

## Empirical Research

# More Problems After Difficult Problems? Behavioral and Electrophysiological Evidence for Sequential Difficulty Effects in Mental Arithmetic

Jochen Andreas Mosbacher<sup>\*a</sup>, Clemens Brunner<sup>a</sup>, Roland Hugo Grabner<sup>ab</sup>**[a]** Section of Educational Neuroscience, Institute of Psychology, University of Graz, Graz, Austria. **[b]** BioTechMed-Graz, Graz, Austria.

## Abstract

This study investigated whether sequential difficulty effects emerge during processing of a mixed set of small, easy and large, more difficult arithmetic problems. Furthermore, we assessed if these sequential difficulty effects are reflected in event-related (de-)synchronization (ERS/ERD) patterns. To this end, we analyzed data of 65 participants, who solved two separate blocks (additions and subtractions) of arithmetic problems while their EEG was recorded. In each block, half of the problems were difficult problems (two-digit/two-digit with carry/borrow), and the other half were easy problems (one-digit/one-digit). Half of the problems were preceded by a problem of the same difficulty (repeat trials), and half were preceded by problems of the other difficulty (switch trials). In subtractions a sequential difficulty effects pattern emerged. Participants solved easy repeat trials faster than easy switch trials, while difficult repeat trials were solved slower and less accurately than difficult switch trials. In the EEG, we found the strongest effects in left hemispheric beta band (13–30 Hz) ERD. Specifically, participants showed a stronger beta band ERD in easy switch trials than in easy repeat trials. Furthermore, beta band ERD was stronger in difficult problems than in easy problems within repeat trials, but stronger in easy problems than in difficult problems within switch trials. In summary, our results are in line with the presence of sequential difficulty effects, as processing of easy and difficult problems was impaired if they were preceded by a difficult problem. Furthermore, these sequential difficulty effects are reflected in ERD patterns.

**Keywords:** EEG, event-related (de-)synchronization, mental arithmetic, sequential difficulty effects, procedural calculation, fact retrieval

Journal of Numerical Cognition, 2020, Vol. 6(1), 108–128, <https://doi.org/10.5964/jnc.v6i1.223>

Received: 2019-05-02. Accepted: 2020-02-06. Published (VoR): 2020-06-15.

\*Corresponding author at: Section of Educational Neuroscience, Institute of Psychology, University of Graz, Universitätsplatz 2, 8010 Graz, Austria. E-mail: [jochen.mosbacher@uni-graz.at](mailto:jochen.mosbacher@uni-graz.at)



This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International License, CC BY 4.0 (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sequential difficulty effects in arithmetic have been first described by [Schneider and Anderson \(2010\)](#). In two experiments using slight manipulations of difficulty in arithmetic problems (with versus without carry/borrow or vertical versus horizontal presentation), these authors could show that performance is worse in trials following on a difficult trial than in trials following on an easy trial. This was not only true for easy switch (easy trial following on a difficult trial) as compared to easy repeat trials (easy trial following on another easy trial) in which costs inflicted by the difficulty switch might also be a cause for worse performance. It was also found in difficult repeat trials (difficult trial following on another difficult trial) which were associated with worse performance as compared to difficult switch trials (difficult trial following on an easy trial). In other words, in difficult trials, participants performed worse in trials where no difficulty switch occurred. The authors' interpretation was that the processing of a difficult task might lead to some kind of resource depletion in working memory or executive control functions that affects the processing of the immediately following task, regardless of whether a switch

(e.g. in task, difficulty, or strategy) occurred or not (Schneider & Anderson, 2010). The more difficult a task or trial, the more resources would be depleted and the stronger would the following one be affected. This notion was later supported by studies on arithmetic performance using approximate multiplication problems with rounding strategies of different difficulty (rounding up versus rounding down; termed strategy sequential difficulty effects (Uittenhove & Lemaire, 2012, 2013; Uittenhove, Poletti, Dufau, & Lemaire, 2013).

An open question however is, whether these patterns of sequential difficulty effects also emerge in a common difficulty switching process in mental arithmetic that occurs in everyday life, namely the switch between fact retrieval and procedural calculation. In general, arithmetic problems can be solved by application of one of these strategies. Fact retrieval, being the mainly applied strategy in small, easy problems (e.g.  $3 + 5 = 8$  or  $2 * 4 = 8$ ), means solving a problem by directly retrieving the solution from long-term memory – a fast, effortless, and highly accurate process. Procedural calculation, on the other hand, occurs primarily in larger, difficult problems (e.g.  $35 + 17$ ). For instance, persons break the problem down in a series of smaller steps to calculate the correct solution ( $35 + 17$  could be broken down into  $35 + 10 = 45$ ;  $45 + 5 = 50$ ;  $7 - 5 = 2$ ;  $50 + 2 = 52$ ). This requires recalling the correct procedure, maintaining the interim results in working memory, and monitoring progress across the single steps. Thus, these two types of strategies strongly differ in complexity and difficulty, and the cognitive load put on executive control functions and working memory is a discriminating feature between them (e.g. Imbo & Vandierendonck, 2007). Furthermore, these two strategies differ starkly in oscillatory patterns assessed by EEG (e.g. De Smedt, Grabner, & Studer, 2009; Grabner & De Smedt, 2011; Tschentscher & Hauk, 2016). Fact retrieval processes are associated with a stronger left hemispheric theta band event-related synchronization (ERS, a power increase in this frequency band from prior to during the task), while a stronger event-related desynchronization (ERD; a power decrease from prior to during the task) in lower and upper alpha and beta bands has been linked to procedural calculations (De Smedt et al., 2009; Grabner & De Smedt, 2011; Tschentscher & Hauk, 2016). Theta band ERS during arithmetic fact retrieval has been interpreted as reflecting retrieval processes of arithmetic facts from verbal long-term memory. Higher ERD in the different alpha and beta bands during procedural calculations, in contrast, is thought to reflect the larger demand of attentional processes (Grabner & De Smedt, 2011) and executive control functions (Tschentscher & Hauk, 2016), respectively.

However, whether these ERS/ERD patterns related to fact retrieval and procedural calculation are also modulated by difficulty switches and consequently by sequential difficulty effects has not been investigated. A previous study on the neurophysiological correlates of sequential difficulty effects in mental arithmetic focused on event-related potentials (ERP; Uittenhove et al., 2013). Results showed slower response latencies and a more negative ERP mean amplitude 200 to 500 ms after trial onset in those trials that were preceded by a difficult trial as compared to those preceded by an easy trial. Uittenhove and colleagues interpreted this result as larger cerebral activity in trials following a more difficult one than an easier one and that sequential difficulty effects interfere with central parts of arithmetic processes (e.g. retrieval of procedures). The high temporal resolution, allowing for a fine-grained analysis of cognitive processes is thereby an advantage of ERPs. ERS/ERD patterns on the other hand allow for the analysis of difficulty switching and sequential difficulty effects on task related activity changes on an oscillatory level in different frequency bands which have been associated with either fact retrieval, or procedural calculation.

The present study aims to extend the current understanding of sequential difficulty effects in mental arithmetic in two ways: First, by investigating sequential difficulty effects in switching between fact retrieval and procedural

calculation, the study tests whether and to what extent previously observed sequential difficulty effects can be generalized to everyday arithmetic demands. And, second, by investigating their oscillatory EEG correlates, the potential impact of these effects on ERS/ERD patterns related to fact retrieval and procedural calculations can be uncovered. Notably, in previous oscillatory EEG studies on mental arithmetic, switching between these trial types is common but sequential difficulty effects on the ERS/ERD patterns have not been investigated so far.

To this end, participants solved two separate sets of arithmetic problems. One consisting of additions and one consisting of subtractions. In each set, half of the problems were easy, fact retrieval problems, and the other half were difficult, procedural calculation problems. Based on prior literature we expected better performance in easy repeat trials than in easy switch trials, but a better performance in difficult switch trials than in difficult repeat trials. In other words, the effects of difficulty switching should be asymmetric, consisting of performance impairments due to possible switching costs inflicted by the difference in difficulty plus sequential difficulty effects in the easy (fact retrieval) switch trials and performance impairments due to sequential difficulty effects in the difficult (procedural) repeat trials. At the electrophysiological level, we expected effects of sequential difficulty effects in the alpha and beta bands as these are related to attentional processes and executive functions (e.g. Grabner & De Smedt, 2011; Tschentscher & Hauk, 2016). If sequential difficulty effects indeed emerge because of resource depletion in these functions, ERD in alpha and beta bands should be stronger in trials following a difficult trial than in trials following an easy trial. Within the easy trials, there should be a stronger ERD in switch than in repeat trials, while within the difficult trials, ERD should be stronger in repeat than in switch trials. Additionally, we explored if the magnitude of sequential difficulty effects relates to basic arithmetic abilities as well as working memory (WM) functions.

## Method

### Sample

The sample consisted of 72 adult students. Seven participants were excluded because they showed an accuracy more than two standard deviations below the mean in at least one type of arithmetic problems. Hence, the final sample consisted of 65 participants with a mean age of 25.98 years ( $SD = 5.33$ ) of which 42 (64.6%) were female. Inclusion criteria consisted of (a) age between 18 and 45 years, (b) no prior or current neurological or psychiatric disease, (c) no regular intake of medication or drugs potentially altering brain states or consciousness (e.g. antidepressants, neuroleptics, alcohol), and (d) right-handedness. None reported any reading or mathematical disabilities. All participants volunteered and gave written and informed consent to take part in the study and received either 20€ or 3.5 hours of participant time (part of course requirement for psychology students) as compensation. The study and procedures were approved by the ethics committee of the University of Graz.

### Tests and Materials

#### Arithmetic Tasks

The arithmetic task (as well as the working memory task) was presented on a 24 inch screen (LG Electronics Inc., Seoul, South Korea) and controlled by PsychoPy software (Peirce, 2007; Peirce et al., 2019). The refresh rate of the screen was set to 120 Hz and participants sat in a comfortable chair with armrests. Responses were

assessed using an RB-series response pad (Cedrus Corporation, San Pedro, USA). Written instructions were given before the start of the first task, and participants could work through four practice trials. After the practice trials, participants had the opportunity to ask questions. After all questions were answered, the investigator left the room and the paradigm was started. Participants then worked on two blocks of arithmetic problems (one with additions and one with subtractions) consisting of 32 easy and 32 difficult problems each. Easy additions consisted of one-digit plus one-digit problems constructed with addends between 2 and 8 and sums below or equal to 10. Tie problems were excluded. As there are only 24 arithmetic problems fulfilling these conditions, eight randomly selected problems were repeated once. Difficult additions were two-digit plus two-digit problems with carry and addends between 12 and 59 and sums below 100. For each participant, the difficult problems were randomly selected from a set of 143 possible problems. Tie problems and problems containing a round number (e.g. 30) were excluded. Subtractions mirrored additions by consisting of the result of one of the additions minus one of its addends, and the difficult problems were again selected randomly. Within each block, the order of easy and difficult problems was pseudorandomized, so that half of the trials were switch trials and half repeat trials for each problem size. A switch trial was defined as a trial consisting of an easy or difficult problem that followed on a problem of the other kind. In contrast, a repeat trial was present if the problem of the current trial was of the same kind as the problem in the trial before. The first trial of each block was excluded from analyses. A single trial (Figure 1) started with a fixation cross for 1 second, followed by the arithmetic problem. The problem was presented on screen until the participant pressed a button, indicating that she or he had solved the problem, or time ran out (the maximal time available were three seconds for easy and five seconds for difficult problems). After button press or timeout, participants had three seconds to select the correct answer from three options, again by pressing a respective button. Before the next trial, an inter-trial interval followed that lasted for the remaining time not used for finding a solution and selecting an answer (min. = 0 seconds; max. = 6 seconds for easy or 8 seconds for difficult trials).

At the end of each arithmetic block, participants were asked separately for easy and difficult problems whether they used more fact retrieval or procedural calculations. Thereby, participants could indicate their strategy use by positioning a cursor on a bar ranging from “retrieved” (meaning 100% of the trials of this type were solved by fact retrieval and 0% by procedural calculations) to “calculated” (meaning 100% of the trials of this type were solved by procedural calculations and 0% by retrieval).

Accuracy and calculation times were used as performance markers. Accuracy was calculated as the percentage of correctly and timely solved trials. Calculation time was assessed as the mean time between start of problem presentation and the first button press, indicating that the participant had come to a solution. Calculation times above or below three standard deviations of a participant’s mean Calculation time in the given tasks were excluded from further analyses. On average, this led to an exclusion of 2.97% ( $SD = 1.75$ ) of the addition trials and 2.63% ( $SD = 1.72$ ) of the subtraction trials. Incorrect trials or trials not solved in time (no button press before calculation time was up) were excluded and, in order to rule out effects of post error slowing, also correct trials that followed an incorrect trial were not considered for analysis.

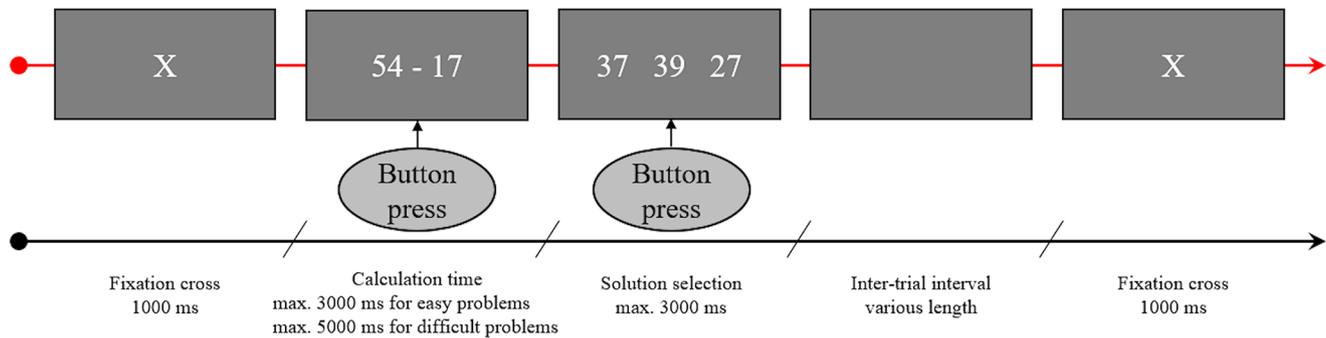


Figure 1. Single trial.

*Note.* Figure 1 shows the procedure of a single arithmetic task. At the beginning of each trial a fixation cross was presented in the middle of the screen for 1 second. Immediately afterwards, the problem was presented and stayed on the screen until the participant pressed a button, indicating that he or she had come to a solution, or calculation time had run out. Maximum calculation times were three seconds for easy problems and five seconds for difficult problems. After the participants pressed the button or the time had run out, they were presented with three possible answers and had three seconds time to indicate the correct one by pressing the corresponding button. Trials were separated by an inter-trial interval consisting of a blank screen of various length, whereby the length corresponded to the unused time of calculation and answer selection parts (duration range: zero to six or eight seconds, depending on difficulty).

### Additional Cognitive Tasks

**Working memory task** — To assess participants' WM, a 2-back letter task was used. The task had a duration of 180 seconds and each letter was presented for 0.5 seconds, followed by a blank screen for 1.5 seconds before the next letter appeared. Hence, the WM task consisted of 90 trials. The letters appeared in a pseudorandomized order to achieve a ratio of 30 target trials to 60 non-target trials. Participants had to press a button if the current letter matched the letter seen two trials ago (target trial) and to refrain from pressing any button if not. Hence, a correct reaction (CR) was present when a target item was displayed and the participant pressed the button and a correct rejection (CRJ) was present if a non-target item was presented and the participant refrained from pressing a button. A miss (M), on the other hand, was present if a target item was presented and the participant failed to press the button, while a false alarm (FA) occurred if a non-target item was presented and the participant pressed the button. Reaction time (WM-RT) was assessed for all correct trials and overall WM accuracy (WM-ACC) was calculated as the percentage of correct reactions and correct rejections over all trials by the equation  $WM-ACC = (1 - ((FA + M) / 90)) * 100$ . As for the arithmetic task, written instructions were given at the beginning and after ten practice trials, participants were given the opportunity to ask final questions.

**IST arithmetic score** — Basic arithmetic abilities were assessed by the IST-2000R (Liepmann, Beauducel, Brocke, & Amthauer, 2007) subtest *Rechenzeichen* (arithmetic operators), a pen & paper test. This subtest consists of 20 items in which participants are given two to three numbers and a result (in the form:  $A ? B = X$  or  $A ? B ? C = X$ ) and have 10 minutes to insert the correct operators in place of the question marks which result in a correct equation. Points were given for every correctly solved item.

### EEG

EEG was recorded using a BioSemi ActiveTwo EEG system (BioSemi, Amsterdam, Netherlands) with 64 channels and hardware low-pass filter of -3 dB at 1/5 of the sampling rate of 256 Hz. Electrodes were mounted in a BioSemi headcap and contact was established using Signagel (Parker Laboratories Inc., Fairfield, USA).

Channels F3, P3, AF8, and F8 were not connected because at these positions transcranial electrical stimulation electrodes were mounted beneath the head cap. Stimulation was not used during the assessments for this study, but followed immediately after as part of another study examining the same participants. EEG data was analyzed using MNE (Gramfort et al., 2013, 2014) and additional custom code in Python. Pre-processing was done in a semi-automatic way, by using an average reference with a high-pass filter at 1 Hz and a notch filter at 50 Hz and performing a visual inspection to remove prominent artefacts and bad/unconnected channels first, followed by an independent component analysis (ICA) to remove ocular artefacts and a second visual inspection to remove remaining artefacts from the signal. After interpolating excluded channels, data were analyzed separately for each frequency band of interest. These consisted of the theta (3-6 Hz), lower alpha (8-10 Hz), upper alpha (10-13 Hz), and beta (13-30 Hz) bands. Frequency ranges of the theta and alpha bands were based on prior studies (e.g. Grabner & De Smedt, 2011, 2012). ERS/ERD values were calculated as follows. First, the EEG was band-pass filtered in the respective band using the IIR forward-backward filtering option in MNE. Second, the mean power was calculated for the resting phase (R) as well as the active phase (A) in all correct trials that followed on a correct trial and where more than 50% of data points remained after artefact correction. The resting phase was defined as the 1000 ms from onset of the fixation cross to the onset of the arithmetic problem. The active phase consisted of the time between onset of the problem and the button press indicating a found solution by the participant. Hence, it spanned the calculation time and was of various length. Third, the mean power for R and A within each trial was averaged over all trials of each respective trial type (easy stay, easy switch, complex stay, and complex switch). Fourth, the ERS/ERD values were calculated for each trial type with  $ERS/ERD = ((A - R) / R) * 100$ , with a positive value indicating ERS (a power increase from rest to activation in the respective EEG band), while a negative value indicates ERD (a power decrease from rest to activation in the respective band). Finally, cluster values of ERS/ERD were calculated as averages over the single channel values of the electrodes contained in each cluster. These clusters (Figure 2) were selected based on prior work (e.g. Grabner & De Smedt, 2011, 2012) and consisted of: (a) left (containing electrodes Fp1, AF3, and AF7) and right (Fp2, AF4) anterior-frontal (AF) clusters; (b) left (F1, F5, F7) and right (F2, F4, F6) frontal (F) clusters; (c) left (FC1, FC3, FC5) and right (FC2, FC4, FC6) fronto-central (FC) clusters; (d) left (C1, C2, C3) and right (C2, C4, C6) central (C) clusters; (e) left (CP1, CP3, CP5) and right (CP2, CP4, CP6) centro-parietal (CP) clusters; (f) left (P1, P5, P7) and right (P2, P4, P6, P8) parietal (P) clusters; (g) left (PO3, PO7, O1) and right (PO4, PO8, O2) parieto-occipital (PO) clusters; and (h) an anterior (Fpz, AFz, Fz, FCz) and a posterior (CPz, Pz, POz, Oz) midline cluster (AM and PM respectively).

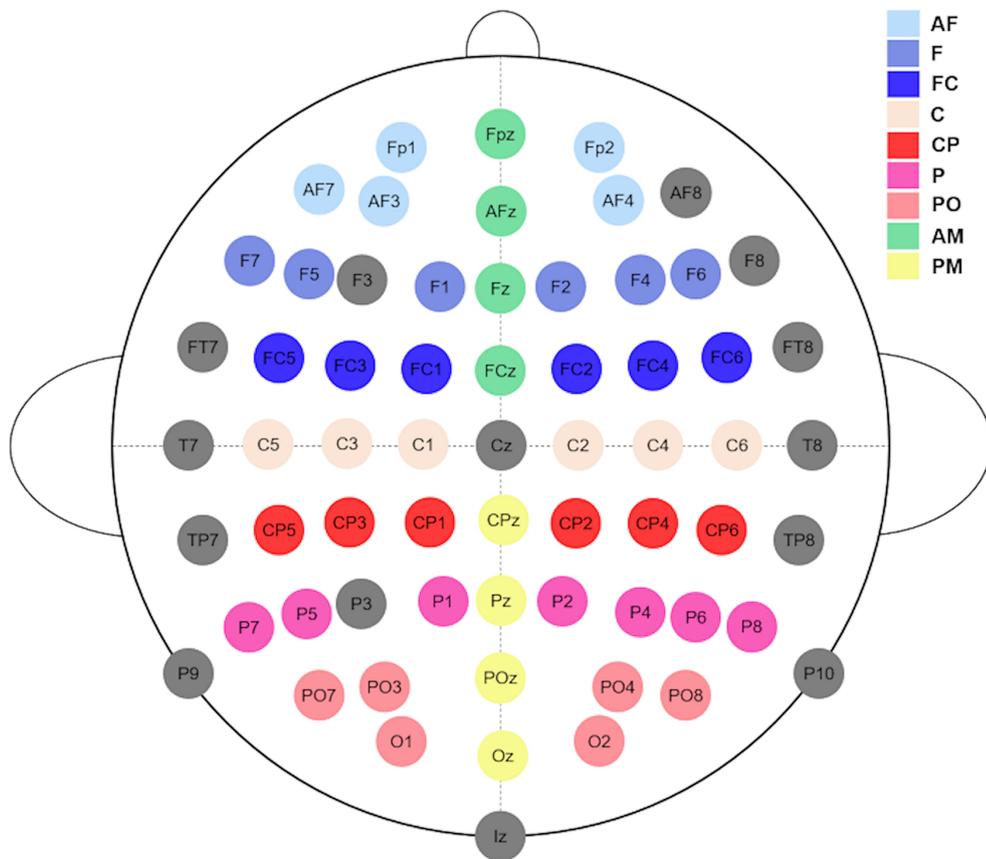


Figure 2. EEG clusters.

*Note.* Figure 2 depicts the EEG clusters used for the analysis. Electrodes were aggregated into seven cluster per hemisphere and two midline cluster. Cluster were based on prior work (e.g. Grabner & De Smedt, 2011), electrodes depicted in grey were not included in analysis.

## Procedure

Examination and EEG measurement took place in normally lit, quiet rooms, in a single person setting. Tests and questionnaires conducted before the EEG measurement consisted of a short demographic questionnaire, followed by a test of hand dominance (HDT; Steingrüber & Lienert, 1971), the *Regensburger Wortflüssigkeitstest* (RWT, a German verbal fluency test; Aschenbrenner, Tucha, & Lange, 2000), the Comprehensive-Trail-Making-Test (CTMT; Reynolds, 2002), and finally the subtest *Rechenzeichen* (arithmetic operators) of the IST 2000R. Immediately afterwards, montage of the EEG electrodes and the main test session followed. In the main test session, participants solve the two blocks of arithmetic problems (one block of additions and one of subtractions) and the WM-task. The order of the arithmetic blocks and the WM task was balanced so that all tasks appeared equally often in each position and separated by short breaks of one minute. Finally, after the EEG measurement was finished, participants were asked to complete the German version of the NEO-Five-Factor-Inventory (NEO-FFI; Borkenau & Ostendorf, 2008; Costa & McCrae, 1992) and a short German questionnaire based on Dweck's mindset assessment (Dweck, Chiu, & Hong, 1995). These two measures as well as RWT and CTMT were not analyzed in this study but are part of research conducted with the same sample.

## Calculations and Analyses

### Switching Costs / Sequential Difficulty Effect

In the arithmetic tasks, switching costs (and with them sequential difficulty effects) were calculated separately for easy and difficult problems as well as additions and subtractions by assessing the difference between switch and repeat trials. For both calculation times and accuracy, the amount of switching costs/sequential difficulty effects was calculated as the difference between performance in switch trials and performance in repeat trials (switch – repeat). Hence, for calculation time positive values indicate better performance in repeat trials and negative values indicate better performance in switch trials. For accuracy, on the other hand, positive values indicate better performance in switch trials, while negative values indicate better performance in repeat trials.

### Statistical Analyses

All statistical analyses were carried out using SPSS 25 (IBM Corporation, Armonk, USA). Behavioral data (calculation times and accuracy) were analyzed separately for additions and subtractions, using repeated measurement ANOVAs with difficulty (easy vs. difficult) and order (switch vs. repeat trials) as within-subject factors. ERS/ERD data was analyzed similarly with separate analyses for each frequency band using repeated measurement ANOVAs with difficulty (easy vs. difficult), order (switch vs. repeat trials), and location cluster as within-subject factors. Furthermore, analyses were conducted separately for left and right hemispheres as well as midline areas. Hence, the location cluster factor consisted of AF, F, FC, C, CP, P, and PO for the analyses regarding the left and right hemispheres and of AM and PM for the analysis of midline areas. Bivariate correlations were calculated to assess the associations between behavioral switching costs/sequential difficulty effects, WM, and IST arithmetic scores. For analyses including WM we had to remove one additional participant as response recording had not worked properly and accuracy and reaction times could not be assessed. Hence, sample size for analyses including WM is 64 instead of 65. The main research question was whether there are switching costs or sequential difficulty effects on a behavioral level when switching between fact retrieval (easy) and procedural calculation (difficult) and, if any, if they are reflected on a neurophysiological level. Hence, for the sake of brevity, ERS/ERD patterns and correlations were only analyzed for those arithmetic operations in which behavioral switching costs/sequential difficulty effects appeared. In all analyses applicable, Greenhouse-Geisser correction was applied if the assumption of sphericity was violated.

## Results

### Performance

Means and standard deviations for accuracy and calculation times are given in [Table 1](#). Additionally, switching costs / sequential difficulty effects in accuracy and calculation times are depicted in [Figure 3](#) to give a better overview.

Table 1

Means and Standard Deviations of Accuracy (%) and Calculation Times (Seconds) in the Different Conditions

Type	Difficulty	Order	<i>M</i>	<i>SD</i>
<b>Accuracy</b>				
Additions	Easy	Repeat	96.43	4.85
		Switch	95.67	4.81
	Difficult	Repeat	86.15	12.18
		Switch	84.58	9.77
Subtractions	Easy	Repeat	96.25	4.91
		Switch	95.87	5.21
	Difficult	Repeat	73.65	16.18
		Switch	78.15	13.95
<b>Calculation Times</b>				
Additions	Easy	Repeat	0.74	0.14
		Switch	0.74	0.13
	Difficult	Repeat	2.49	0.64
		Switch	2.50	0.70
Subtractions	Easy	Repeat	0.82	0.19
		Switch	0.85	0.20
	Difficult	Repeat	2.94	0.77
		Switch	2.83	0.68

### Additions

For accuracy, the repeated measurements ANOVA showed a significant main effect of difficulty,  $F(1, 64) = 82.855$ ;  $p < .001$ ;  $\eta_p^2 = .564$ , with easy problems being solved more accurately ( $M = 96.05$ ;  $SD = 3.28$ ) than difficult problems ( $M = 85.37$ ;  $SD = 8.97$ ). The main effect order,  $F(1, 64) = 1.685$ ;  $p = .199$ ;  $\eta_p^2 = .026$ , and the interaction difficulty \* order,  $F(1, 64) = 0.198$ ;  $p = .658$ ;  $\eta_p^2 = .003$ , were not significant.

Similarly, for calculation times the ANOVA only showed a significant main effect of difficulty,  $F(1, 64) = 576.910$ ;  $p < .001$ ;  $\eta_p^2 = .900$ , with easy problems ( $M = 0.74$ ;  $SD = 0.13$ ) being solved faster than difficult problems ( $M = 2.50$ ;  $SD = 0.66$ ). The main effect order,  $F(1, 64) = 0.238$ ;  $p = .627$ ;  $\eta_p^2 = .004$ , and the interaction difficulty \* order,  $F(1, 64) = 0.114$ ;  $p = .737$ ;  $\eta_p^2 = .002$ , were not significant.

### Subtractions

For accuracy, the repeated measurements ANOVA showed significant main effects of difficulty,  $F(1, 64) = 159.242$ ;  $p < .001$ ;  $\eta_p^2 = .713$ , and order,  $F(1, 64) = 4.827$ ;  $p = .032$ ;  $\eta_p^2 = .070$ . Importantly, also the interaction difficulty \* order proved significant,  $F(1, 64) = 6.290$ ;  $p = .015$ ;  $\eta_p^2 = .089$ . Pairwise comparisons showed that in difficult subtractions, switch trials were solved more accurately than repeat trials ( $p = .009$ ), while in easy subtractions there was no difference between switch and repeat trials ( $p = .675$ ). Furthermore, both easy repeat and switch trials were solved more accurately than difficult repeat ( $p < .001$ ) and switch trials ( $p < .001$ ).

Analysis of calculation times in subtractions showed significant main effects of difficulty,  $F(1, 64) = 779.774$ ;  $p < .001$ ;  $\eta_p^2 = .924$ , and order,  $F(1, 64) = 4.110$ ;  $p = .047$ ;  $\eta_p^2 = .060$ . Again, also the interaction difficulty \* order was significant,  $F(1, 64) = 9.117$ ;  $p = .004$ ;  $\eta_p^2 = .125$ . Pairwise comparisons revealed that in difficult subtractions, switch trials were solved faster than repeat trials ( $p = .010$ ), while in easy subtractions, repeat

trials were solved faster than switch trials ( $p = .009$ ). Furthermore, both easy repeat and switch trials were solved faster than difficult repeat ( $p < .001$ ) and switch trials ( $p < .001$ ).

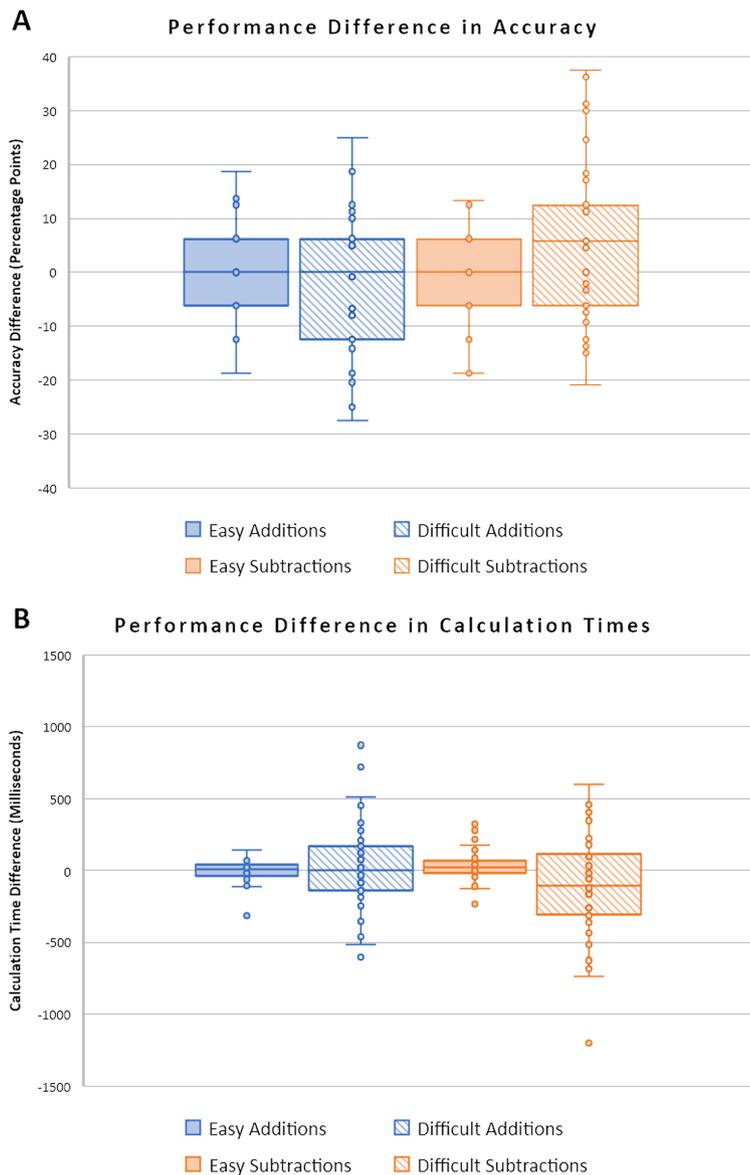


Figure 3. Switching costs / sequential difficulty effects in accuracy and calculation times.

Note. Figure 3 depicts boxplots for the switch costs / sequential difficulty effects (differences between switch and repeat trials) in accuracy (A) and calculation times (B). Circles indicate individual values.

## Event-Related (De-)Synchronization in Subtractions

### Beta Band

Results for the beta band are depicted in Figure 4 (topographic maps are given in Figure 5 as additional information).

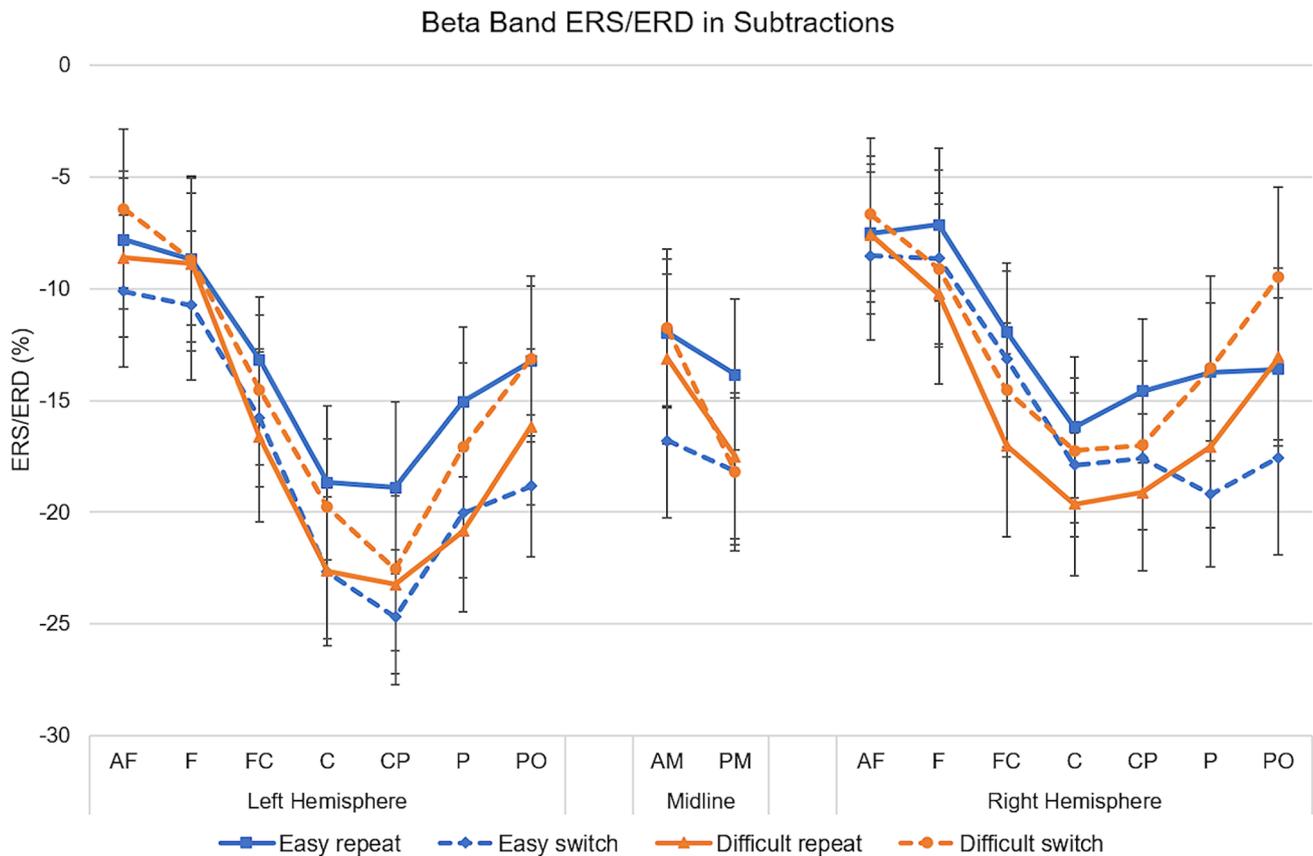


Figure 4. Beta band – Results.

Note. Figure 4 shows beta band ERS/ERD during subtractions over the left and right hemispheres as well as over midline areas. Error bars reflect the confidence intervals (95%), AF = antero-frontal; F = frontal; FC = fronto-central; C = central; CP = centro-parietal; P = parietal; PO = parieto-occipital; AM = anterior midline; PM = posterior midline.

In the left hemisphere, there was a significant interaction difficulty \* order,  $F(1, 64) = 15.286$ ;  $p < .001$ ;  $\eta_p^2 = .193$ . Pairwise comparisons showed that there was more beta band ERD (higher negative values) in easy switch ( $M = -17.55$ ;  $SD = 9.54$ ) than in easy repeat trials ( $M = -13.65$ ;  $SD = 10.48$ ;  $p < .001$ ), while in difficult problems there was a pattern in the opposite direction but no significant difference in ERD between repeat ( $M = -16.72$ ;  $SD = 12.06$ ) and switch trials ( $M = -14.60$ ;  $SD = 11.32$ ;  $p = .081$ ). Furthermore, in repeat trials ERD was stronger in difficult problems than in easy problems ( $p = .036$ ), while in switch trials, ERD was stronger in easy problems than in difficult ones ( $p = .011$ ). Additionally, there was a significant effect of location,  $F(2.92, 187.08) = 45.922$ ;  $p < .001$ ;  $\eta_p^2 = .418$ .

In the right hemisphere, there was also a significant interaction difficulty \* order,  $F(1, 64) = 7.814$ ;  $p = .007$ ;  $\eta_p^2 = .109$ . Thereby, ERD was higher in easy switch trials ( $M = -14.65$ ;  $SD = 11.11$ ) than in easy repeat trials ( $M = -12.10$ ;  $SD = 9.97$ ;  $p = .028$ ), while there again was only a pattern in the opposite direction but no significant difference between difficult repeat ( $M = -14.82$ ;  $SD = 11.93$ ) and switch trials ( $M = -12.52$ ;  $SD = 11.61$ ;  $p = .075$ ). Furthermore, there were no significant differences between easy and difficult repeat trials ( $p = .073$ ) nor between easy and difficult switch trials ( $p = .125$ ). Additionally, there were significant effects of location,  $F(2.35, 150.60) = 22.911$ ;  $p < .001$ ;  $\eta_p^2 = .264$ , and the interaction difficulty \* location,  $F(2.96, 189.29) =$

5.718;  $p = .001$ ;  $\eta_p^2 = .082$ . This interaction was mainly driven by a higher ERD in difficult ( $M = -15.77$ ;  $SD = 11.72$ ) than in easy trials ( $M = -12.53$ ;  $SD = 12.27$ ;  $p = .023$ ) in fronto-central regions, but a higher ERD in easy ( $M = -15.58$ ;  $SD = 13.31$ ) than in difficult trials ( $M = -11.27$ ;  $SD = 14.40$ ;  $p = .010$ ) in parieto-occipital regions.

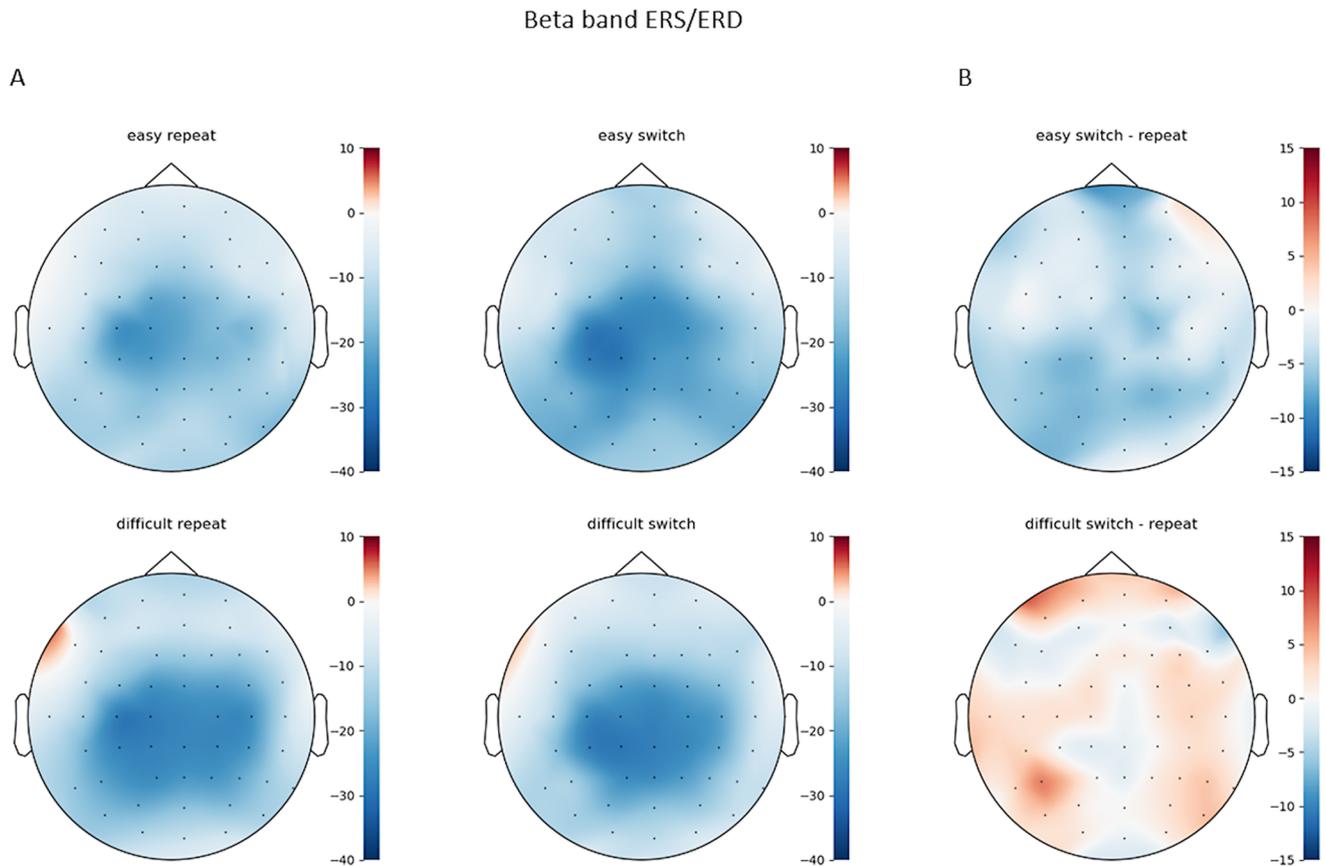


Figure 5. Beta band – Topographic maps.

Note. Figure 5 **A** shows the beta band ERS/ERD values across the scalp for the different conditions of subtractions and **B** shows the difference between repeat and switch trials (difference = switch – repeat).

In midline areas, there was a significant main effect of order,  $F(1, 64) = 4.066$ ;  $p = .048$ ;  $\eta_p^2 = .060$ , however, also a significant interaction difficulty \* order emerged,  $F(1, 64) = 6.916$ ;  $p = .011$ ;  $\eta_p^2 = .098$ . Thereby, ERD was stronger in easy switch trials ( $M = -17.49$ ;  $SD = 11.52$ ) than in easy repeat trials ( $M = -12.91$ ;  $SD = 12.29$ ;  $p = .002$ ), while there was no difference between difficult switch ( $M = -14.99$ ;  $SD = 12.93$ ) and repeat trials ( $M = -15.32$ ;  $SD = 13.38$ ;  $p = .814$ ). Furthermore, there was no significant difference between easy and difficult repeat trials ( $p = .178$ ), nor between easy and difficult switch trials ( $p = .074$ ). Additionally, there were significant effects of location,  $F(1, 64) = 10.131$ ;  $p = .002$ ;  $\eta_p^2 = .137$ , and the interaction difficulty \* location,  $F(1, 64) = 8.517$ ;  $p = .005$ ;  $\eta_p^2 = .117$ . Here, this interaction was driven by a higher amount of ERD in difficult trials in posterior midline areas ( $M = -17.85$ ;  $SD = 13.18$ ) than in anterior midline areas ( $M = -12.46$ ;  $SD = 12.99$ ;  $p < .001$ ), while there was no difference between anterior ( $M = -14.38$ ;  $SD = 11.18$ ) and posterior ( $M = -16.02$ ;  $SD = 11.82$ ;  $p = .170$ ) in easy trials.

### Upper Alpha Band

Results for upper alpha band are depicted in Figure 6 (topographic maps are given in Figure 7 as additional information). Regarding the left hemisphere, a significant interaction difficulty \* order emerged,  $F(1, 64) = 4.476$ ;  $p = .038$ ;  $\eta_p^2 = .065$ . Thereby, there was a significant difference in switch trials, with more upper alpha band ERD in easy switch trials ( $M = -23.35$ ;  $SD = 17.16$ ) than in difficult switch trials ( $M = -16.65$ ;  $SD = 24.98$ ;  $p = .042$ ). In repeat trials there was no significant difference between easy ( $M = -19.70$ ;  $SD = 20.16$ ) and difficult problems ( $M = -21.02$ ;  $SD = 23.61$ ;  $p = .670$ ). Furthermore, there was no significant difference between easy repeat and switch trials ( $p = .125$ ), nor between difficult repeat and switch trials ( $p = .118$ ). Additionally, there was an effect of location,  $F(2.91, 186.48) = 7.088$ ;  $p < .001$ ;  $\eta_p^2 = .100$ .

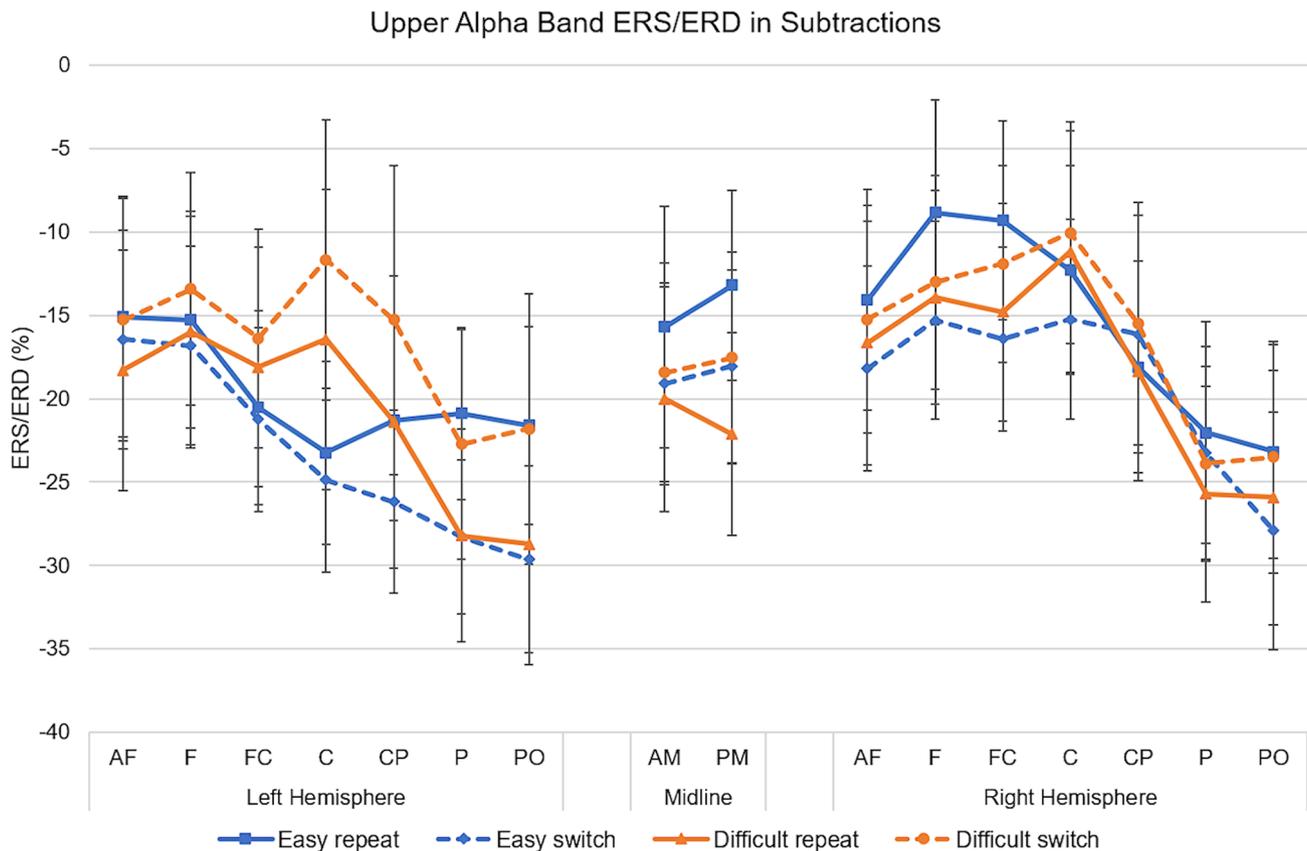


Figure 6. Upper alpha band – Results.

Note. Figure 6 shows upper alpha band ERS/ERD during subtractions over the left and right hemispheres as well as over midline areas. Error bars reflect the confidence intervals (95%), AF = antero-frontal; F = frontal; FC = fronto-central; C = central; CP = centro-parietal; P = parietal; PO = parieto-occipital; AM = anterior midline; PM = posterior midline.

In the right hemisphere only an effect of location,  $F(2.62, 167.67) = 13.650$ ;  $p < .001$ ;  $\eta_p^2 = .176$ , emerged.

Regarding midline areas, there were no significant effects of difficulty, order, location, or any interactions.

## Upper alpha band ERS/ERD

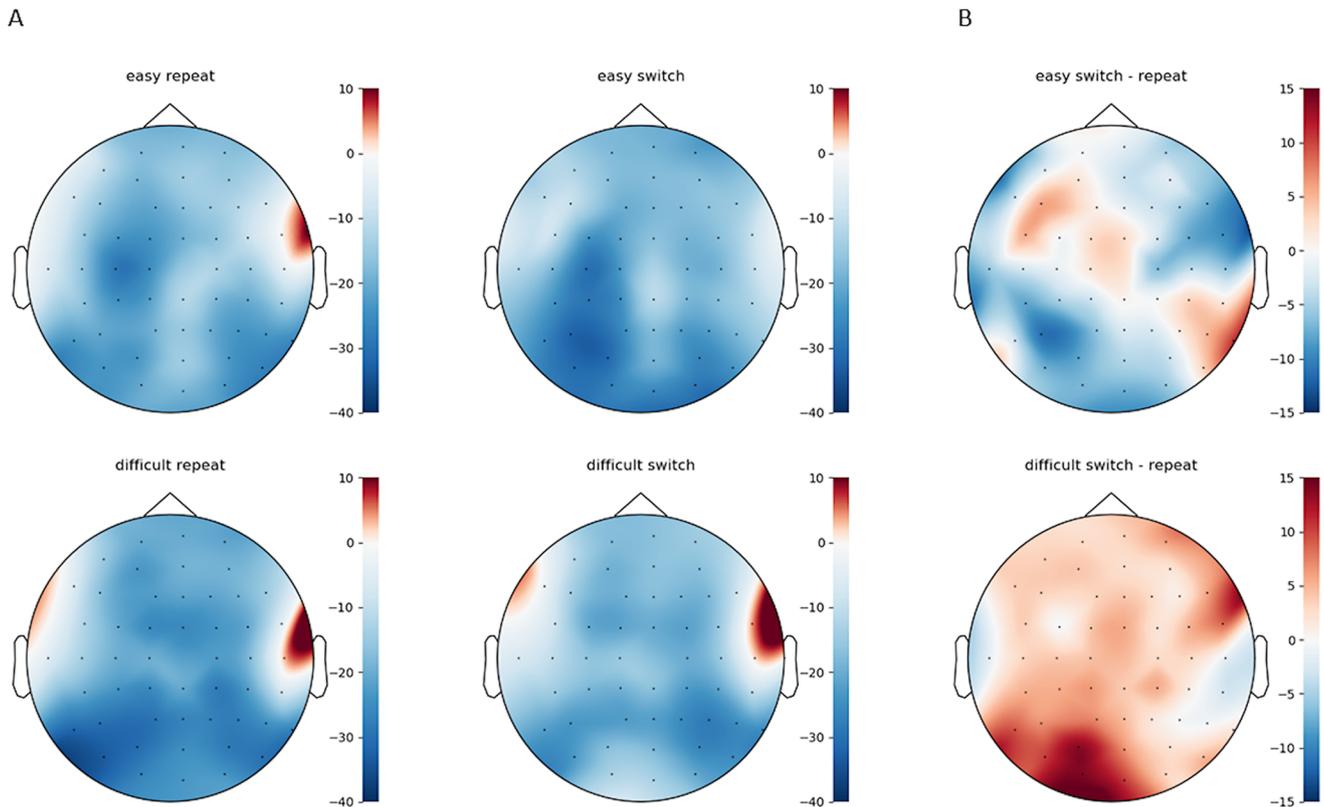


Figure 7. Upper alpha band – Topographic maps.

Note. Figure 7 **A** shows the upper alpha band ERS/ERD values across the scalp for the different conditions of subtractions and **B** shows the difference between repeat and switch trials (difference = switch – repeat).

### Lower Alpha Band

In the left hemisphere, only effects of location,  $F(2.94, 188.13) = 14.554$ ;  $p < .001$ ;  $\eta_p^2 = .185$ , and the interaction difficulty \* location,  $F(2.86, 183.31) = 6.212$ ;  $p = .001$ ;  $\eta_p^2 = .088$ , emerged. The interaction was mainly driven by a stronger ERD in easy ( $M = -16.51$ ;  $SD = 26.19$ ) as compared to difficult problems ( $M = -7.67$ ;  $SD = 37.58$ ;  $p = .035$ ) in fronto-central areas, but a higher ERD in difficult ( $M = -28.13$ ;  $SD = 25.13$ ) as compared to easy problems ( $M = -22.27$ ;  $SD = 27.02$ ;  $p = .020$ ) in parietal areas.

In the right hemisphere, the only significant effect was location,  $F(3.07, 196.50) = 10.341$ ;  $p < .001$ ;  $\eta_p^2 = .139$ .

Regarding midline areas, results were similar to the left hemisphere, with the only significant effects emerging from location,  $F(1, 64) = 9.856$ ;  $p = .003$ ;  $\eta_p^2 = .133$ , and the interaction difficulty \* location,  $F(1, 64) = 6.102$ ;  $p = .016$ ;  $\eta_p^2 = .087$ . The interaction is driven by a stronger ERD in anterior ( $M = -22.43$ ;  $SD = 27.92$ ) than in posterior areas ( $M = -11.92$ ;  $SD = 35.73$ ;  $p < .001$ ) in difficult problems, while there is no ERD difference between anterior ( $M = -16.00$ ;  $SD = 30.67$ ) and posterior ( $M = -13.89$ ;  $SD = 31.23$ ;  $p = .392$ ) in easy problems. ERD did not differ between easy and difficult problems in anterior ( $p = .050$ ) or posterior midline areas ( $p = .548$ ).

### Theta Band

In the left hemisphere, the repeated measurements ANOVA only revealed a significant effect of difficulty,  $F(1, 64) = 28.174$ ,  $p < .001$ ;  $\eta_p^2 = .306$ , with an ERS in easy problems ( $M = 9.31$ ;  $SD = 19.33$ ), while in difficult problems there even was a slight ERD ( $M = -2.54$ ;  $SD = 14.83$ ), and of location,  $F(2.90, 185.41) = 7.022$ ;  $p = .002$ ;  $\eta_p^2 = .076$ . Neither order, nor any of the interactions showed an effect.

Similarly in the right hemisphere, there only was an effect of difficulty,  $F(1, 64) = 16.556$ ;  $p < .001$ ;  $\eta_p^2 = .206$ , with an ERS in easy problems ( $M = 6.38$ ;  $SD = 16.58$ ) and an ERD in the difficult problems ( $M = -3.03$ ;  $SD = 14.03$ ), and location,  $F(3.08, 197.20) = 3.857$ ,  $p = .010$ ;  $\eta_p^2 = .057$ . Again, neither order, nor any of the interactions showed an effect.

For midline areas, only difficulty showed an effect,  $F(1, 64) = 15.888$ ;  $p < .001$ ;  $\eta_p^2 = .199$ , again with an ERS in easy problems ( $M = 11.34$ ;  $SD = 20.35$ ) and an ERD in difficult problems ( $M = -0.29$ ;  $SD = 19.31$ ).

### Correlations Between Switching Costs / Sequential Difficulty Effects, Working Memory, and IST Arithmetic Scores

A larger amount of sequential difficulty effects in accuracy in difficult subtractions (higher accuracy in switch as compared to repeat trials) was significantly related to lower IST arithmetic scores ( $r = -.267$ ;  $p = .032$ ). No significant correlations between the amounts of switching costs/sequential difficulty effects and overall accuracy ( $M = 89.3\%$ ;  $SD = 5.8$ ) or reaction times ( $M = 0.59$  seconds;  $SD = 0.60$ ) in the WM task were found (all  $p > .05$ ).

### Additional Analyses

#### Strategy Use

Participants indicated that in additions they solved 87.8% ( $SD = 21.1$ ) of the easy problems by fact retrieval and 77.4% ( $SD = 20.8$ ) of the difficulty problems by procedural calculation. In subtractions, participants indicated that they solve 83.8% ( $SD = 23.4$ ) of the easy problems by fact retrieval and 83.7% ( $SD = 18.3$ ) of the difficult problems by procedural calculation.

#### Inter-Trial Interval Durations

There was no difference between the average durations of the inter-trial intervals after easy ( $M = 4.74$  seconds;  $SD = 0.18$ ) and difficult additions ( $M = 4.72$  seconds;  $SD = 0.72$ ),  $t(64) = 0.284$ ;  $p = .777$ . In subtractions however, there was a difference between the average inter-trial intervals after easy ( $M = 4.62$  seconds;  $SD = 0.24$ ) and difficult problems ( $M = 4.34$  seconds;  $SD = 0.78$ ),  $t(64) = 3.653$ ;  $p < .001$ .

## Discussion

The main objective of this study was to investigate whether sequential difficulty effect patterns, which appear when switching between procedural calculation problems of different difficulty or applied strategy (Schneider & Anderson, 2010; Uittenhove & Lemaire, 2012), are also present when switching between easy arithmetic problems, mainly solved by fact retrieval, and difficult problems, solved primarily by procedural calculations. As the processing of easy and difficult arithmetic problems leads to distinct ERS/ERD patterns related to the different involvement of cognitive functions like executive control and WM (Grabner & De Smedt, 2011;

Tschentscher & Hauk, 2016), an additional aspect was to assess whether sequential difficulty effects also affect these ERS/ERD patterns.

On the behavioral level, the main findings were sequential difficulty effect patterns in calculation times and, to a lesser extent, accuracy in subtractions. In line with our predictions, both easy and difficult subtractions were solved more slowly if they were preceded by a difficult problem. A partially similar pattern emerged in accuracy. In easy problems, however, there were no differences in accuracy between switch and repeat trials. Hence, for calculation time in subtractions and to a lesser extent also for accuracy, the results are well in line with prior literature on sequential difficulty effects (e.g. Schneider & Anderson, 2010; Uittenhove & Lemaire, 2012).

In additions, however, we did not find any differences between repeat and switch trials. The absence in the more difficult addition problems is thereby partly in line with prior literature, as asymmetries are assumed to arise from less switch costs in the more difficult or less automatized task or strategy as compared to the switch costs in the easy or more automatized one (e.g. Allport et al., 1994; Campbell, 2005; Schneider & Anderson, 2010; Uittenhove & Lemaire, 2012). A possible explanation for the absence of any differences is that difficulty switching costs and sequential difficulty effects might have a comparable magnitude in the additions used and balance each other out. Switching costs should lead to poorer performance in switch trials, while sequential difficulty effects should lead to poorer performance in repeat trials (Schneider & Anderson, 2010). This explanation seems plausible for the complex additions in this study, because additions are easier than subtractions (as indicated by higher accuracy and faster calculation times) and resource depletion might therefore be less pronounced. Regarding the absence of differences in the easy problems, a similar explanation cannot hold, as difficulty switching costs and sequential difficulty effects should add up in these trials. Both should negatively affect performance in the easy switch trials. However, the small, easy additions were the easiest problems in this study, with an average calculation time of just 0.74 seconds and an average accuracy of 96.05%. Hence, it is possible that there were some kind of ceiling effects in easy additions and the costs of switching and resource depletion might not have been strong enough to impact the performance in these very easy problems.

An additional interesting aspect regarding sequential difficulty effects is whether these effects are related to arithmetic performance and WM. In case of basic arithmetic abilities, only the magnitude of the sequential difficulty effects in accuracy (less accuracy in repeat than in switch trials) in difficult subtractions was related to performance. However, basic arithmetic abilities were assessed by a single test consisting of a series of more complex problems. This might be a reason why only the sequential difficulty effects in difficult problems appear to be related. Nevertheless, this is an interesting new aspect because it indicates that the magnitude of sequential difficulty effects might relate to arithmetic performance in general (although it has to be mentioned that the correlation was rather low and has to be interpreted cautiously). Furthermore, there were no correlations between sequential difficulty effects or switching costs and WM performance. This is somewhat surprising, as Uittenhove and Lemaire (2013) found correlations between the magnitude of sequential difficulty effects and three measures of working memory. However, while these authors used operation span, running span, and reading span tasks, in the current study a 2-Back task with letters was applied. This difference in operationalization of WM might be the reason for the lack of correlation found in this study, as N-Back tasks and WM span tasks might assess different aspects of WM (Kane, Conway, Miura, & Colflesh, 2007).

At the electrophysiological level, the sequential difficulty effects in subtractions were reflected in EEG, albeit only partially as expected. Based on prior research (De Smedt et al., 2009; Grabner & De Smedt, 2011;

Schneider & Anderson, 2010; Tschentscher & Hauk, 2016), we anticipated a stronger alpha and beta band ERD in easy switch than in easy repeat trials, and a stronger alpha and beta band ERD in difficult repeat than in difficult switch trials. Regarding alpha bands, an effect of order only emerged in the upper alpha band in the left hemisphere, with a stronger ERD in easy switch than in difficult switch trials. This was somewhat unexpected, as prior research pointed to a generally stronger ERD in difficult, procedural arithmetic problems, especially over posterior regions (De Smedt et al., 2009; Grabner & De Smedt, 2011). Such an effect of difficulty was absent in this data. The stronger ERD in easy switch trials as compared to difficult switch trials indicate that sequential difficulty effects might play an important role in ERS/ERD patterns in upper alpha band during mental arithmetic, which has not been considered so far.

The most interesting results regarding sequential difficulty effects were found in the beta band, with the strongest effects in the left hemisphere. Specifically, ERD was stronger in easy switch than in easy repeat trials (with a trend to but not significantly more ERD in difficult repeat than in difficult switch trials). Similar to the upper alpha band, ERD was even stronger in easy switch trials than in difficult switch trials, but lower in easy repeat than in complex repeat trials. A comparable pattern emerged over the right hemisphere and midline areas, although only the differences between easy repeat and switch trials proved to be significant. Hence, regardless of difficulty of the trial at hand, ERD in beta band (especially in the left hemisphere) seems stronger in the trials preceded by a difficult one. A stronger decrease of beta band power in right frontal, parietal, and temporal areas during procedural calculations (difficult) as compared to fact retrieval (easy) has been assumed to reflect higher demands on an executive control network (Tschentscher & Hauk, 2016). Hence, the finding of sequential difficulty effects in beta band ERD is novel and supports the view that executive control is one of the possible reasons for sequential difficulty effects (Schneider & Anderson, 2010). Furthermore, these results extend our understanding of sequential difficulty effects, as they show that these effects not only affect the strength of cortical activity over the time of task processing as assessed by ERPs (Uittenhove et al., 2013), but also modulate task related oscillatory patterns associated with the involvement of different cognitive functions. Uittenhove et al. (2013) argued that their findings indicate sequential difficulty effects mainly affecting early to central stages of task processing, mostly the retrieval of procedural strategies. This might be in line with the results of this study, as sequential difficulty effects in beta band ERS/ERD patterns points to an involvement of executive control functions (Tschentscher & Hauk, 2016). These functions could be important for the selection and conduction of the correct / best strategy to solve a task and to inhibit competing ones. Depletion of executive control functions might impair early parts of mental arithmetic processes, but future research is needed to confirm this notion.

Additionally, we could replicate earlier findings of stronger theta band ERS in easy, fact retrieval problems as compared to difficult, procedural calculation problems (De Smedt et al., 2009; Grabner & De Smedt, 2011; Tschentscher & Hauk, 2016), while no sequential difficulty effects emerged. This absence is in line with both our expectations and prior research on theta band in the context of mental arithmetic, mainly interpreting theta band ERS as reflecting memory retrieval processes (De Smedt et al., 2009) while sequential difficulty effects are thought to be based on depletion of executive control or WM functions (Schneider & Anderson, 2010).

There are some limitations in this study. While the differences in average accuracy and calculation times between fact retrieval (easy) and procedural calculation (difficult) problems were as expected (lower accuracy and longer calculation times in the difficult procedural calculation problems), we did not assess the applied strategy trial by trial, but as a general self-rating at the end of each block. Thereby, participants indicated that

they solved easy problems primarily by fact retrieval, and difficult problems primarily by procedural calculation, but not all of them were solved by application of the assumed strategy. Hence, it is possible that not every switch of problem difficulty was accompanied by a switch of strategy and vice versa. Applying a trial by trial assessment of strategy use and only analyzing those trials in which the reported strategy fits to the given problem difficulty might lead to even clearer sequential difficulty effect patterns in future studies. Furthermore, in the current study participants had to calculate the results and were then presented with three possible answers to choose from. This is different from a production task, where participants have to produce the answer directly, used in some other studies (Lemaire & Lecacheur, 2010; Uittenhove & Lemaire, 2012). The procedure used has some advantages for ERS/ERD analysis as the relevant EEG data is less contaminated by moving/talking activation and artifacts. However, this procedure might be bit easier than a production task and it cannot be ruled out that participants reconsidered some results upon seeing the solution options, especially in the difficult problems. As sequential difficulty effects are thought to be based in the difference of difficulty between the two task or strategies these effects might be even more pronounced if using a production task. Additionally, correlations between WM performance and sequential difficulty effects were also found using a productions task, hence the use of a, probably easier procedure in the current study might be another reason for the absence of a similar correlation in this study. Another possible limitation lies in the use of variable inter-trial intervals. The variability was a requisite of the neuro-stimulation part (ensuring equal set durations) that followed this experiment. Thereby, the difference in duration based on the remaining time from calculating, solution selection and strategy selection ensured that all participants solved the same amount of problems in the same amount of total task time. However, as switching costs and sequential difficulty effects seem to be moderated by the time between two trials (Bryck & Mayr, 2008; Uittenhove & Lemaire, 2013) this variability might have led to some trials with long inter-trial intervals before and some with short ones. In order to prevent differences in inter-trial intervals between easy and difficult problems (easy are solved faster) the maximal calculation time for easy problems was two seconds shorter, also shortening the inter-trial interval by the same amount. In the additions this approach worked well, the inter-trial intervals after easy and difficult problems were nearly identical, but in subtractions, inter-trial intervals were still shorter after difficult problems than after easy problems. However, with about 280 ms, the difference between intervals after easy and difficult subtractions was rather small compared to the average inter-trial intervals of 4.48 seconds. Furthermore, the average response-stimulus intervals were close to the long response-stimulus interval used by Bryck and Mayr (2008) and much longer than the other intervals investigated in these prior studies (Uittenhove & Lemaire, 2013). Hence, we expect the influence to be small and argue that the occurrence of sequential difficulty effects is a sign that the generally long and variable intervals were no problem in this setup. However, investigating whether effects are different if very short, constant intervals are used might be interesting for future research.

Overall, these results extend prior knowledge by showing that switching between easy and difficult subtractions induces sequential difficulty effects and, more importantly, that these effects are reflected in differences in ERS/ERD patterns and relate to basic arithmetic performance. Especially the effects on ERS/ERD patterns are valuable information for future studies, as they show that arithmetic strategy effects found earlier (e.g. De Smedt et al., 2009; Grabner & De Smedt, 2011; Tschentscher & Hauk, 2016) can be moderated by effects resulting from the order of easy and difficult problems. Hence, it will be important to consider these effects in studies on arithmetic problem solving. Regarding mental arithmetic in general and sequential difficulty effects in particular, oscillations in the beta band might yield valuable new information on underlying mechanisms.

Furthermore, the association between the magnitude of sequential difficulty effects and basic arithmetic performance adds a new aspect to individual differences in arithmetic abilities.

### Funding

The authors have no funding to report.

### Competing Interests

The authors have declared that no competing interests exist.

### Acknowledgments

The authors have no support to report.

### Data Availability

For this study, a dataset is freely available (see the [Supplementary Materials](#) section).

## Supplementary Materials

The dataset and short descriptions of the variables used for the analyses in this paper are available on the PsychArchives open-access repository (see Index of [Supplementary Materials](#) below).

### Index of Supplementary Materials

Mosbacher, J. A., Brunner, C., & Grabner, R. H. (2020). *Supplementary materials to "More problems after difficult problems? Behavioral and electrophysiological evidence for sequential difficulty effects in mental arithmetic"* [Research data and codebook]. PsychOpen. <https://doi.org/10.23668/psycharchives.3011>

## References

- Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and unconscious information processing* (pp. 421–452). Cambridge, MA, USA: MIT Press.
- Aschenbrenner, S., Tucha, O., & Lange, K. W. (2000). *Regensburger Wortflüssigkeits-Test: RWT*. Göttingen, Germany: Hogrefe.
- Borkenau, P., & Ostendorf, F. (2008). *NEO-Fünf-Faktoren-Inventar nach Costa und McCrae*. Göttingen, Germany: Hogrefe.
- Bryck, R. L., & Mayr, U. (2008). Task selection cost asymmetry without task switching. *Psychonomic Bulletin & Review*, *15*(1), 128–134. <https://doi.org/10.3758/PBR.15.1.128>
- Campbell, J. I. D. (2005). Asymmetrical language switching costs in Chinese-English bilinguals' number naming and simple arithmetic. *Bilingualism*, *8*(1), 85–91. <https://doi.org/10.1017/S136672890400207X>
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PIR) and NEO Five Factor Inventory (NEO-FFI) professional manual*. Odessa, FL, USA: Psychological Assessment Resources.

- De Smedt, B., Grabner, R. H., & Studer, B. (2009). Oscillatory EEG correlates of arithmetic strategy use in addition and subtraction. *Experimental Brain Research*, *195*(4), 635-642. <https://doi.org/10.1007/s00221-009-1839-9>
- Dweck, C. S., Chiu, C.-y., & Hong, Y.-y. (1995). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry*, *6*(4), 267-285. [https://doi.org/10.1207/s15327965pli0604\\_1](https://doi.org/10.1207/s15327965pli0604_1)
- Grabner, R. H., & De Smedt, B. (2011). Neurophysiological evidence for the validity of verbal strategy reports in mental arithmetic. *Biological Psychology*, *87*(1), 128-136. <https://doi.org/10.1016/j.biopsycho.2011.02.019>
- Grabner, R. H., & De Smedt, B. (2012). Oscillatory EEG correlates of arithmetic strategies: A training study. *Frontiers in Psychology*, *3*, Article 428. <https://doi.org/10.3389/fpsyg.2012.00428>
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., . . . Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, Article 267. <https://doi.org/10.3389/fnins.2013.00267>
- Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., . . . Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *NeuroImage*, *86*, 446-460. <https://doi.org/10.1016/j.neuroimage.2013.10.027>
- Imbo, I., & Vandierendonck, A. (2007). The role of phonological and executive working memory resources in simple arithmetic strategies. *The European Journal of Cognitive Psychology*, *19*(6), 910-933. <https://doi.org/10.1080/09541440601051571>
- Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attention control, and the N-Back Task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 615-622. <https://doi.org/10.1037/0278-7393.33.3.615>
- Lemaire, P., & Lecacheur, M. (2010). Strategy switch costs in arithmetic problem solving. *Memory & Cognition*, *38*(3), 322-332. <https://doi.org/10.3758/MC.38.3.322>
- Liepmann, D., Beauducel, A., Brocke, B., & Amthauer, R. (2007). *I-S-T 2000R Intelligenz-Struktur-Test 2000 R*. Göttingen, Germany: Hogrefe.
- Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1-2), 8-13. <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., . . . Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195-203. <https://doi.org/10.3758/s13428-018-01193-y>
- Reynolds, C. R. (2002). *Comprehensive trail making test (CTMT)*. Austin, TX, USA: Pro-Ed.
- Schneider, D. W., & Anderson, J. R. (2010). Asymmetric switch costs as sequential difficulty effects. *Quarterly Journal of Experimental Psychology*, *63*(10), 1873-1894. <https://doi.org/10.1080/17470211003624010>
- Steingrüber, H. J., & Lienert, G. A. (1971). *Hand-Dominanz-Test: HDT*. Göttingen, Germany: Hogrefe.
- Tschentscher, N., & Hauk, O. (2016). Frontal and parietal cortices show different spatiotemporal dynamics across problem-solving stages. *Journal of Cognitive Neuroscience*, *28*(8), 1098-1110. [https://doi.org/10.1162/jocn\\_a\\_00960](https://doi.org/10.1162/jocn_a_00960)

- Uittenhove, K., & Lemaire, P. (2012). Sequential difficulty effects during strategy execution – A study in arithmetic. *Experimental Psychology*, 59(5), 295-301. <https://doi.org/10.1027/1618-3169/a000157>
- Uittenhove, K., & Lemaire, P. (2013). Strategy sequential difficulty effects vary with working-memory and response-stimulus-intervals: A study in arithmetic. *Acta Psychologica*, 143(1), 113-118. <https://doi.org/10.1016/j.actpsy.2013.02.007>
- Uittenhove, K., Poletti, C., Dufau, S., & Lemaire, P. (2013). The time course of strategy sequential difficulty effects: An ERP study in arithmetic. *Experimental Brain Research*, 227(1), 1-8. <https://doi.org/10.1007/s00221-012-3397-9>