* p < 0.05.
** p < 0.01.

(e.g., Holroyd and Umemoto, 2016; Shenhav et al., 2013, 2017; Vassena et al., 2017; Verguts et al., 2015). Here we explored the role of ACC in valuating and regulating control levels in order to sustain performance on an extended task, as revealed by electrophysiological indices of ACC activity (Holroyd and Umemoto, 2016).

Participants performed a standard time estimation task for 2 h while their EEG was recorded (Miltner et al., 1997). They received a small monetary reward for each "on-time" estimate, the difficulty of which was adjusted across trials so that all of the participants received the rewards on approximately 50% of the trials. Moreover, in order to increase their motivation the participants were told, truthfully, that better performance increased their chances for receiving one of two $100 Amazon gift cards, which in fact were later awarded to two of the participants via a lottery. We observed that time-window size decreased over time, following both a linear and a quadratic trend (Fig. 1a); that is, performance initially improved with time-on-task and later deteriorated towards the end of the experiment. A straightforward interpretation of these findings is that the initial improvement in task performance reflects an early learning process, and that the later decrement in performance reflects an impairment due to mental fatigue. Several observations support this interpretation. First, post-experiment self-reports revealed that better (smaller) time-window size correlated with overall task engagement, suggesting that task performance depended strongly on the control levels applied; thus, a withdrawal of control should result in impaired performance. Second, the amplitude of the P300, which is said to index attentional resource allocation to the task at hand (Polich, 2007), decreased with time on task (Fig. 1c & d). This gradual decline in P300 amplitude over time replicates a study by Hopstaken and colleagues (2015; 2016), who reported a time-on-task decrease in P300 amplitude that was accompanied by subjective measures of decreasing task engagement and increasing mental fatigue. Third, the amplitude of the RewP– which is derived as a difference wave in order to control for overlapping ERP components like the P300 – also declined with time on task. The strong association between RewP amplitude and reward sensitivity (Holroyd and Umemoto, 2016) suggests that the participants valued the

performance feedback less as the task progressed.

By contrast, FMT power increased with time on task (Fig. 1b), as observed previously (Barwick et al., 2012; Paus et al., 1997; Wascher et al., 2014). These findings suggest that, as predicted, reward valuation and attention levels decreased while control levels increased with time-on-task. Taken together, they appear to indicate that greater application of control in order to counteract deteriorating levels of attention and reward valuation is associated with accumulating mental fatigue, as suggested by a deterioration in performance at the end of the task (see also Hopstaken et al., 2015; c.f. Milyavskaya et al., this issue).

We also predicted that RewP amplitude and FMT power should interact to support task performance. However, the between-person analyses did not reveal any significant associations among time-window size, RewP amplitude, and FMT power, even though both of the electrophysiological phenomena are believed to be produced by ACC (Holroyd and Umemoto, 2016). Although low statistical power may have obscured these relationships, it is also possible that reward valuation and effort expenditure provide semi-independent determinants of task performance, such that RewP amplitude and FMT power would not be expected to co-vary across participants (e.g., Mueller et al., 2015). For this reason, we conducted within-subject analyses to gain insight into these factors. In particular, we applied a multi-level modeling approach, which provides a statistically powerful and flexible means to analyze how variations in within-subject measures interrelate over time (Hopstaken et al., 2015; Saunders et al., 2015).

Our primary analysis of interest explored whether FMT power and RewP amplitude, measured with respect to their within-subjects mean values, predicted subject performance (time-window size) across the task. This analysis revealed significant linear and quadratic effects on task performance of time-on-task that were modulated by FMT power (Fig. 2): participants performed the task better (i.e., time-window size was smaller) at the start of the task when they produced more FMT power relative to their average FMT power across the task, as compared to when they produced less FMT power relative to their average FMT power across the task. And their performance was more stable later in the task (exhibiting relatively shallow changes with time) when they produced greater FMT power relative to their average FMT power across the task as compared to when they produced less FMT power relative to their average FMT power across the task. These findings provide insight into the role of control in regulating task performance: Although increased FMT power is associated with deteriorating performance at the end of the task, as discussed above (see Fig. 1a and b), the within-subjects analysis reveals that greater FMT power is also associated with better performance at the start of the task (Fig. 2). Greater FMT power is further associated with less steep of a fall in performance toward the end of the task. Thus, FMT power appears to reflect a neural process that enhances performance, both during an early stage when subjects learn to perform the task better, and a later stage when performance suffers due to cognitive fatigue.

By contrast, contrary to our prediction, RewP amplitude did not predict task performance (Table 1, lines 4, 6, and 8). Further, several exploratory analyses also failed to reveal any significant interactions between RewP amplitude and FMT power predicting time-window size, or a mediating effect of RewP amplitude on FMT power to predict time-window size. Several possibilities might account for this. First, the variation in the subjective reward value of the feedback may have been insufficient to permit significant changes in performance, despite the lengthy and tedious nature of the task. In particular, the objective reward value was held constant across the task, which may have constrained the participants' subjective appraisals of the feedback. Second, the task design may have been suboptimal for detecting the interrelationships between reward valuation, effortful control and task performance. For example, task difficulty in this study was titrated so that participants received both reward and no-reward feedback approximately 50% of the time. This may have restricted the variability of the control signal, thereby weakening the potential RewP-FMT

relationship. Third, the interplay between reward value and task performance may be more complicated than the simple linear and curvilinear models that we tested here, especially given presumed individual variation in these processes across subjects. Consistent with this, we found that better task performers scored higher on trait measures of reward responsiveness and persistence. It may be that the performance levels of a subset of participants, in particular those scoring high in persistence, are less likely to be influenced by individual rewards received during the task (as this group may be more concerned about sustaining effort to complete the task successfully rather than accumulating rewards along the way). Although our study did not have sufficient statistical power to examine such individual differences, this possibility could be investigated in the future.

In addition to testing whether RewP amplitude and FMT power predict extended performance on an effortful task, we also explored factors that predict RewP amplitude and FMT power. Toward this end, we conducted within-person multi-level modeling analyses on these measures (Table 1). The analysis predicting RewP amplitude revealed significant interactions of time-window size with both linear- and quadratic-time trends. That is, smaller (better) time-window size (relative to each participant's own average time-window size across the task) was associated with smaller declines in RewP amplitude over time, whereas larger (worse) time-window size was associated with a steeper decline in RewP amplitude with time on task (Supplementary Fig. 3). Similarly, the within-person analysis predicting FMT revealed a significant interaction of linear-time and time-window size: Better task performance was associated with larger FMT power at the beginning of the task and smaller increases in FMT power with time-on-task (Supplementary Fig. 1). By contrast, we did not observe any associations between RewP amplitude and FMT power, suggesting their independence (Holroyd and Umemoto, 2016; Holroyd et al., 2012). It appears that greater task engagement, as reflected in smaller time-window size, was associated with more gradual reductions in RewP amplitude and more gradual increases in FMT power. That is, good task performance was associated with more stable indicators of reward valuation and cognitive control throughout the task.

Finally, we found from post-experiment self-reports that higher levels of task engagement correlated significantly with higher reward responsiveness and persistence scores, indicating that participants who scored high on these personality traits were more engaged during the entire experiment. An exploration of whether persistence and reward sensitivity were associated with task performance, RewP amplitude, and FMT power failed to reveal any significant relations among them (see Supplementary Table 1). Although a number of studies have reported individual differences in personality associated with RewP amplitude (e.g., Bress and Hajcak, 2013, Cooper et al., 2014; Liu et al., 2014; Parvaz et al., 2016; Umemoto and Holroyd, 2017) and phasic FMT power (e.g., Cavanagh and Shackman, 2014), we did not replicate any of these findings. Nevertheless, such effects may have been obscured by the low statistical power in the present study for conducting analyses related to individual differences.

A challenging question is whether increasing FMT power with time-on-task reflects growing mental fatigue per se, or the application of effortful control needed to overcome that fatigue, as these processes are necessarily correlated. In fact, several neuroimaging studies have implicated ACC in cognitive fatigue (e.g., Cook et al., 2007; Dobryakova et al., 2013; Wylie et al., 2017). However, we argue against the simple association between FMT and fatigue itself. Specifically, mental fatigue is generally associated with impaired task performance (as reflected by increased error rates and greater reaction times), but performance here improved over the first 90 min of the study while FMT power continued to rise. Furthermore, during the initial stage of the task greater FMT was associated, within subjects, with better task performance. Our observation that changes in FMT power closely followed time-on-task performance, as well as relatively more stable (shallower changes in) task performance, appears inconsistent with the account that FMT

simply reflects mental fatigue itself.

Past studies have shown that a brief, phasic burst of FMT power is commonly elicited by the delivery of external negative feedback stimuli (e.g., loss of money, punishment; see e.g., Bernat et al., 2015; Foti et al., 2015) and by response conflict and error commission (e.g., Cavanagh and Frank, 2014). Error commission has been studied in the literature examining mental fatigue, where the amplitude of the error related negativity (ERN) – a performance-related ERP component associated with both the RewP (Holroyd and Coles, 2002) and FMT power (Cavanagh and Frank, 2014) – has been observed to decline with time on task (Lorist et al., 2005; Boksem et al., 2006; Inzlicht and Gutsell, 2007). By contrast, ongoing FMT power, as we measured in this study, is associated with prolonged mental challenges and high working memory load (Hsieh and Ranganath, 2014; Mitchell et al., 2008). Understanding how phasic and ongoing FMT signals relate to each other is an important question for future studies. For example, past studies have shown that the adverse effects of mental fatigue on task performance and ERN amplitude, which is associated with phasic FMT (Cavanagh and Frank, 2014), can be counteracted with motivational incentives (e.g., Boksem et al., 2006; Hopstaken et al., 2015; Sarter et al., 2006). Whether fatigue-related changes in tonic FMT power can also be reversed with task incentives is an interesting avenue for future research.

In summary, although task performance improved until the last quarter of the experiment, when it started to reverse, decreasing RewP and P300 amplitudes indicated that the participants gradually disengaged from the task and devalued the outcomes of their performance. By contrast, FMT power increased monotonically with time on task, suggesting greater recruitment of cognitive control even as the task became harder and less rewarding. Although task performance and the electrophysiological measures did not correlate with each other across participants, a within-person increase in FMT power was associated with 1) better task performance at the start of the experiment, suggesting that high control levels facilitated learning; and 2) more stable performance overall. Conversely, better task performance was associated with higher self-reports of task engagement and relatively more stable RewP amplitudes and FMT power over time. Together, we interpret these results as reflecting the role of ACC in sustaining behavior over an extended period, especially for tasks that demand high levels of cognitive control and low immediate reward value.

Acknowledgements

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.neuropsychologia.2018.06.010.

References

Aiken, L.S., West, S.G., 1991. Multiple Regression: Testing and Interpreting Interactions. Sage, Newbury Park, CA.

Asada, H., Fukuda, Y., Tsunoda, S., Yamaguchi, M., Tonoike, M., 1999. Frontal midline theta rhythms reflect alternative activation of prefrontal cortex and anterior cingulate cortex in humans. Neurosci. Lett. 274, 29–32. http://dx.doi.org/10.1016/S0304-3940(99)00679-5.

Barwick, F., Arnett, P., Slobounov, S., 2012. EEG correlates of fatigue during administration of a neuropsychological test battery. Clin. Neurophysiol. 123, 278–284. http://dx.doi.org/10.1016/j.clinph.2011.06.027.

Becker, M.P.I., Nitsch, A.M., Miltner, W.H.R., Straube, T., 2014. A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. J. Neurosci. 34, 3005–3012. http://dx.doi.org/10.1523/JNEUROSCI.3684-13.2014.

Bernat, E.M., Nelson, L.D., Baskin-Sommers, A.R., 2015. Time-frequency theta and delta measures index separable components of feedback processing in a gambling task. Psychophys 52, 626–637.

Boksem, M.A.S., Meijman, T.F., Lorist, M.M., 2006. Mental fatigue, motivation and action monitoring. Biol. Psychol. 72, 123–132. http://dx.doi.org/10.1016/j.biopsycho.2005.08.007.

Boksem, M.A.S., Tops, M., 2008. Mental fatigue: costs and benefits. Brain Res. Rev. 59, 125–139. http://dx.doi.org/10.1016/j.brainresrev.2008.07.001.

Botvinick, M., Braver, T., 2015. Motivation and cognitive control: from behavior to neural mechanism. Annu. Rev. Psychol. 66, 83–113. http://dx.doi.org/10.1146/annurev-psych-010814-015044.

Botvinick, M.M., 2012. Hierarchical reinforcement learning and decision making. Curr. Opin. Neurobiol. 22, 956–962. http://dx.doi.org/10.1016/j.conb.2012.05.008.

Botvinick, M.M., Niv, Y., Barto, A.C., 2009. Hierarchically organized behavior and its neural foundations: a reinforcement learning perspective. Cognition 113, 262–280. http://dx.doi.org/10.1016/j.cognition.2008.08.011.

Brainard, D.H., 1997. The psychophysics toolbox. Spat. Vis. 10, 433–436. http://dx.doi.org/10.1163/156856897X00357.

Bress, J.N., Hajcak, G., 2013. Self-report and behavioral measures of reward sensitivity predict the feedback negativity. Psychophysiology 50, 610–616. http://dx.doi.org/10.1111/psyp.12053.

Cavanagh, J.F., Frank, M.J., 2014. Frontal theta as a mechanism for cognitive control. Trends Cogn. Sci. 18, 414–421. http://dx.doi.org/10.1016/j.tics.2014.04.012.

Cavanagh, J.F., Shackman, A.J., 2014. Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. J. Physiol. Paris 109, 3–15. http://dx.doi.org/10.1016/j.jphysparis.2014.04.003.

Cloninger, C.R., Svrakic, D.M., Przybeck, T.R., 1993. A psychobiological model of temperament and character. Arch. Gen. Psychiatry 50, 975–990.

Cook, D.B., O'Connor, P.J., Lange, G., Steffener, J., 2007. Functional neuroimaging correlates of mental fatigue induced by cognition among chronic fatigue syndrome patients and controls. Neuroimage 36, 108–122. http://dx.doi.org/10.1016/j.neuroimage.2007.02.033.

Cooper, A.J., Duke, E., Pickering, A.D., Smillie, L.D., 2014. Individual differences in reward prediction error: contrasting relations between feedback-related negativity and trait measures of reward sensitivity, impulsivity and extraversion. Front. Hum. Neurosci. 8, 248. http://dx.doi.org/10.3389/fnhum.2014.00248.

Dobryakova, E., DeLuca, J., Genova, H.M., Wylie, G.R., 2013. Neural correlates of cognitive fatigue: cortico-striatal circuitry and effort–reward imbalance. J. Int. Neuropsychol. Soc. 19, 849–853. http://dx.doi.org/10.1017/S1355617713000684.

Donchin, E., Coles, M.G.H., 1988. Is the P300 component a manifestation of context updating? Behav. Brain Sci. 11, 357–374. http://dx.doi.org/10.1017/S0140525X00058015.

Ebitz, R.B., Hayden, B.Y., 2016. Dorsal anterior cingulate: a Rorschach test for cognitive neuroscience. Nat. Neurosci. 19, 1278–1279. http://dx.doi.org/10.1038/nn.4387.

Erdfelder, E., Faul, F., Buchner, A., 1996. GPOWER: a general power analysis program. Behav. Res. Methods, Instrum. Comput. 28, 1–11. http://dx.doi.org/10.3758/BF03203630.

Foti, D., Weinberg, A., Bernat, E.M., Proudfit, G.H., 2015. Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. Clin. Neurophys. 126, 1338–1347.

Gard, D.E., Gard, M.G., Kring, A.M., John, O.P., 2006. Anticipatory and consummatory components of the experience of pleasure: a scale development study. J. Res. Pers. 40, 1086–1102. http://dx.doi.org/10.1016/j.jrp.2005.11.001.

Gratton, G., Coles, M.G.H., Donchin, E., 1983. A new method for off-line removal of ocular artifact. Electroencephalogr. Clin. Neurophysiol. 55, 468–484. http://dx.doi.org/10.1016/0013-4694(83)90135-9.

Gusnard, D. a., Ollinger, J.M., Shulman, G.L., Cloninger, C.R., Price, J.L., Van Essen, D.C., Raichle, M.E., 2003. Persistence and brain circuitry. Proc. Natl. Acad. Sci. USA 100, 3479–3484. http://dx.doi.org/10.1073/pnas.0538050100.

Hockey, G.R.J., 2011. A motivational control theory of cognitive fatigue. Cogn. Fatigue Multidiscip. Perspect. Curr. Res. Futur. Appl. 167–187. http://dx.doi.org/10.1037/12343-008.

Holroyd, C.B., Coles, M.G.H., 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. 109, 679–709. http://dx.doi.org/10.1037//0033-295X.109.4.679.

Holroyd, C.B., HajiHosseini, A., Baker, T.E., 2012. Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. Trends Cogn. Sci. 16, 192. http://dx.doi.org/10.1093/cercor/bhr332.

Holroyd, C.B., Krigolson, O.E., 2007. Reward prediction error signals associated with a modified time estimation task. Psychophysiology 44, 913–917. http://dx.doi.org/10.1111/j.1469-8986.2007.00561.x.

Holroyd, C.B., McClure, S.M., 2015. Hierarchical control over effortful behavior by rodent medial frontal cortex: a computational model. Psychol. Rev. 122, 54–83.

Holroyd, C.B., Umemoto, A., 2016. The research domain criteria framework: the case for anterior cingulate cortex. Neurosci. Biobehav. Rev. 71, 418–443. http://dx.doi.org/10.1016/j.neubiorev.2016.09.021.

Holroyd, C.B., Yeung, N., 2012. Motivation of extended behaviors by anterior cingulate cortex. Trends Cogn. Sci. 16, 122–128. http://dx.doi.org/10.1016/j.tics.2011.12.008.

Hopstaken, J.F., van der Linden, D., Bakker, A.B., Kompier, M.A.J., 2015. A multifaceted investigation of the link between mental fatigue and task disengagement. Psychophysiology 52, 305–315. http://dx.doi.org/10.1111/psyp.12339.

Hopstaken, J.F., van der Linden, D., Bakker, A.B., Kompier, M.A., Leung, Y.K., 2016. Shifts in attention during mental fatigue: evidence from subjective, behavioral, physiological, and eye-tracking data. J. Exp. Psychol.: Hum. Percept. Perform. 42 (6), 878.

Hsieh, L.T., Ranganath, C., 2014. Frontal midline theta oscillations during working

memory maintenance and episodic encoding and retrieval. Neuroimage 85, 721–729. http://dx.doi.org/10.1016/j.neuroimage.2013.08.003.

Ishii, R., Shinosaki, K., Ukai, S., Inouye, T., Ishihara, T., Yoshimine, T., Hirabuki, N., Asada, H., Kihara, T., Robinson, S.E., Takeda, M., 1999. Medial prefrontal cortex generates frontal midline theta rhythm. Neuroreport 10, 675–679. http://dx.doi.org/10.1097/00001756-199903170-00003.

Inzlicht, M., Bartholow, B.D., Hirsh, J.B., 2015. Emotional foundations of cognitive control. Trends Cogn. Sci. 19, 126–132. http://dx.doi.org/10.1016/j.tics.2015.01.004.

Inzlicht, M., Gutsell, J.N., 2007. Running on empty: neural signals for self-control failure. Psychol. Sci. 18, 933–937. http://dx.doi.org/10.1111/j.1467-9280.2007.02004.x.

Inzlicht, M., Schmeichel, B.J., Macrae, C.N., 2014. Why self-control seems (but may not be) limited. Trends Cogn. Sci. 18, 127–133. http://dx.doi.org/10.1016/j.tics.2013.12.009.

Janssen, D.J.C., Poljac, E., Bekkering, H., 2016. Binary sensitivity of theta activity for gain and loss when monitoring parametric prediction errors. Soc. Cogn. Affect. Neurosci. 11, 1280–1289. http://dx.doi.org/10.1093/scan/nsw033.

Jasper, H.H., 1958. The ten twenty electrode system of the international federation. Electroencephalogr. Clin. Neurophysiol. 10, 371–375.

Kool, W., McGuire, J.T., Rosen, Z.B., Botvinick, M.M., 2010. Decision making and the avoidance of cognitive demand. J. Exp. Psychol. Gen. 139, 665–682. http://dx.doi.org/10.1037/a0020198.

Kurzban, R., Duckworth, A., Kable, J.W., Myers, J., 2013. An opportunity cost model of subjective effort and task performance. Behav. Brain Sci. 36, 661–679. http://dx.doi.org/10.1017/S0140525X12003196.

Lefcheck, J.S., 2016. piecewiseSEM: piecewise structural equation modelling in r for ecology, evolution, and systematics. Methods Ecol. Evol. 7, 573–579.

Lin, H., Saunders, B., Hutcherson, C.A., Inzlicht, M., 2017. August. Midfrontal theta and pupil dilation track subjective conflict in value-based decisions. Poster session presented at the 13th International Conference for Cognitice Neuroscience, Amsterdam, Netherlands.

Liu, W. hua, Wang, L. zhi, Shang, H. rui, Shen, Y., Li, Z., Cheung, E.F.C., Chan, R.C.K., 2014. The influence of anhedonia on feedback negativity in major depressive disorder. Neuropsychologia 53, 213–220. http://dx.doi.org/10.1016/j.neuropsychologia.2013.11.023.

Lorist, M.M., Boksem, M.A.S., Ridderinkhof, K.R., 2005. Impaired cognitive control and reduced cingulate activity during mental fatigue. Cogn. Brain Res. 24, 199–205. http://dx.doi.org/10.1016/j.cogbrainres.2005.01.018.

Lovibond, P.F., Lovibond, S.H., 1995. The structure of negative emotional states: comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behav. Res. Ther. 33, 335–343. http://dx.doi.org/10.1016/0005-7967(94)00075-U.

Luu, P., Tucker, D.M., 2001. Regulating action: alternating activation of midline frontal and motor cortical networks. Clin. Neurophysiol. 112, 1295–1306. http://dx.doi.org/10.1016/S1388-2457(01)00559-4.

Marin, R.S., Biedrzycki, R.C., Firinciogullari, S., 1991. Reliability and validity of the apathy evaluation scale. Psychiatry Res. 38, 143–162. http://dx.doi.org/10.1016/0165-1781(91)90040-V.

McGuire, J.T., Botvinick, M.M., 2010. Prefrontal cortex, cognitive control, and the registration of decision costs. Proc. Natl. Acad. Sci. 107, 7922–7926. http://dx.doi.org/10.1073/pnas.0910662107.

Milyavskaya, M., Inzlicht, M., Johnson, T., Larson, M., 2018. Reward sensitivity following boredom and cognitive effort: A high-powered neurophysiological investigation. Neuropsychologia (this issue).

Miltner, W.H.R., Braun, C.H., Coles, M.G.H., 1997. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a generic neural system for error detection. J. Cogn. Neurosci. 9, 788–798.

Mitchell, D.J., McNaughton, N., Flanagan, D., Kirk, I.J., 2008. Frontal-midline theta from the perspective of hippocampal "theta". Prog. Neurobiol. 86, 156–185. http://dx.doi.org/10.1016/j.pneurobio.2008.09.005.

Mueller, E.M., Pechtel, P., Cohen, A.L., Douglas, S.R., Pizzagalli, D.A., 2015. Potentiated processing of negative feedback in depression is attenuated by anhedonia. Depression. Anxiety 32 (4), 296–305.

Murphy, P.R., Robertson, I.H., Balsters, J.H., O'connell, R.G., 2011. Pupillometry and P3 index the locus coeruleus–noradrenergic arousal function in humans. Psychophys 48, 1532–1543.

Nelson, B.D., Kessel, E.M., Jackson, F., Hajcak, G., 2016. The impact of an unpredictable context and intolerance of uncertainty on the electrocortical response to monetary gains and losses. Cogn. Affect. Behav. Neurosci. 16, 153–163. http://dx.doi.org/10.3758/s13415-015-0382-3.

Norman, D.A., Shallice, T., 1986. Attention to Action. Conscious. Self-Regulation 4, 1–18. 〈http://dx.doi.org/10.1007/978-1-4757-0629-1_1〉.

Parvaz, M.A., Gabbay, V., Malaker, P., Goldstein, R.Z., 2016. Objective and specific tracking of anhedonia via event-related potentials in individuals with cocaine use disorders. Drug Alcohol Depend. 164, 158–165. http://dx.doi.org/10.1016/j.drugalcdep.2016.05.004.

Paus, T., Zatorre, R.J., Hofle, N., Caramanos, Z., Gotman, J., Petrides, M., Evans, A.C., 1997. Time-related changes in neural systems underlying attention and arousal during the performance of an auditory vigilance task. J. Cogn. Neurosci. 9, 392–408. http://dx.doi.org/10.1162/jocn.1997.9.3.392.

Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat. Vis. http://dx.doi.org/10.1163/156856897X00366.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118, 2128–2148.

Proudfit, G.H., 2015. The reward positivity: from basic research on reward to a biomarker for depression. Psychophysiology 52, 449–459. http://dx.doi.org/10.1111/psyp.12370.

Sambrook, T.D., Goslin, J., 2015. A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. Psychol. Bull. 141, 213–235. http://dx.doi.org/10.1037/bul0000006.

Sarter, M., Gehring, W.J., Kozak, R., 2006. More attention must be paid: the neurobiology of attentional effort. Brain Res. Rev. 51, 145–160. http://dx.doi.org/10.1016/j.brainresrev.2005.11.002.

Saunders, B., Milyavskaya, M., Inzlicht, M., 2015. What does cognitive control feel like? Effective and ineffective cognitive control is associated with divergent phenomenology. Psychophysiology 52, 1205–1217. http://dx.doi.org/10.1111/psyp.12454.

Scheeringa, R., Bastiaansen, M.C.M., Petersson, K.M., Oostenveld, R., Norris, D.G., Hagoort, P., 2008. Frontal theta EEG activity correlates negatively with the default mode network in resting state. Int. J. Psychophysiol. 67, 242–251. http://dx.doi.org/10.1016/j.ijpsycho.2007.05.017.

Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, Andrew S., Winter, J.J., Davidson, R.J., 2011. Nat. Rev. Neurosci. 12, 154–167. http://dx.doi.org/10.1038/nrn2994.

Shenhav, A., Botvinick, M.M., Cohen, J.D., 2013. The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron 79, 217–240. http://dx.doi.org/10.1016/j.neuron.2013.07.007.

Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T.L., Cohen, J.D., Botvinick, M.M., 2017. Toward a rational and mechanistic account of mental effort. Annu. Rev. Neurosci. 40, 99–124. http://dx.doi.org/10.1146/annurev-neuro-072116-031526.

Treynor, W., Gonzalez, R., Nolen-Hoeksema, S., 2003. Rumination reconsidered: a psychometric analysis. Cogn. Ther. Res. 27, 247–259.

Umemoto, A., Holroyd, C.B., 2016. Exploring individual differences in task switching: persistence and other personality traits related to anterior cingulate cortex function. Prog. Brain Res. 229, 189–212.

Umemoto, A., Holroyd, C.B., 2017. Neural mechanisms of reward processing associated with depression-related personality traits. Clin. Neurophysiol. 128, 1184–1196. http://dx.doi.org/10.1016/j.clinph.2017.03.049.

Van den Berg, I., Franken, I.H.A., Muris, P., 2010. A new scale for measuring reward responsiveness. Front. Psychol. 1, 1–7. http://dx.doi.org/10.3389/fpsyg.2010.00239.

Vassena, E., Holroyd, C.B., Alexander, W.H., 2017. Computational models of anterior cingulate cortex: At the crossroads between prediction and effort. Front. Neurosci. 11, 1–9. http://dx.doi.org/10.3389/fnins.2017.00316.

Verguts, T., Vassena, E., Silvetti, M., 2015. Adaptive effort investment in cognitive and physical tasks: a neurocomputational model. Front. Behav. Neurosci. 9. http://dx.doi.org/10.3389/fnbeh.2015.00057.

Vogt, B.A., 2009. Cingulate Neurobiology and Disease. Oxford University Press, Oxford, UK.

Walsh, M.M., Anderson, J.R., 2012. Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. Neurosci. Biobehav. Rev. 36, 1870–1884. http://dx.doi.org/10.1016/j.neubiorev.2012.05.008.

Wascher, E., Rasch, B., Sanger, J., Hoffmann, S., Schneider, D., Rinkenauer, G., Heuer, H., Gutberlet, I., 2014. Frontal theta activity reflects distinct aspects of mental fatigue. Biol. Psychol. 96, 57–65. http://dx.doi.org/10.1016/j.biopsycho.2013.11.010.

Westbrook, A., Braver, T.S., 2015. Cognitive effort: a neuroeconomic approach. Cogn. Affect. Behav. Neurosci. 15, 395–415. http://dx.doi.org/10.3758/s13415-015-0334-y.

Wylie, G.R., Genova, H.M., DeLuca, J., Dobryakova, E., 2017. The relationship between outcome prediction and cognitive fatigue: a convergence of paradigms. Cogn. Affect. Behav. Neurosci. http://dx.doi.org/10.3758/s13415-017-0515-y.