The Cognitive Boost at the Peak of Circadian Arousal is Not as General and Robust as Previously Thought!

Alodie Rey-Mermet & Nicolas Rothen

Faculty of Psychology

UniDistance Suisse

Author Note

Alodie Rey-Mermet https://orcid.org/0000-0002-6385-9328

Nicolas Rothen https://orcid.org/0000-0002-8874-8341

Some results of the present work were presented at the 62nd Annual Meeting of the Psychonomic Society, virtual conference, USA, November 2021, at the 22nd meeting of the European Society for Cognitive Psychology, Lille, France, August 2022, and at the 17th Conference of the Swiss Psychological Society, Zurich, Switzerland, September 2022. The design of the present study and its analyses were pre-registered on the Open Science Framework (OSF). The preregistration can be accessed at `https://osf.io/tywu7`. The data, the experiment codes, the analysis codes, and the results are available at the OSF and can be accessed at `https://osf.io/ngfxv`. The preprint of this article can be accessed on PsyArXiv. A. Rey-Mermet is currently supported by a grant from the Swiss National Science Foundation (Grant 100014_207865). N. Rothen is currently supported by a grant from the Swiss National Science Foundation (Grant 10001CM_204314).

We have no known conflict interest to disclose.

Word count: 9'967

The authors made the following contributions. Alodie Rey-Mermet: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing - Original Draft, Writing - Review & Editing, Visualization; Nicolas Rothen: Conceptualization, Methodology, Validation, Investigation, Resources, Writing - Original Draft, Writing - Review & Editing.

Correspondence concerning this article should be addressed to Alodie Rey-Mermet, Schinerstrasse 18, 3900 Brig, Switzerland. E-mail: `alodie.rey-mermet@fernuni.ch`

Abstract

The synchrony effect – that is, the observation of better cognitive performance at the peak of circadian arousal than at off peak – is accepted as common knowledge. However, empirical evidence for this effect is mixed. This may be explained by two methodical challenges. First, most studies are underpowered. Second, they include one task, but tasks differ across studies. Here, we tested the synchrony effect by focusing on two universal psychological constructs (i.e., the ability to maintain information temporarily and the ability to avoid being distracted by irrelevant information). We addressed the methodical challenges by asking 191 young adults with morning or evening chronotypes to perform eight tasks at on- and off-peak times. Moreover, in addition to analysing at the individual task-level, we estimated the constructs at the latent-variable level (i.e., as the common variance across the tasks without the impact of task-specific requirements). The results showed strong evidence against a synchrony effect. We conclude that the synchrony effect is most likely a methodical artefact and discuss the impact of our research on psychological science and scientific research more widely.

Public significance statements: It seems to be common knowledge that morning types perform better in the morning and evening types perform better in the evening. However, the present study suggests that this seemingly well-established knowledge is based on methodical shortcomings. Thus, requests to adapt working schedules are premature.

Keywords: synchrony effect, attentional control, executive functions, working memory, individual differences

The Cognitive Boost at the Peak of Circadian Arousal is Not as General and Robust as Previously Thought!

Environmental conditions at day and night are fundamentally different. Therefore, it is very hard to imagine that they do not influence human behavior and experience. Indeed, common knowledge as well as a large body of literature in psychology and neurosciences suggest that systematic variation in circadian arousal over the course of 24 hours differentially affects cognitive performance (e.g., Murray et al., 2009; Schmidt et al., 2007). However, due to methodical challenges, the evidence for effects of circadian arousal on cognitive performance is not as clear as it seems. Here, we address these methodical challenges to clarify whether variation in circadian arousal affects human cognition.

The synchrony effect describes the observation of better cognitive performance on the peak of circadian arousal in comparison to cognitive performance at off-peak times (e.g., May et al., 2005; May et al., 1993; May, 1999; May & Hasher, 1998). Individual differences in circadian arousal follows a normal distribution according to which individuals can be subdivided into different chronotypes. Morning types experience their peak of circadian arousal in the morning. Accordingly, they raise earlier and go earlier to bed. By contrast, evening types experience their peak of circadian arousal later during the day in the evening. Accordingly, they raise later and go later to bed (Griefahn et al., 2001; Horne & Ostberg, 1976). In line with the synchrony effect, morning types should exhibit better cognitive performance in the morning than in the evening. In contrast, evening types should exhibit better cognitive performance in the evening than in the morning.

Although the synchrony effect is accepted as true and is not really questioned in daily life, empirical evidence is more equivocal. For instance, evidence is mixed for tasks measuring attentional control or executive functions (i.e., the ability to maintain goal-relevant information when facing distraction; Draheim et al., 2022; von Bastian et al., 2020). While some studies present evidence for the synchrony effect on attentional control (e.g., Bennett et al., 2008; Hahn et al., 2012; Hasher et al., 2002; Intons-Peterson et al., 1998; Lara et al., 2014; May, 1999; May & Hasher, 1998), others are not able to detect such an effect on attentional control (e.g., Bennett et al., 2008; Heimola et al., 2021; Knight & Mather, 2013; Li et al., 1998; Matchock & Mordkoff, 2009; May & Hasher, 1998; Schmidt et al., 2012). The same situation applies to working memory (i.e., the ability to manipulate and maintain information for a short duration; e.g., Baddeley, 2012). That is, some studies find evidence for a synchrony effect on working memory (e.g., Rowe et al., 2009; Schmidt et al., 2015; West et al., 2002), while other studies suggest no synchrony effect on working memory (e.g., Ceglarek et al., 2021; Heimola et al., 2021; Lewandowska et al., 2018; Schmidt et al., 2015).

Based on this mixed evidence, it is tempting to conclude that the synchrony effect is not general and robust. However, the current evidence may be distorted by at least two methodical challenges. First, almost all studies investigating the synchrony effect on human cognition were statistically underpowered. For example, more than half of the studies included small sample sizes in a between-subjects design. Second, most studies included one task, and the tasks differed across the studies. Only including one task per study is questionable because of the task-impurity problem. That is, each cognitive task measures not only the construct of interest but also other constructs and random noise (e.g., Miyake & Friedman, 2012). Thus, the mixed evidence regarding the synchrony effect may result from different studies controlling more or less well for this task-impurity problem. Moreover, including different tasks across different studies is problematic because these different tasks may not assess the same construct. For example, working memory is often synonymously used for short-term memory (i.e., the ability only to maintain information for a short duration; e.g., Aben et al., 2012). Thus, the mixed evidence regarding the synchrony effect in working memory may result from different constructs being measured. For attentional control, this issue is even greater. Recent research has emphasized that different tasks used to assess attentional control are not measuring the same construct (e.g., Karr et al., 2018; Rey-Mermet et al., 2019; Rey-Mermet et al., 2020, 2018). Therefore, the mixed evidence regarding the synchrony effect on attentional control may be explained by the missing convergent validity of the attentional-control tasks.

The goal of the present study was to determine the scope and robustness of the synchrony effect on human cognition by focusing on universal psychological constructs and by addressing the outlined methodical challenges. Specifically, we aimed to investigate the synchrony effect on working memory and attentional control using two sets of four tasks. In one set, the tasks required temporary maintenance of information (i.e., short-term memory). In the other set, the tasks required temporary maintenance and manipulation of information (i.e., working memory). Using structural equation modeling, we modeled attentional control as the goal-directed nature of the working-memory tasks without their maintenance aspects. The modeling approach also enables us to solve the task-impurity problem by capturing the constructs of interest (i.e., attentional control and short-term maintenance) as the shared variance across the measures. Furthermore, we used a within-subject design in which a large sample of participants were tested at both peak and off-peak times.

We hypothesized that if there is an effect of circadian arousal on working memory and attentional control (e.g., May & Hasher, 1998; Rowe et al., 2009; West et al., 2002), we should find better performance at peak than at off-peak times for the short-term memory and working-memory tasks as well as for the latent constructs of short-term maintenance and attentional control. In contrast, if circadian arousal has no effect on working memory and attentional control (e.g., Matchock & Mordkoff, 2009; May & Hasher, 1998; Schmidt et al., 2015), performance should not differ between peak and off-peak times at the individual-task level and the latent-variable level.

To anticipate our results, we find strong evidence against an effect of circadian arousal on human cognition. We must conclude that the long-standing observation of a synchrony effect in cognitive performance is most likely a methodical artefact.

Methods

Transparency and openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study (Simmons et al., 2012). All deidentified data, experiment codes, analysis codes, and research materials are publicly accessible at `https://osf.io/ngfxv`. This study's design and its analysis were pre-registered. The preregistration can be accessed at `https://osf.io/tywu7`. For all analyses, we used R (Version 4.2.1; R Core Team, 2021) and the R-packages *afex* (Version 1.1.1; Singmann et al., 2021), *BayesFactor* (Version 0.9.12.4.4; Morey & Rouder, 2021), *ggplot2* (Version 3.3.6; Wickham, 2016), *lavaan* (Version 0.6.12; Rosseel, 2012), *papaja* (Version 0.1.1; Aust & Barth, 2020), *psych* (Version 2.2.9; Revelle, 2021), *semTools* (Version 0.5.6; Jorgensen et al., 2021), and *splithalf* (Version 0.8.2; Parsons, 2020).

Participants

All participants were recruited and tested by the students from UniDistance Suisse who took part in the course M08 in fall term 2020 and in the course M1 in spring term 2021. To avoid any selection bias, we tested all participants who were willing to participate. Due to this data-collection procedure, the analyses were planned to be performed on all chronotypes (morning, evening, and neutral) as well as on morning and evening types only. Therefore, we first determined the target sample size for a sample including all chronotypes. Based on Brysbaert's (2019) recommendation according to which an effect size of .196 – that is, a Cohen's d of .40 – represents "a good first estimate of the smallest effect size of interest in psychological research" (see p. 1), we hypothesized a measurement model including four latent variables and 16 manifest variables with an anticipated effect size of .196. We determined the sample size with the a-priori Sample Size Calculator for Structural Equation Models (retrieved from http://www.danielsoper.com/statcalc) by using a statistical power of .90 and a probability level of .05 (Soper, 2018). In this case, the target sample size is 453. Nevertheless, if circadian arousal affects human cognition (e.g., May et al., 2005; May et al., 1993; May, 1999; May & Hasher, 1998), the synchrony effect should be more clear-cut in the morning and evening types. In this case, we opted for a slightly larger effect size (i.e., .287 or a Cohen's d of .50, see Brysbaert, 2019), and the target sample size is 194. We first focus on the analyses on the morning and evening types. The analyses on the complete sample are then presented as part of the multiverse-analysis approach.

In total, 689 young participants were tested. Participants were students' acquaintance and were not paid. A complete description of the exclusion criteria used in the present study is presented in Table 1. The final sample consisted of 191 young adults. One hundred and thirty-one participants were morning types, and 60 were evening types. Demographic characteristics for the final sample are summarized in Table 2. The study was approved by the ethics committee of the University (approval number: 2020-06-00001), and all participants gave written informed consent.

Material

All tasks and questionnaires were programmed using lab.js (Henninger et al., 2019) on a computer using a 31 x 17.4 cm screen size and a 1920 x 1080 px screen resolution. A set of four tasks was used to measure short-term memory. These tasks were simple-span tasks with either digits, letters, matrices, or arrows as materials. They were programmed following Kane et al. (2004). A set of four tasks was used to measure working memory. These tasks were complex-span tasks and updating tasks with numerical or spatial materials. They were programmed following Rey-Mermet et al. (2019). Each task had two versions (i.e., Version 1 and Version 2). The two versions only differed in the presentation order of the stimulus exemplars. In each version, the same pseudorandom presentation order was administered for all participants. For all tasks, each stimulus exemplar was presented approximately equally often. Unless specified otherwise, each event (e.g., stimulus or prompt) was presented centrally in black color and in 36-point sans-serif font. In all tasks, the feedback consisted of a smiling face after a correct response and a frowning face after an error. Both feedback stimuli comprised a width and a height of 2.29° visual angle at a viewing distance of 60 cm. In the next paragraphs, we present the short-term memory and working-memory tasks as well as the questionnaires separately.

Short-term memory and working-memory tasks

In all short-term memory tasks, a trial consisted of an encoding phase followed by a recall phase. In the working-memory tasks, there was in addition either a distractor task for the complex spans or updating steps for the updating tasks. In all spans, set size refers to the number of memoranda to be remembered during each trial. Although the number of memoranda was limited in most tasks (e.g., the digits one until nine for the digit simple span), memoranda did not repeat within a trial. Next, we describe each task in detail.

In the digit simple span, the encoding phase consisted of memorizing sequences of digits. The digits were presented in set sizes ranging from two to nine digits. In the recall phase, digits had to be recalled in correct serial order. Thus, a text field was presented in the center of the screen. In addition, a counter (e.g., “digit 1” for the first digit to be recalled) was presented on the upper part of the screen (2.86° visual angle) to keep track of the serial order.

The letter simple span was similar to the digit simple span, except for the following modifications. First, the memoranda were the uppercase letters B, F, H, J, L, M, Q, R, and X. Second, set sizes ranged from two to eight letters. Third, at recall, letters had to be recalled in correct serial order by clicking sequentially on the corresponding letters in the matrix. To this end, all nine letters were presented in a 3 x 3 matrix (with a length of 4.68° visual angle). A counter (e.g., “letter 1” for the first letter to be recalled) was also presented in 32-point font on the upper part of the screen (3.82° visual angle).

The matrix simple span was similar to the letter simple span, except for the following modifications. First, the memoranda were the positions of squares in a 4 x 4 matrix. The matrix had a length of 4.96° visual angle and each square of the matrix had a length of 1.24° visual angle. Second, set sizes ranged from two to seven squares. Third, at recall, an empty 4 x 4 matrix was displayed. The counter included the word "position" (e.g., "position 1" for the first position to be recalled).

The arrow simple span was similar to the letter and matrix simple spans, except for the following modifications. First, the memoranda were short and long arrows radiating out from the center of the screen. Short arrows had a length of 0.86° visual angle, whereas long arrows had a length of 1.72° visual angle. Each arrow pointed at either 0°, 45°, 90°, 135°, 180°, 225°, 270°, or 315°. Second, set size ranged from two to six arrows. Third, at recall, all short arrows were presented in a 3 x 3 matrix on the left side of the screen, and all long arrows were also presented in a 3 x 3 matrix but on the right side of the screen. Both matrices had a length of 6.30°, and they were separated by a length of 2.10° visual angle. The counter included the word "arrow" (e.g., "arrow 1" for the first arrow to be recalled).

In the numerical complex span, the encoding phase consisted of memorizing sequences of three to five two-digit numbers. Between the presentation of these memoranda, the distractor task was presented. Thus, one equation that was either valid (e.g., "6-6=0") or invalid (e.g., "5+11=18") was displayed. The validity of the equation was judged by pressing the left- and right-pointing arrow key, respectively. This response mapping was presented in 24-point monospace font with a width of 12.35° visual angle on the lower part of the screen (4.30° visual angle). At recall, the numbers had to be recalled in correct serial order. Thus, a text field was presented in the center of the screen. In addition, a counter (e.g., "number 1" for the first number to be recalled) was presented on the upper part of the screen (2.86° visual angle) to keep track of the serial order.

The spatial complex span task was similar to the numerical complex span, except for the following modifications. First, four to six red squares were presented sequentially in a 5 x 5 matrix. The matrix had a length of 6.2° visual angle and each square of the matrix had a length of 1.24° visual angle. Second, the distractor task consisted of judging whether the pattern emerging from four squares presented concurrently and arranged in an L-shape was vertical or horizontal. Third, at recall, positions had to be recalled in correct serial order by clicking sequentially on the corresponding positions in the matrix. Thus, a 5 x 5 matrix was presented. In addition, a counter (e.g., "position 1" for the first position to be recalled) was displayed above the matrix (i.e., 4.30° visual angle).

In the numerical updating task, the encoding phase consisted of memorizing four digits (ranging from one to nine) presented in four different colors (i.e., red, blue, green, and orange). The digits were displayed centrally, each separated by a width of 1.43° visual angle. In the following updating steps, the digit to be updated was presented centrally in one of the four colors. In the recall phase, the most recent digit of each color had to be recalled. Thus, the word "digit" was displayed in the color corresponding to the digit to be recalled.

In the spatial updating task, the encoding phase consisted of memorizing the positions of three to five dots presented in a 4 x 4 matrix. The matrix had a length of 6.2° visual angle and each dot had a length of 1.24° visual angle. In each updating step, a new position of the to-be-updated dot was indicated by an arrow pointing in the direction of the required mental shift of that dot. Thus, one of the colored dots was presented centrally below a black arrow pointing either left, right, up, or down (with a width of 1.24° visual angle and a height of 0.48° visual angle). After each updating step, the most recent position of the dot to be updated had to be recalled. Thus, an empty 4 x 4 matrix is presented.

Questionnaires

In the present study, five questionnaires were used. The first questionnaire was the German version of the Morningness-Eveningness Questionnaire (D-MEQ, Griefahn et al.,

2001) to assess the chronotype of each participant. The second questionnaire was implemented to have some descriptive information about our sample. Thus, this questionnaire assessed socio-demographic variables, such as age, gender, handedness, color blindness, nationality, native language, foreign language(s), number of education years, socio-economic status, synesthetic experience, and leisure activities (e.g., music, sport, video game) as well as information about the general health status and the preferences for dorsal/ventral processing types. The third questionnaire was used to have some information about the current health status of our sample. The questions were about medication use and sleep in the last 24 hours as well as nicotine, alcohol and drug consumption in the last 2 hours. Because of the known effects of depression and stress on cognition (e.g., McDermott & Ebmeier, 2009; Rock et al., 2014; Saenger et al., 2014; Starcke et al., 2016), the last two questionnaires were the German Version of the Beck Depression Inventory II (BDI-II, Hautzinger et al., 2006) and the German version of the Perceived Stress Questionnaire (PSQ, Fliege et al., 2005).

Procedure

Participants were tested remotely by means of a browser-based online experiment during three sessions. Participants were alone for all three sessions, but they could phone the student who recruited them in case of problems or questions. At the beginning of all three sessions, participants were required to confirm that they performed the experiment in a laboratory-like setting (e.g. they were alone in a quiet room, without distraction, and they closed all computer programs, but the browser window with the tasks). At the beginning of the second and third sessions, participants were asked to perform a scaling task. This task was implemented to ensure that all stimuli were presented with the same size because all tasks were run on participants' computers with different screen sizes and resolutions. In this scaling task, participants had to adapt the size of a rectangle to the size of a credit card. Then, based on this result, the size of the stimuli was computed so that it was the same across the different screen sizes and resolutions. In the middle of each session of the three sessions, participants could take a break of about 10 minutes. At the end of each session, they were asked about their current health status.

The three sessions were so organized that the first session was a screening session lasting for approximately 30 minutes. At the beginning of this session, after being informed, participants confirmed their consent to participate. Then, they were asked to perform the German version of the Morningness-Eveningness Questionnaire (D-MEQ, Griefahn et al., 2001) in addition to a general socio-demographic questionnaire. The following two sessions lasted approximately for 1.5 hour each. These two sessions were separated by at least one night and maximally 11 days. They were performed either at 08:00 or at 17:00. Based on the D-MEQ scores, half of the participants were tested on peak in the second session and off peak in the third session, whereas the other half was tested in the reversed order (i.e., off peak in the second session and on peak in the third session). During the on- and off-peak sessions, participants performed all short-term memory and working-memory tasks. At the end of the third session, participants were asked to complete the German version of the Beck Depression Inventory II (BDI-II, Hautzinger et al., 2006) and of the Perceived Stress Questionnaire (PSQ, Fliege et al., 2005).

Across the two on- and off- peaks sessions, the same order of short-term memory and working-memory tasks was used. For half of the participants, the following task order was used: verbal simple span, spatial complex span, numerical simple span, spatial updating, numerical complex span, arrow simple span, numerical updating, and spatial simple span. This task order was reversed for the other half of participants to control for practice effects. Moreover, in each session, one version of the tasks - that is, either Version 1 or Version 2 - was used. These versions and their order were counterbalanced across the two sessions so that half of the participants started with Version 1 and the other half with Version 2.

The task structure was similar across the different short-term memory and working-memory tasks. That is, each task started with the presentation of instructions explaining how participants had to carry out the task. These instructions were followed by a practice block, which could be repeated in case the participants required it. Following the practice block, participants performed one experimental block for the short-term memory tasks and two experimental blocks for the working-memory tasks. For the short-term memory tasks, the practice block included three trials with a set size of two, and the experimental block included three trials of each set size. In this experimental block, the set size ranged from three to nine for the digit simple span, three to eight for the letter simple span, two to seven for the matrix simple span, and two to six for the arrow simple span. For both numerical and spatial complex spans as well as for the spatial updating task, the practice block included two trials, and the two experimental blocks included a total of 12 trials. For the numerical updating task, the practice block included three trials (with four, six, and seven updating steps, respectively). The two experimental blocks included a total of 25 trials (each trial including seven updating steps). In addition, in this task, recall was probed in five out of the 25 trials immediately after the initial encoding. This was implemented to ensure that the initial set of memoranda was encoded. In all tasks, participants could take brief rests after each block.

The trial sequence was similar across all short-term memory tasks. That is, each trial started with the prompt “Ready?” until the participant pressed the space key. Then, the memorandum was presented for 1000 ms in the digit, letter, and arrow simple spans, and for 650 ms in the matrix simple span. The presentation of each memorandum was followed by a blank screen for 500 ms. This sequence was repeated, depending on the set size. After all memoranda were presented, participants were required to recall the sequence of memoranda in correct serial order. In the digit simple span, participants were asked to enter their response with the keyboard and then to press “Next” on the screen using the mouse. In the letter, matrix, and arrow simple spans, they were asked to select their responses by clicking on the screen. During the practice block only, the feedback about the accuracy of the given response was presented for 500 ms. At the end of each trial, a blank screen was displayed for 500 ms.

Except for the prompt "Ready?" and the feedback during the practice block, the trial sequence was more diverse for the working-memory tasks. In the numerical and spatial complex spans, each memorandum was presented for 1000 ms, and each distractor task was presented for 3000 ms maximally. During the distractor task, the stimulus-response mapping was additionally presented but in the practice block only. This encoding sequence was repeated, depending on the set size. After all memoranda were presented, participants were asked to recall memoranda in correct serial order. In the numerical complex span, participants were asked to enter their response with the keyboard and then to press "Next" on the screen using the mouse. In the spatial complex span, they were asked to select their responses by clicking on the screen. In both complex spans, at the end of the trial, a blank screen was displayed for 500 ms. In the numerical updating task, the encoding phase lasted for 5000 ms, followed by a blank screen for 250 ms. Each updating step was then displayed for 1250 ms, followed by a blank screen for 250 ms. The recall lasted until the participant responded by entering the digit with the keyboard. In the spatial updating task, memoranda were simultaneously presented for 500 ms per colored dot (e.g., four memoranda were presented for 2000 ms). Each updating step lasted for 500 ms. After each updating step, participants were asked to recall the most recent position of the dot to be updated. To this end, the matrix was displayed again until the participant clicked on it to give a response.

Data preparation

For all short-term memory and working-memory tasks, the dependent measure consisted of the accuracy rates computed as the proportion of memoranda recalled at the correct position (partial-credit unit score, see Conway et al., 2005). Mean accuracy rates were then computed for each participant, each task, and each session (on peak vs. off peak). For the numerical updating task, performance on the immediate probes were not included in the computation of the dependent measure. Standardized questionnaires were analyzed following their manual.

Data analysis

We used an alpha level of .05 for all tests from the null hypothesis significance testing (NHST) framework. Effect sizes were calculated as Cohen's *d*. For the Bayesian hypothesis testing, default prior scales were used. Moreover, the Bayes Factors (BFs) were interpreted using Raftery's (1995) classification scheme. According to this classification, a BF between 1-3 is considered as weak evidence, a BF between 3-20 is considered as positive evidence, a BF between 20-150 is considered as strong evidence, and a BF larger than 150 is considered as very strong evidence.

Model estimation

Model fit was evaluated via multiple fit indices (Hu & Bentler, 1999, 1998): the χ^2 goodness-of-fit statistic, the Bentler's comparative fit index (CFI), the root mean square error of approximation (RMSEA), and the standardized root-mean-square residual (SRMR). For the χ^2 statistic, a small, non-significant value indicates good fit. For the CFI, values larger than .95 indicate good fit, and values between .90 and .95 indicate acceptable fit. RMSEA values smaller than .06 and SRMR values smaller than .08 indicate good fit. In addition, the following criteria had to be met for a model to be considered a "good" fitting model: (1) the Kaiser-Meyer-Olkin (KMO) index – a measure of whether the correlation matrix is factorable – should be larger than .60 (Tabachnick & Fidell, 2019); (2) most of the error variances had to be lower than .90; (3) most of the factor loadings had to be significant and larger than .30; (4) no factor should be dominated by a large loading from one task; (5) The quality of how well the factor was represented by a set of measures – a measure sometimes called construct reliability or replicability – had to be good. This was assessed by the index H and it had to meet the standard criterion of .70 (Rodriguez et al., 2016).

Results

Results are reported in three steps. First, we investigated the reliability estimates and the correlational pattern for all measures. Second, we replicated previous research by examining the synchrony effect at the individual-task level (see, e.g., Ceglarek et al., 2021; Lewandowska et al., 2018; Rowe et al., 2009; Schmidt et al., 2015; West et al., 2002) . Third, we used structural equation modeling (SEM) to investigate the synchrony effect at the latent-variable level for the constructs of short-term maintenance and attentional control. In this step, we aimed to model these constructs for on and off peaks separately and to estimate a latent-change model between both peaks. This assessment is equivalent to a paired *t*-test which is applied to latent constructs (see Kievit et al., 2018).

Reliability and correlations

As shown in Table 3, all measures had acceptable skew and kurtosis (i.e., between -1.01 and 0.50). The reliability estimates for all measures in both sessions were good, ranging from .77 to .95.

Pearson correlation coefficients as well as their upper and lower confidence intervals are shown in Table 4. Bayes factors (BFs) for the correlations are presented in Table 5. These assessed the weight of evidence in favor of the alternative hypothesis (BF_{10}, i.e., in favor of a correlation) and in favor of the null hypothesis (BF_{01}, i.e., in favor of the absence of the correlation). The correlations were moderate to strong, ranging from .24 to .73. All correlations were significant (ps $\leq .001$), and all Bayes Factors (BFs) suggested strong to very strong evidence for the correlations (all $BFs_{10} \geq 38.92$).

Results at the individual-task level: Synchrony effect in short-term memory and working-memory tasks

Consistent with previous research (e.g., Ceglarek et al., 2021; Lewandowska et al., 2018; Rowe et al., 2009; Schmidt et al., 2015; West et al., 2002), we then investigated the synchrony effect at the individual-task level. This is displayed in Figure 1. Statistically, we compared short-term memory and working-memory performance between on- and off-peak sessions by computing a paired two-tailed *t*-test for each short-term memory and working-memory task separately. In addition, we computed Bayesian *t*-tests. This allowed us to assess not only the strength of evidence for the alternative hypothesis (i.e., the presence of the synchrony effect) but also the strength of evidence for the null hypothesis (i.e., the absence of a synchrony effect). These results are presented in Table 6. As shown in this table, no synchrony effect was observed for most tasks. Only in the numerical updating task and in the arrow simple span, a significant synchrony effet was found. However, the effect sizes were small for these tasks, and Bayesian evidence only suggests weak to positive evidence in favor of the effect (see Table 6).

Results at the latent-variable level: Synchrony effect in short-term maintenance and attentional control

The synchrony effect in short-term maintenance and attentional control was assessed by modeling a latent change between both on and off peaks for these constructs. A bifactor model was used to model short-term maintenance and attentional control. Overall, four latent-change models were thus modeled (see Figures 2 and 3). In each model, we took into account measure-specific variance across both sessions by allowing error variances from the measures of the same task to correlate. Factor loadings were also constrained to be positive. Moreover, the measurement invariance across time was modeled by applying equality constraints over off and on peaks for the factor loadings, the error variances, and the intercepts. Furthermore, we assessed the synchrony effect for each latent construct by modeling a latent change between the on-peak factor and the off-peak factor of the construct. That is, the on-peak factor was regressed on the off-peak factor by fixing the unstandardized regression weight to 1. The latent-change factor Δ was then measured by fixing the unstandardized factor loading to the on-peak factor to 1. A covariance was also added between the off-peak factor and the latent-change factor Δ to capture whether the latent change was dependent or proportional to the scores at off peak.

Following this strategy, we estimated the first model (Model 1) by modeling short-term maintenance as the common variance between short-term memory and working-memory measures, and attentional control as the working-memory variance remaining after controlling for short-term memory (see Engle et al., 1999). More precisely, all short-term memory and working-memory measures at off and on peaks were forced to load on an off-peak maintenance factor and an on-peak maintenance factor, respectively. In addition, all working-memory tasks at off and on peaks were forced to load on an off-peak attentional-control factor and an on-peak attentional-control, respectively. We assessed the synchrony effect for maintenance by modeling a latent change between the on-peak maintenance factor and the off-peak maintenance factor. Similarly, the synchrony effect for attentional control was assessed by modeling a latent change between the on-peak attentional-control factor and the off-peak attentional-control factor. This model is depicted in Figure 2a. However, it provided a bad fit to the data, KMO = .86, $\chi^2(97, N = 191) = 266.44$, $p < .001$, CFI = .91, RMSEA [90% CI] = .10 [.08, .11], SRMR = .09.

Short-term memory and working-memory tasks were reported to involve different maintenance processes for verbal-numerical and spatial materials, but common attentional-control processes (see, e.g., Kane et al., 2004). Accordingly, we fitted a second model (Model 2), which was similar to Model 1, except that the off- and on-peak maintenance factors were modeled separately for both material types. This model is depicted in Figure 2b. It provided a good fit to the data, KMO = .86, $\chi^2(87, N = 191) = 133.68$, $p = .001$, CFI = .97, RMSEA [90% CI] = .05 [.03, .07], SRMR = .07. However, the model parameters were not fully identified as the covariance matrix was not positive definite.

Applying a model-reduction strategy, we fitted a third model (Model 3) in which we assumed no synchrony effect for short-term maintenance. The synchrony effect was modeled only for attentional control. Thus, this model was similar to Model 2, except that there was one maintenance factor for verbal-numerical material and one maintenance factor for spatial material across both on- and off-peak sessions. This model is depicted in Figure 3a. It provided an acceptable fit to the data, KMO = .86, $\chi^2(101, N = 191) = 210.16$, $p < .001$, CFI = .94, RMSEA [90% CI] = .08 [.06, .09], SRMR = .08. Inspection of the key parameter for the latent change shows no difference between off and on peaks, thus indicating no synchrony effect for attentional control (the unstandardized intercept of the latent-change factor = .01, se = .01, $p = .493$, 95% CI = [-.01, .03]). In this model, however, the construct reliability for the attentional-control factors was low ($H = .42$ and $H = .39$ for off and on peak, respectively). This indicates that the attentional-control factors did not represent much common variance across the measures. Therefore, this model has low explanatory power.

In the last model (Model 4), we followed previous work in which attentional control has been modeled as a common factor across all short-term memory and working-memory measures (see Kane et al., 2004). Thus, short-term memory and working-memory tasks at off and on peaks were forced to load on an off-peak general factor and an on-peak general factor, respectively. In addition, short-term memory and working-memory tasks with verbal-numerical material were forced to load on a verbal-numerical maintenance factor, whereas those tasks with spatial material were forced to load on a spatial maintenance factor. Thus, similar to Model 3, the synchrony effect was modeled only for attentional control. This model is depicted in Figure 3b. It provided a good fit to the data, KMO = .86, $\chi^2(94, N = 191) = 110.49$, $p = .118$, CFI = .99, RMSEA [90% CI] = .03 [.00, .05], SRMR = .05. In this case, the construct reliability for all factors was good (all $Hs \geq .73$).

Inspection of the key parameter for the latent change shows no difference between off and on peaks, thus indicating no synchrony effect for attentional control (the unstandardized intercept of the latent-change factor = .001, se = .01, p = .937, 95% CI = [-.01, .01]).

Multiverse-analysis approach

In a final step, we tested for the robustness of the results in a multiverse-analysis approach. To this end, we re-ran the analyses by applying different data transformations, participants' selections, trimming procedures, and SEM approaches. The different procedures we used are described in Table 7. Please note that as preregistered, we modeled the data using a second-order latent growth curve modeling approach. However, these model assessments never resulted in acceptable or good fit statistics. Therefore, we opted for the latent-change models, which is an extension of the latent growth curve models focusing more directly on the difference between the two measurement time points (Ghisletta & McArdle, 2012).

An overview of the results from the multiverse-analysis approach is presented in Figure 4. As shown in the upper part of the figure, no robust synchrony effect was observed at the individual-task level. At the latent-variable level, the results showed the difficulty of measuring short-term maintenance and attentional control as good factors. When these were modeled satisfactorily (i.e., with acceptable to good fit statistics and good indices H), the synchrony effect was in most cases not significant. In the two cases showing a significant effect, the latent change was small (maximum value for the unstandardized intercept of the latent-change factor = 0.02).

Finally, for the sake of completeness and following previous work (e.g., Allen et al., 2008; Bonnefond et al., 2003; Fabbri et al., 2008; Lewandowska et al., 2018; Matchock & Mordkoff, 2009), we also investigated the time-of-day effect, that is, the difference in performance between the morning and evening sessions (but independently of the peak of circadian arousal). The results are summarized in the Appendix. Consistent with the results about the synchrony effect, the results showed no robust time-of-day effect either at the individual-task level or at the latent-variable level.

Discussion

The effect of circadian arousal on human cognition is well-established as common knowledge, but its empirical evidence is more equivocal. In the present study, we aimed to empirically clarify this effect. Specifically, we determined the scope and robustness of the synchrony effect on human cognition by focusing on universal psychological constructs and by addressing the methodical challenges typically observed in previous research. Thus, we investigated the synchrony effect on working memory and attentional control in a large sample of participants, who were tested at their optimal time of day and their non-optimal time of day. All participants performed four short-term memory tasks and four working-memory tasks. Attentional control was assessed on the latent-variable level as the goal-directed nature of working-memory tasks without their maintenance aspects. In all tasks, we observed performance similar to previous research (e.g., in the mean, reliability and correlation estimates; cf. Kane et al., 2004; Rey-Mermet et al., 2019). Quite surprisingly, however, the results showed no systematic and robust synchrony effect for any of the constructs we measured (i.e., short-term memory, working memory, and attentional control). Moreover, this pattern of results was confirmed when we applied different data transformations, participants' selection criteria, and trimming procedures.

On the individual-task level, the results showed no synchrony effect for most tasks. A synchrony effect was only detected in one out of four short-term memory tasks (i.e., arrow simple span) and one out of four working-memory tasks (i.e., numerical updating task). However, their effect sizes were at best small. From a positive perspective, our results can be interpreted that the synchrony effect does exist and that it would emerge in about 25% of cognitive tasks (i.e., 2/8 of our tasks). From this perspective and based on the assumption that a large portion of null-findings are not published (e.g., Kühberger et al., 2014), our findings would be broadly consistent with previous research showing mixed evidence with regards to the synchrony effect on short-term memory and working memory (e.g., Rowe et al., 2009; Schmidt et al., 2015; West et al., 2002; but Ceglarek et al., 2021; Heimola et al., 2021; Lewandowska et al., 2018). This interpretation of the findings would be valid if the circumstances under which a synchrony effect occurs and under which it does not occur were clear. For example, the circumstances would have been clear if the synchrony effect was observed in one task type (e.g., working-memory tasks) or – although theoretically less plausible – only in tasks with the same stimulus materials (e.g., the digit simple span in short-term memory and the numerical complex span in working memory). However, we observed a synchrony effect in two tasks (arrow simple span and numerical updating) with different types (short-term memory and working memory, respectively) and different materials (spatial and numerical, respectively). Therefore, these results indicate no systematic pattern. Given that, our positive findings with regards to a synchrony effect on cognition is most likely coincidental. Therefore, the synchrony effect at the individual-task level is rather the exception than the norm.

On the latent-variable level, we were also not able to observe any evidence for a synchrony effect on cognition. Specifically, we used structural equation modeling to assess short-term maintenance and attentional control separately for on- and off-peak times, and we estimated the synchrony effect as a latent change between both peaks. Across the different analyses we run, we found no systematic synchrony effect for short-term maintenance or attentional control. The reasons were multiple: (1) The models did not provide good fit statistics or were not fully identified. (2) The factors did not represent shared variances across the measures. (3) The latent synchrony effect was not significant. There were only two cases with a specific combination of data transformation, participants' selection and trimming procedure where the model provided good fit statistics, substantial common variance, and a significant – but small – latent synchrony effect. However, consistent with the results at the individual-task level, these two cases are the exception and represent most likely coincidental findings. Therefore, the systematic picture of the SEM approach unequivocally reveals strong evidence against a synchrony effect on cognition.

Overall, the present results are mostly in line with the findings challenging the synchrony effect on cognition (e.g., Ceglarek et al., 2021; Heimola et al., 2021, 2021; Knight & Mather, 2013; Lewandowska et al., 2018; Li et al., 1998; Matchock & Mordkoff, 2009). Since less than 5% of our participants qualified themselves as definite morning or evening types (i.e., the most extreme categories from the chronotype distribution), we were not able to perform all analyses to investigate whether there is a reliable synchrony effect on cognition in these extreme chronotypes. Please also note that while we found no evidence for a synchrony effect on cognition more generally, this does not necessarily preclude that off-peak working schedules have no impact on performance and well-being. For instance, some compensation might be possible on a daily basis for traditional working schedules. By contrast, prolonged periods of off-peak schedules might be harmful to health and performance in shift-workers with more extreme chronotypes.

The results of the present study also convey an important message which applies to psychological science and scientific research more widely. Namely, reliable and robust insights into cognitive processes result from well-powered experimental designs in which several tasks measure the same underlying constructs. Had we conducted our study with only one of the two tasks which revealed positive findings, we would have come to the wrong conclusion that the synchrony effect does exist. Moreover, given our large sample, the within-subject design and the reliability of our measures, we would have interpreted the findings as very robust and representative. As our results show no systematic synchrony effect across the different tasks and the different levels of analyses (individual-task vs. latent-variable level), such an interpretation of our results is not warranted. Given the overall picture of our results, we must conclude that the impact of circadian arousal on cognitive performance is not as general and robust as previously thought.

References

Aben, B., Stapert, S., & Blokland, A. (2012). About the Distinction between Working Memory and Short-Term Memory. *Frontiers in Psychology*, *3*. https://www.frontiersin.org/articles/10.3389/fpsyg.2012.00301

Allen, P. A., Grabbe, J., McCarthy, A., Bush, A. H., & Wallace, B. (2008). The Early Bird Does Not Get the Worm: Time-of-Day Effects on College Students' Basic Cognitive Processing. *The American Journal of Psychology*, *121*(4), 551–564. https://doi.org/10.2307/20445486

Aust, F., & Barth, M. (2020). *papaja: Prepare reproducible APA journal articles with R Markdown*. https://github.com/crsh/papaja

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Bennett, C. L., Petros, T. V., Johnson, M., & Ferraro, F. R. (2008). Individual Differences in the Influence of Time of Day on Executive Functions. *The American Journal of Psychology*, *121*(3), 349–361. https://doi.org/10.2307/20445471

Bonnefond, A., Gisselbrecht, D., Hoeft, A., Eschenlauer, R., Muzet, A., & Tassi, P. (2003). Cognitive performance in middle-aged adults as a function of time of day and task load. *Neurobiology of Sleep-Wakefulness Cycle*, *3*, 1–8.

Brysbaert, M. (2019). How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *Journal of Cognition*, *2*, 16. https://doi.org/10.5334/joc.72

Ceglarek, A., Hubalewska-Mazgaj, M., Lewandowska, K., Sikora-Wachowicz, B., Marek, T., & Fafrowicz, M. (2021). Time-of-day effects on objective and subjective short-term memory task performance. *Chronobiology International*, *38*(9), 1330–1343. https://doi.org/10.1080/07420528.2021.1929279

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R.

W. (2005). Working memory span tasks: A methodological review and userâ€™s guide. *Psychonomic Bulletin & Review*, *12*(5), 769–786. https://doi.org/10.3758/BF03196772

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.20982/tqmp.01.1.p042

Draheim, C., Pak, R., Draheim, A. A., & Engle, R. W. (2022). The role of attention control in complex real-world tasks. *Psychonomic Bulletin & Review*, *29*(4), 1143–1197. https://doi.org/10.3758/s13423-021-02052-2

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309–331. https://doi.org/10.1037/0096-3445.128.3.309

Fabbri, M., Natale, V., & Adan, A. (2008). Effect of time of day on arithmetic fact retrieval in a number-matching task. *Acta Psychologica*, *127*(2), 485–490. https://doi.org/10.1016/j.actpsy.2007.08.011

Fliege, H., Rose, M., Arck, P., Walter, O. B., Kocalevent, R.-D., Weber, C., & Klapp, B. F. (2005). The Perceived Stress Questionnaire (PSQ) Reconsidered: Validation and Reference Values From Different Clinical and Healthy Adult Samples. *Psychosomatic Medicine*, *67*(1), 78–88. https://doi.org/10.1097/01.psy.0000151491.80178.78

Ghisletta, P., & McArdle, J. J. (2012). Latent Curve Models and Latent Change Score Models Estimated in R. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(4), 651–682. https://doi.org/10.1080/10705511.2012.713275

Griefahn, B., Künemund, C., Bröde, P., & Mehnert, P. (2001). Zur Validität der deutschen Ãœbersetzung des Morningness-Eveningness-Questionnaires von Horne und Östberg. *Somnologie*, *5*(2), 71–80. https://doi.org/10.1046/j.1439-054X.2001.01149.x

Hahn, C., Cowell, J. M., Wiprzycka, U. J., Goldstein, D., Ralph, M., Hasher, L., & Zelazo,

P. D. (2012). Circadian rhythms in executive function during the transition to adolescence: The effect of synchrony between chronotype and time of day. *Developmental Science*, *15*(3), 408–416. https://doi.org/10.1111/j.1467-7687.2012.01137.x

Hasher, L., Chung, C., May, C. P., & Foong, N. (2002). Age, time of testing, and proactive interference. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *56*(3), 200–207. https://doi.org/10.1037/h0087397

Hautzinger, M., Keller, F., & Kühner, C. (2006). *Beck Depression Inventar II (BDI 2)*. Harcourt Test Service.

Heimola, M., Paulanto, K., Alakuijala, A., Tuisku, K., Simola, P., Ämmälä, A.-J., Räisänen, P., Parkkola, K., & Paunio, T. (2021). Chronotype as self-regulation: Morning preference is associated with better working memory strategy independent of sleep. *SLEEP Advances*, *2*(1), zpab016. https://doi.org/10.1093/sleepadvances/zpab016

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2019). Lab.js: A free, open, online study builder. *PsyArXiv*. https://doi.org/10.31234/osf.io/fqr49

Horne, J. A., & Ostberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology*, *4*(2), 97–110. https://doi.org/10.3109/07420529709001458

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424–453. http://psycnet.apa.org/journals/met/3/4/424/

Intons-Peterson, M. J., Rocchi, P., West, T., McLellan, K., & Hackney, A. (1998). Aging, Optimal Testing Times, and Negative Priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 362–376.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2021). *semTools: Useful tools for structural equation modeling.* https://CRAN.R-project.org/package=semTools

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The Generality of Working Memory Capacity: A Latent-Variable Approach to Verbal and Visuospatial Memory Span and Reasoning. *Journal of Experimental Psychology: General*, *133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189

Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The Unity and Diversity of Executive Functions: A Systematic Review and Re-Analysis of Latent Variable Studies. *Psychological Bulletin*, *144*(11), 1147–1185. https://doi.org/10.1037/bul0000160

Kievit, R. A., Brandmaier, A. M., Ziegler, G., van Harmelen, A.-L., de Mooij, S. M. M., Moutoussis, M., Goodyer, I. M., Bullmore, E., Jones, P. B., Fonagy, P., Lindenberger, U., & Dolan, R. J. (2018). Developmental cognitive neuroscience using latent change score models: A tutorial and applications. *Developmental Cognitive Neuroscience*, *33*, 99–117. https://doi.org/10.1016/j.dcn.2017.11.007

Knight, M., & Mather, M. (2013). Look Out—It's Your Off-Peak Time of Day! Time of Day Matters More for Alerting than for Orienting or Executive Attention. *Experimental Aging Research*, *39*(3), 305–321. https://doi.org/10.1080/0361073X.2013.779197

Kühberger, A., Fritz, A., & Scherndl, T. (2014). Publication Bias in Psychology: A Diagnosis Based on the Correlation between Effect Size and Sample Size. *PLOS ONE*, *9*(9), e105825. https://doi.org/10.1371/journal.pone.0105825

Lara, T., Madrid, J. A., & Correa, Á. (2014). The Vigilance Decrement in Executive

Function Is Attenuated When Individual Chronotypes Perform at Their Optimal Time of Day. *PLOS ONE*, *9*(2), e88820. https://doi.org/10.1371/journal.pone.0088820

Lewandowska, K., Wachowicz, B., Marek, T., Oginska, H., & Fafrowicz, M. (2018). Would you say "yes" in the evening? Time-of-day effect on response bias in four types of working memory recognition tasks. *Chronobiology International*, *35*(1), 80–89. https://doi.org/10.1080/07420528.2017.1386666

Li, K. Z. H., Hasher, L., Jonas, D., Rahhal, T. A., & May, C. P. (1998). Distractibility, circadian arousal, and aging: A boundary condition? *Psychology and Aging*, *13*(4), 574–583. https://doi.org/10.1037/0882-7974.13.4.574

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519–530. https://doi.org/10.1093/biomet/57.3.519

Matchock, R. L., & Mordkoff, T. J. (2009). Chronotype and time-of-day influences on the alerting, orienting, and executive components of attention. *Experimental Brain Research*, *192*(2), 189–198. https://doi.org/10.1007/s00221-008-1567-6

May, C. P. (1999). Synchrony effects in cognition: The costs and a benefit. *Psychonomic Bulletin & Review*, *6*(1), 142–147.

May, C. P., & Hasher, L. (1998). Synchrony Effects in Inhibitory Control Over Thought and Action. *Journal of Experimental Psychology*, *24*(2), 363–379.

May, C. P., Hasher, L., & Foong, N. (2005). Implicit memory, age, and time of day. *Psychological Science*, *16*(2), 96–100. https://doi.org/10.1111/j.0956-7976.2005.00788.x

May, C. P., Hasher, L., & Stoltzfus, E. R. (1993). Optimal Time of Day and the Magnitude of Age Differences in Memory. *Psychological Science*, *4*(5), 326–330. https://doi.org/10.1111/j.1467-9280.1993.tb00573.x

McDermott, L. M., & Ebmeier, K. P. (2009). A meta-analysis of depression severity and cognitive function. *Journal of Affective Disorders*, *119*(1), 1–8.

https://doi.org/10.1016/j.jad.2009.04.022

Miyake, A., & Friedman, N. P. (2012). The Nature and Organization of Individual Differences in Executive Functions: Four General Conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. https://doi.org/10.1177/0963721411429458

Morey, R. D. (2008). Confidence Intervals from Normalized Data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Morey, R. D., & Rouder, J. N. (2021). *BayesFactor: Computation of bayes factors for common designs*. https://CRAN.R-project.org/package=BayesFactor

Murray, G., Nicholas, C. L., Kleiman, J., Dwyer, R., Carrington, M. J., Allen, N. B., & Trinder, J. (2009). Nature's Clocks and Human Mood: The Circadian System Modulates Reward Motivation. *Emotion*, *9*(5), 705–716. https://doi.org/10.1037/a0017080

Parsons, S. (2020). *Splithalf; robust estimates of split half reliability*. https://doi.org/10.6084/m9.figshare.5559175.v5

Parsons, S., Kruijt, A.-W., & Fox, E. (2019). Psychological Science Needs a Standard Practice of Reporting the Reliability of Cognitive-Behavioral Measurements. *Advances in Methods and Practices in Psychological Science*, *2*(4), 378–395. https://doi.org/10.1177/2515245919879695

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, *25*, 111–163. https://doi.org/10.2307/271063

Revelle, W. (2021). *Psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://CRAN.R-project.org/package=psych

Rey-Mermet, A., Gade, M., & Oberauer, K. (2018). Should we stop thinking about inhibition? Searching for individual and age differences in inhibition ability. *Journal of*

Experimental Psychology: Learning, Memory, and Cognition, *44*(4), 501–526. https://doi.org/10.1037/xlm0000450

Rey-Mermet, A., Gade, M., Souza, A. S., von Bastian, C. C., & Oberauer, K. (2019). Is executive control related to working memory capacity and fluid intelligence? *Journal of Experimental Psychology: General*, *148*(8), 1335–1372. https://doi.org/10.1037/xge0000593

Rey-Mermet, A., Singh, K. A., Gignac, G. E., Brydges, C. R., & Ecker, U. K. H. (2020). Interference control in working memory: Evidence for discriminant validity between removal and inhibition tasks. *PLOS ONE*, *15*(12), e0243053. https://doi.org/10.1371/journal.pone.0243053

Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis. *Psychological Medicine*, *44*(10), 2029–2040. https://doi.org/10.1017/S0033291713002535

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods*, *21*(2), 137–150. https://doi.org/10.1037/met0000045

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rowe, G., Hasher, L., & Turcotte, J. (2009). Short article: Age and synchrony effects in visuospatial working memory. *Quarterly Journal of Experimental Psychology*, *62*(10), 1873–1880. https://doi.org/10.1080/17470210902834852

Saenger, J., Bechtold, L., Schoofs, D., Blaszkewicz, M., & Wascher, E. (2014). The influence of acute stress on attention mechanisms and its electrophysiological correlates. *Frontiers in Behavioral Neuroscience*, *8*. https://www.frontiersin.org/article/10.3389/fnbeh.2014.00353

Schmidt, C., Collette, F., Cajochen, C., & Peigneux, P. (2007). A time to think: Circadian rhythms in human cognition. *Cognitive Neuropsychology*, *24*(7), 755–789.

https://doi.org/10.1080/02643290701754158

Schmidt, C., Collette, F., Reichert, C. F., Maire, M., Vandewalle, G., Peigneux, P., & Cajochen, C. (2015). Pushing the Limits: Chronotype and Time of Day Modulate Working Memory-Dependent Cerebral Activity. *Frontiers in Neurology*, *6*. https://doi.org/10.3389/fneur.2015.00199

Schmidt, C., Peigneux, P., Leclercq, Y., Sterpenich, V., Vandewalle, G., Phillips, C., Berthomier, P., Berthomier, C., Tinguely, G., Gais, S., Schabus, M., Desseilles, M., Dang-Vu, T., Salmon, E., Degueldre, C., Balteau, E., Luxen, A., Cajochen, C., Maquet, P., & Collette, F. (2012). Circadian Preference Modulates the Neural Substrate of Conflict Processing across the Day. *PLoS ONE*, *7*(1), e29658. https://doi.org/10.1371/journal.pone.0029658

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *Dialogue: The Official Newsletter of the Society for Personality and Social Psychology*, *26*, 4–7.

Singmann, H., Bolker, B., Westfall, J., Aust, F., & Ben-Shachar, M. S. (2021). *Afex: Analysis of factorial experiments*. https://CRAN.R-project.org/package=afex

Soper, D. S. (2018). *A-priori Sample Size Calculator for Structural Equation Models [Software]*. http://www.danielsoper.com/statcalc

Starcke, K., Wiesen, C., Trotzke, P., & Brand, M. (2016). Effects of Acute Laboratory Stress on Executive Functions. *Frontiers in Psychology*, *7*. https://www.frontiersin.org/article/10.3389/fpsyg.2016.00461

Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed.). Pearson.

von Bastian, C. C., Blais, C., Brewer, G. A., Gyurkovics, M., Hedge, C., Kałamała, P., Meier, M. E., Oberauer, K., Rey-Mermet, A., Rouder, J. N., Souza, A. S., Bartsch, L. M., Conway, A. R. A., Draheim, C., Engle, R. W., Friedman, N. P., Frischkorn, G. T., Gustavson, D. E., Koch, I., … Wiemers, E. A. (2020). *Advancing the understanding of individual differences in attentional control: Theoretical, methodological, and analytical considerations*. PsyArXiv. https://doi.org/10.31234/osf.io/x3b9k

West, R., Murphy, K. J., Armilio, M. L., Craik, F. I. M., & Stuss, D. T. (2002). Effects of Time of Day on Age Differences in Working Memory. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *57*(1), P3–P10. https://doi.org/10.1093/geronb/57.1.P3

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Table 1

Exclusion Criteria.

Reasons	Number of exclusions
Participants were not aged between 18 and 28.	9
Participants did not report Swiss German or German as native language.	5
Participants reported colorblindness or no normal vision.	0
Participants reported neurological or psychiatric disorders.	5
Participants did not complete the whole experiment.	46
The session order or the task order across the sessions was incorrect.	6
Participants took a rest longer than 30 minutes.	8
The data timestamp was not in chronological order.	2
Participants did not perform the morning session between 07:30 and 10:00 and the evening session between 16:30 and 19:00.	65
The morning or evening session lasted more than two hours.	2
A task was missing.[a]	69
Participants were multivariate outliers.[b]	26
Participants had a neutral chronotype.[c]	255

[a] A task was missing or discarded if the computer malfunctioned or if the participants did not respond within three minutes when performing the task. This was implemented to ensure laboratory-like setting. [b] We checked for multivariate normality across all measures using the Mardia's (1970) kurtosis index. Participants were considered as multivariate outliers when Mahalanobis's d2 values were significant. [c] Following Griefahn et al. (2001), neutral chronotype was indicated with a chronoscore ranging from 42 to 58.

Table 2

Sample Characteristics and Background Measures.

Measure	Sample
Sample size	191
Age, years	23.8 (3)
Age range	18-28
Gender (female / male / other)	135 / 54 / 2
Education, years	13.5 (3.8)
Education level[a]	4.9 (1.5)
BDI-II score[b]	6.9 (6.3)
PSQ score[c]	30 (15.9)
D-MEQ chronoscore[d] (morning / evening)	64.5 (4) / 36.5 (4.1)
D-MEQ chronoscore range (morning / evening)	59-78 / 25-41

Note. Standard deviations are given in parentheses. BDI-II = Beck Depression Inventory II (Hautzinger et al., 2006); PSQ = Perceived Stress Questionnaire (Fliege et al., 2005); D-MEQ = German version of the Morningness-Eveningness-Questionnaire (Griefahn et al., 2001).

[a] Education level ranged from 1 (no or less than 9 school years) to 8 (Ph.D.). [b] Depression score ranged from 0 (minimal depression) to 63 (severe depression). [c] PSQ score ranged from 0 to 100. [d] Chronoscore ranged from 16 (definite evening type) to 86 (definite morning type). Neutral types are indicated by a chronoscore ranging from 42 to 58.

Table 3

Synchrony Effect: Descriptive Statistics.

Session	Task	Mean	SD	Min.	Max.	Skew	Kurtosis	Reliability
Off peak	Digit simple span	.77	.11	.44	1	-0.36	-0.23	.89 [.87, .91]
	Letter simple span	.75	.12	.37	1	-0.46	0.37	.87 [.84, .89]
	Matrix simple span	.79	.12	.43	.99	-0.49	-0.32	.87 [.84, .89]
	Arrow simple span	.64	.12	.28	.98	-0.13	0.19	.77 [.72, .81]
	Numerical complex span	.44	.19	.02	.98	0.28	-0.01	.88 [.85, .90]
	Spatial complex span	.36	.20	.02	.83	0.48	-0.70	.92 [.91, .94]
	Numerical updating	.58	.23	.11	1	-0.05	-1.01	.95 [.94, .96]
	Spatial updating	.67	.16	.11	.98	-0.64	0.30	.93 [.91, .94]
On peak	Digit simple span	.77	.11	.46	.98	-0.32	-0.32	.89 [.87, .91]
	Letter simple span	.76	.12	.33	1	-0.40	-0.03	.88 [.86, .91]
	Matrix simple span	.80	.11	.46	1	-0.75	0.33	.84 [.81, .88]
	Arrow simple span	.66	.13	.32	1	-0.07	-0.20	.80 [.76, .84]
	Numerical complex span	.45	.18	.04	1	0.21	-0.20	.86 [.84, .89]
	Spatial complex span	.37	.20	.02	1	0.50	-0.39	.92 [.91, .94]
	Numerical updating	.61	.23	.10	1	-0.24	-0.89	.95 [.94, .96]
	Spatial updating	.68	.15	.22	1	-0.65	0.30	.92 [.90, .93]

Note. Short-term memory and working memory were measured using accuracy rates. Permutation-based split-half reliability estimates were computed (see Parsons et al., 2019). The split-half correlations were adjusted with the Spearman–Brown prophecy formula, and the results of 5000 random splits were averaged. The 95% confidence intervals are presented in brackets. SD = Standard Deviation; Min. = minimum; Max. = maximum.

Table 4

Synchrony Effect: Pearson Correlation Coefficients.

		Off peak								On peak						
Session	Task	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
Off peak	1. Digit s.s.	-														
	2. Letter s.s.	**.68***	-													
		[.60, .75]														
	3. Matrix s.s.	**.26***	**.28***	-												
		[.12, .39]	[.14, .40]													
	4. Arrow s.s.	**.37***	**.33***	**.59***	-											
		[.24, .48]	[.19, .45]	[.49, .68]												
	5. Numerical c.s.	**.54***	**.54***	**.34***	**.40***	-										
		[.43, .64]	[.43, .63]	[.20, .46]	[.27, .51]											
	6. Spatial c.s.	**.32***	**.35***	**.55***	**.40***	**.45***	-									
		[.18, .44]	[.22, .47]	[.44, .64]	[.27, .51]	[.33, .56]										
	7. Numerical upd.	**.34***	**.36***	**.39***	**.47***	**.40***	**.43***	-								
		[.21, .46]	[.23, .48]	[.27, .51]	[.35, .57]	[.28, .52]	[.31, .54]									
	8. Spatial upd.	**.30***	**.29***	**.56***	**.53***	**.34***	**.50***	**.50***	-							
		[.17, .43]	[.15, .41]	[.45, .65]	[.41, .62]	[.21, .46]	[.38, .60]	[.39, .60]								
On peak	9. Digit s.s.	**.67***	**.63***	**.32***	**.38***	**.44***	**.36***	**.28***	**.25***	-						
		[.59, .75]	[.53, .71]	[.19, .44]	[.25, .49]	[.32, .55]	[.23, .48]	[.15, .41]	[.12, .38]							
	10. Letter s.s.	**.60***	**.68***	**.27***	**.32***	**.49***	**.29***	**.25***	**.27***	**.66***	-					
		[.50, .68]	[.60, .75]	[.13, .39]	[.18, .44]	[.37, .59]	[.15, .41]	[.11, .38]	[.13, .40]	[.57, .73]						
	11. Matrix s.s.	**.28***	**.34***	**.70***	**.52***	**.25***	**.42***	**.25***	**.45***	**.38***	**.31***	-				
		[.14, .41]	[.21, .46]	[.62, .77]	[.41, .62]	[.11, .38]	[.30, .53]	[.12, .38]	[.33, .56]	[.26, .50]	[.18, .44]					
	12. Arrow s.s.	**.29***	**.26***	**.58***	**.70***	**.31***	**.38***	**.39***	**.49***	**.41***	**.32***	**.55***	-			
		[.15, .41]	[.12, .38]	[.47, .66]	[.61, .76]	[.18, .43]	[.25, .50]	[.26, .50]	[.37, .59]	[.28, .52]	[.18, .44]	[.44, .64]				
	13. Numerical c.s.	**.44***	**.53***	**.31***	**.34***	**.67***	**.35***	**.37***	**.24***	**.52***	**.52***	**.34***	**.42***	-		
		[.32, .55]	[.42, .63]	[.17, .43]	[.21, .46]	[.58, .74]	[.22, .47]	[.24, .49]	[.10, .37]	[.41, .62]	[.41, .62]	[.21, .46]	[.29, .53]			
	14. Spatial c.s.	**.26***	**.31***	**.51***	**.32***	**.32***	**.72***	**.29***	**.45***	**.42***	**.39***	**.46***	**.39***	**.43***	-	
		[.13, .39]	[.17, .43]	[.40, .61]	[.19, .44]	[.18, .44]	[.64, .78]	[.16, .42]	[.33, .56]	[.30, .53]	[.26, .50]	[.34, .56]	[.26, .50]	[.31, .54]		
	15. Numerical upd.	**.28***	**.35***	**.33***	**.34***	**.31***	**.40***	**.64***	**.41***	**.37***	**.33***	**.36***	**.42***	**.50***	**.44***	-
		[.14, .40]	[.22, .47]	[.20, .45]	[.21, .46]	[.18, .44]	[.27, .51]	[.54, .71]	[.29, .53]	[.24, .48]	[.20, .45]	[.23, .48]	[.29, .53]	[.39, .60]	[.32, .55]	
	16. Spatial upd.	**.27***	**.28***	**.48***	**.44***	**.28***	**.41***	**.33***	**.73***	**.30***	**.35***	**.51***	**.46***	**.29***	**.46***	**.42***
		[.14, .40]	[.15, .41]	[.36, .58]	[.32, .55]	[.14, .40]	[.28, .52]	[.20, .46]	[.65, .79]	[.17, .42]	[.22, .47]	[.39, .61]	[.34, .56]	[.15, .41]	[.34, .56]	[.30, .53]

Note. Ninety-five percent confidence intervals are presented in brackets. Correlations for which the Bayes factor suggested positive to strong evidence for the alternative hypothesis (BF_{10}) are presented in bold; correlations for which the Bayes factor suggested positive to strong evidence for the null hypothesis (BF_{01}) are presented in italics. S.s. = simple span; c.s. = complex span; upd. = updating.

* $p < .05$.

Table 5

Synchrony Effect: Bayes Factors in Favor of the Alternative Hypothesis (BF_{10}) and in Favor of the Null Hypothesis (BF_{01}) for the Pearson Correlation Coefficients.

			Off peak								On peak						
Session	Task	BF	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.
Off peak	1. Digit s.s.	BF_{10}	-														
		BF_{01}															
	2. Letter s.s.	BF_{10}	$1.60*10^{24}$	-													
		BF_{01}	$6.24*10^{-25}$														
	3. Matrix s.s.	BF_{10}	98.48	278.69	-												
		BF_{01}	0.01	$3.59*10^{-3}$													
	4. Arrow s.s.	BF_{10}	$8.46*10^{4}$	$5.32*10^{3}$	$2.27*10^{16}$	-											
		BF_{01}	$1.18*10^{-5}$	$1.88*10^{-4}$	$4.40*10^{-17}$												
	5. Numerical c.s.	BF_{10}	$1.34*10^{13}$	$6.32*10^{12}$	$9.83*10^{3}$	$1.19*10^{6}$	-										
		BF_{01}	$7.48*10^{-14}$	$1.58*10^{-13}$	$1.02*10^{-4}$	$8.37*10^{-7}$											
	6. Spatial c.s.	BF_{10}	$2.74*10^{3}$	$2.88*10^{4}$	$2.79*10^{13}$	$1.40*10^{6}$	$2.10*10^{8}$	-									
		BF_{01}	$3.64*10^{-4}$	$3.47*10^{-5}$	$3.58*10^{-14}$	$7.16*10^{-7}$	$4.76*10^{-9}$										
	7. Numerical upd.	BF_{10}	$1.65*10^{4}$	$5.62*10^{4}$	$8.10*10^{5}$	$1.21*10^{9}$	$2.06*10^{6}$	$2.60*10^{7}$	-								
		BF_{01}	$6.05*10^{-5}$	$1.78*10^{-5}$	$1.23*10^{-6}$	$8.29*10^{-10}$	$4.85*10^{-7}$	$3.84*10^{-8}$									
	8. Spatial upd.	BF_{10}	$1.08*10^{3}$	523.30	$7.90*10^{13}$	$1.28*10^{12}$	$1.24*10^{4}$	$3.99*10^{10}$	$4.95*10^{10}$	-							
		BF_{01}	$9.25*10^{-4}$	$1.91*10^{-3}$	$1.27*10^{-14}$	$7.83*10^{-13}$	$8.07*10^{-5}$	$2.51*10^{-11}$	$2.02*10^{-11}$								
On peak	9. Digit s.s.	BF_{10}	$2.44*10^{23}$	$1.45*10^{19}$	$3.27*10^{3}$	$1.89*10^{5}$	$1.01*10^{8}$	$5.62*10^{4}$	360.76	76.89	-						
		BF_{01}	$4.11*10^{-24}$	$6.91*10^{-20}$	$3.06*10^{-4}$	$5.28*10^{-6}$	$9.92*10^{-9}$	$1.78*10^{-5}$	$2.77*10^{-3}$	0.01							
	10. Letter s.s.	BF_{10}	$8.73*10^{16}$	$1.87*10^{24}$	135.18	$2.56*10^{3}$	$1.09*10^{10}$	510.81	68.38	188.91	$6.20*10^{21}$	-					
		BF_{01}	$1.15*10^{-17}$	$5.35*10^{-25}$	$7.40*10^{-3}$	$3.90*10^{-4}$	$9.15*10^{-11}$	$1.96*10^{-3}$	0.01	$5.29*10^{-3}$	$1.61*10^{-22}$						
	11. Matrix s.s.	BF_{10}	300.22	$1.31*10^{4}$	$2.42*10^{26}$	$1.02*10^{12}$	62.66	$1.09*10^{7}$	78.89	$1.45*10^{8}$	$3.68*10^{5}$	$2.28*10^{3}$	-				
		BF_{01}	$3.33*10^{-3}$	$7.61*10^{-5}$	$4.14*10^{-27}$	$9.81*10^{-13}$	0.02	$9.21*10^{-8}$	0.01	$6.88*10^{-9}$	$2.72*10^{-6}$	$4.38*10^{-4}$					
	12. Arrow s.s.	BF_{10}	456.42	81.46	$2.21*10^{15}$	$3.02*10^{25}$	$1.92*10^{3}$	$3.11*10^{5}$	$4.15*10^{5}$	$1.08*10^{10}$	$3.05*10^{6}$	$2.45*10^{3}$	$4.09*10^{13}$	-			
		BF_{01}	$2.19*10^{-3}$	0.01	$4.53*10^{-16}$	$3.31*10^{-26}$	$5.22*10^{-4}$	$3.22*10^{-6}$	$2.41*10^{-6}$	$9.29*10^{-11}$	$3.28*10^{-7}$	$4.08*10^{-4}$	$2.44*10^{-14}$				
	13. Numerical c.s.	BF_{10}	$5.19*10^{7}$	$3.20*10^{12}$	$1.66*10^{3}$	$1.41*10^{4}$	$3.47*10^{22}$	$2.19*10^{4}$	$1.10*10^{5}$	38.92	$9.85*10^{11}$	$7.80*10^{11}$	$1.67*10^{4}$	$5.92*10^{6}$	-		
		BF_{01}	$1.93*10^{-8}$	$3.13*10^{-13}$	$6.04*10^{-4}$	$7.08*10^{-5}$	$2.88*10^{-23}$	$4.58*10^{-5}$	$9.08*10^{-6}$	0.03	$1.02*10^{-12}$	$1.28*10^{-12}$	$5.98*10^{-5}$	$1.69*10^{-7}$			
	14. Spatial c.s.	BF_{10}	128.81	$1.43*10^{3}$	$2.25*10^{11}$	$3.54*10^{3}$	$2.83*10^{3}$	$1.10*10^{28}$	586.53	$1.84*10^{8}$	$1.26*10^{7}$	$5.34*10^{5}$	$4.78*10^{8}$	$4.13*10^{5}$	$2.97*10^{7}$	-	
		BF_{01}	$7.76*10^{-3}$	$6.99*10^{-4}$	$4.45*10^{-12}$	$2.83*10^{-4}$	$3.53*10^{-4}$	$9.13*10^{-29}$	$1.70*10^{-3}$	$5.43*10^{-9}$	$7.93*10^{-8}$	$1.87*10^{-6}$	$2.09*10^{-9}$	$2.42*10^{-6}$	$3.36*10^{-8}$		
	15. Numerical upd.	BF_{10}	274.86	$2.55*10^{4}$	$6.34*10^{3}$	$1.20*10^{4}$	$2.25*10^{3}$	$1.30*10^{6}$	$9.09*10^{19}$	$5.23*10^{6}$	$9.16*10^{4}$	$5.93*10^{3}$	$5.68*10^{4}$	$6.39*10^{6}$	$6.40*10^{10}$	$6.35*10^{7}$	-
		BF_{01}	$3.64*10^{-3}$	$3.92*10^{-5}$	$1.58*10^{-4}$	$8.36*10^{-5}$	$4.45*10^{-4}$	$7.71*10^{-7}$	$1.10*10^{-20}$	$1.91*10^{-7}$	$1.09*10^{-5}$	$1.69*10^{-4}$	$1.76*10^{-5}$	$1.56*10^{-7}$	$1.56*10^{-11}$	$1.57*10^{-8}$	
	16. Spatial upd.	BF_{10}	225.33	335.53	$2.66*10^{9}$	$5.67*10^{7}$	267.41	$3.02*10^{6}$	$9.14*10^{3}$	$1.85*10^{29}$	959.14	$3.78*10^{4}$	$1.24*10^{11}$	$3.55*10^{8}$	467.77	$4.35*10^{8}$	$1.17*10^{7}$
		BF_{01}	$4.44*10^{-3}$	$2.98*10^{-3}$	$3.76*10^{-10}$	$1.76*10^{-8}$	$3.74*10^{-3}$	$3.31*10^{-7}$	$1.09*10^{-4}$	$5.41*10^{-30}$	$1.04*10^{-3}$	$2.65*10^{-5}$	$8.04*10^{-12}$	$2.81*10^{-9}$	$2.14*10^{-3}$	$2.30*10^{-9}$	$8.57*10^{-8}$

Note. S.s. = simple span; c.s. = complex span; upd. = updating.

Table 6

Synchrony Effect: Inferential Statistical Values for the t*-Tests Comparing On- and Off-Peak Sessions for Each Short-Term Memory and Working-Memory Task Separately, and Bayes Factors (BF) from the Bayesian* t*-Tests.*

Task	$t(190)$	p	Cohen's d	BF_{10}	BF_{01}
Digit simple span	*0.01*	*.988*	*0.00*	*0.08*	*12.37*
Letter simple span	*0.96*	*.338*	*0.07*	*0.13*	*7.86*
Matrix simple span	*1.53*	*.128*	*0.11*	*0.25*	*3.93*
Arrow simple span	**2.75**	**.007***	**0.20**	**3.13**	**0.32**
Numerical complex span	*0.66*	*.509*	*0.05*	*0.10*	*9.98*
Spatial complex span	*0.65*	*.514*	*0.05*	*0.10*	*10.03*
Numerical updating	2.38	.018*	0.17	1.27	0.79
Spatial updating	*1.15*	*.252*	*0.08*	*0.15*	*6.47*

Note. For the sake of clarity, results with a BF_{10} larger than 3 are presented in bold, whereas results with a BF_{01} larger than 3 are presented in italics.

* $p < .05$.

Table 7

Multiverse-analysis approach: Description of the Different Procedures Used in the Present Study.

Procedure	Description
Data transformation	1. **raw accuracy rates**
	2. arcus-sinus transformed accuracy rates
Participants' selection	1. all chronotypes[a]
	2. **moderate and definite morning and evening types**
	3. definite morning and evening types
Trimming	1. with missing data (but no more than 2 short-term or working-memory tasks were missing in each session)
	2. after removing missing data
	3. after removing outliers in time-out responses (i.e., number of time-out responses smaller or larger than 3 SDs) in the processing parts of the complex spans
	4. after removing multivariate outliers in all span tasks
	5. **after removing multivariate outliers in all span tasks and the processing parts of the complex spans**
	6. after removing outliers in depression (i.e., BDI-II score larger than 20) and stress (i.e., PSQ score larger or smaller than 3 SDs)
	7. after removing outliers in depression (i.e., BDI-II score larger than 20) and stress (i.e., PSQ score larger or smaller than 3 SDs) as well as those participants reporting medicament intake in the last 24 hours or alcohol, drug, caffeine or nicotine intake in the last 2 hours before a session
SEM approach	1. latent-change model without constraints

Table 7

Multiverse-analysis approach: Description of the Different Procedures Used in the Present Study. (continued)

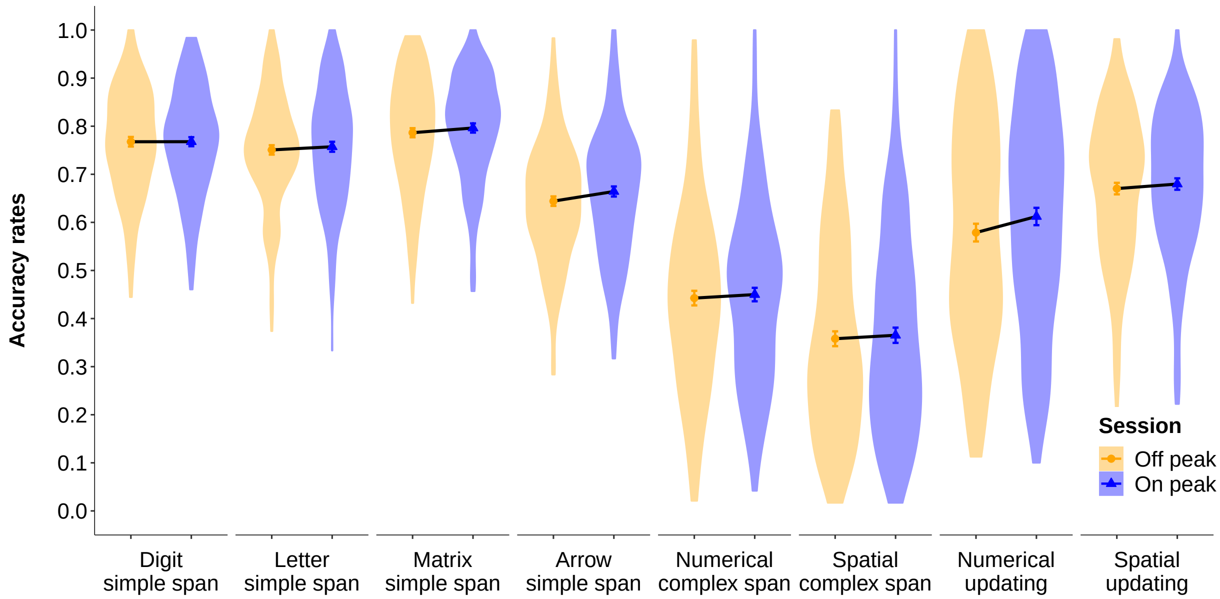
Procedure	Description
	2. **latent-change model with the constraints of positive factor loadings**
	3. latent-change model with the constraints of positive error variances
	4. latent-change model with the constraints of positive factor loadings and error variances
	5. latent-change model without the covariance between the latent-change factors in case of bivariate latent-change model
	6. latent-change model without the measures which were used to assess attentional control and which showed a significant synchrony effect at the individual-task level
	7. second-order latent growth model

Note. Structural equation modeling was performed only if the sample size was larger than 80 participants. The procedure presented in the main text is displayed in bold. SD = standard deviation; BDI-II = Beck Depression Inventory II (Hautzinger et al., 2006); PSQ = Perceived Stress Questionnaire (Fliege et al., 2005); SEM = Structural Equation Modeling.

[a] Participants with a score smaller than 50 were classified as evening types, whereas participants with a score larger than 50 were classified as morning types. For participants with a score of 50, they were classified as morning or evening types according to the question: “One hears about ‘morning types’ and ‘evening types’. Which one of these types do you consider yourself to be?” The potential responses were: “Definitely a morning type”, “Rather more a morning type than an evening type”, “Rather more an evening type than a morning type”, and “Definitely an evening type” (see Griefahn et al., 2001).

Figure 1

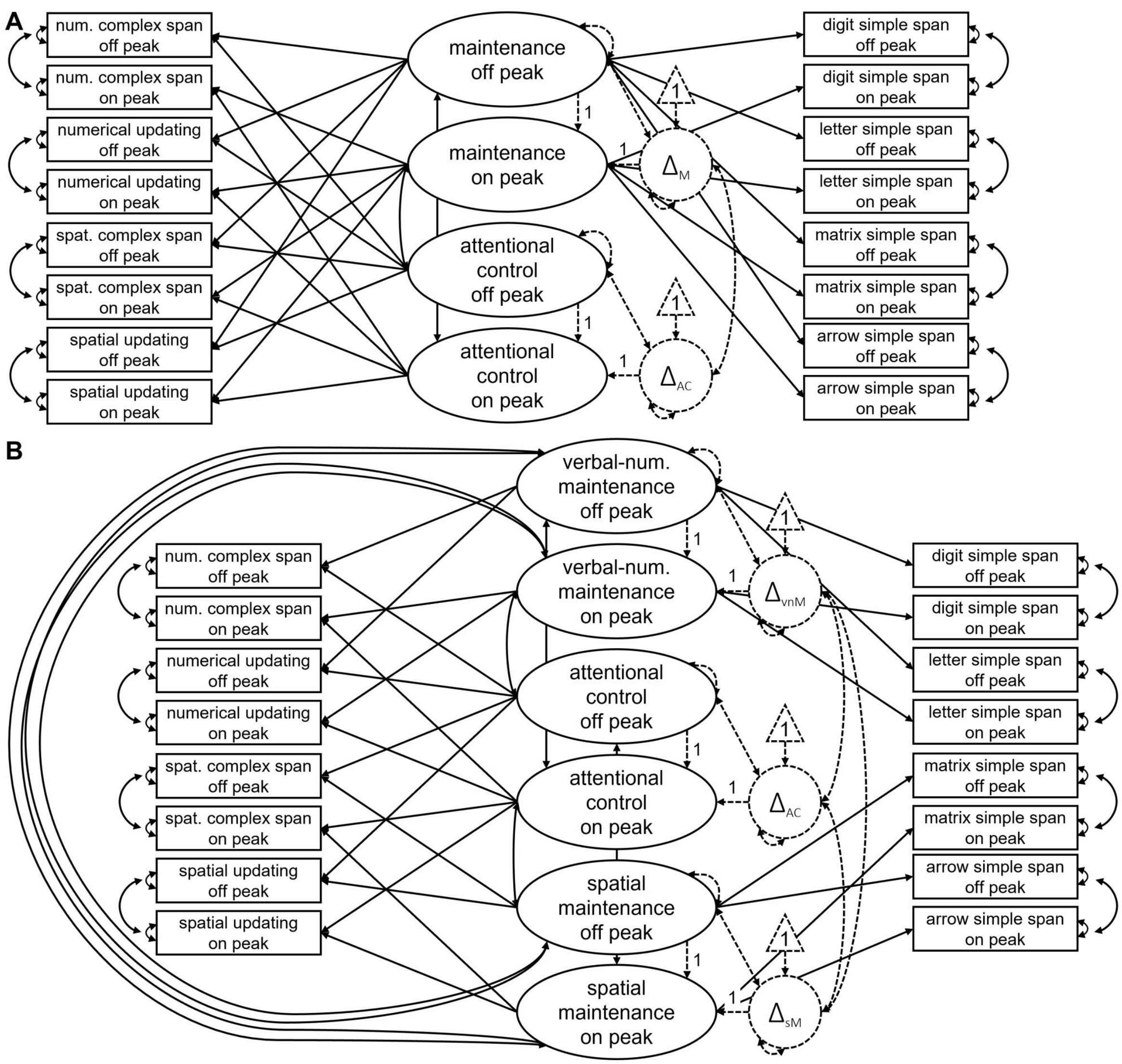
Synchrony Effect (i.e., Mean Accuracy Difference between On- and Off-Peak Sessions) for Each Short-Term and Working-Memory Task



Note. Error bars represent within-subject confidence intervals (Cousineau, 2005; Morey, 2008). Shaded areas display data density plots.

Figure 2

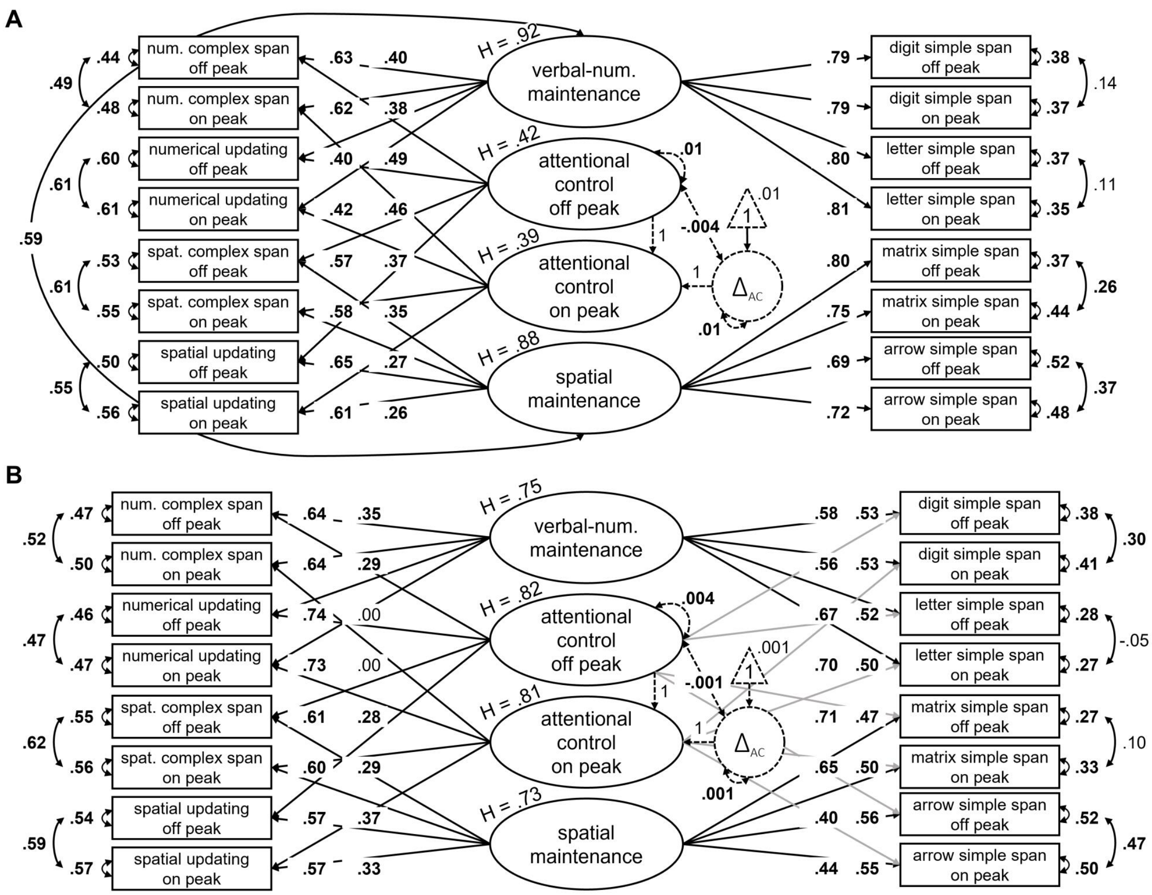
Latent-Change Model 1 and Latent-Change Model 2



Note. Panel A: Model 1. Panel B: Model 2. Please see the text for a detailed description of the model parameters. In both models, the synchrony effect was estimated as the latent change between off and on peaks for each latent construct. Paths and variables specific to the estimation of the latent changes are displayed in dashed lines. Means are omitted for visual clarity. Num. = numerical; spat. = spatial; Δ_{AC} = latent-change factor for attentional control; Δ_M = latent-change factor for short-term maintenance; Δ_{vnM} = latent-change factor for verbal-numerical short-term maintenance; Δ_{sM} = latent-change factor for spatial short-term maintenance.

Figure 3

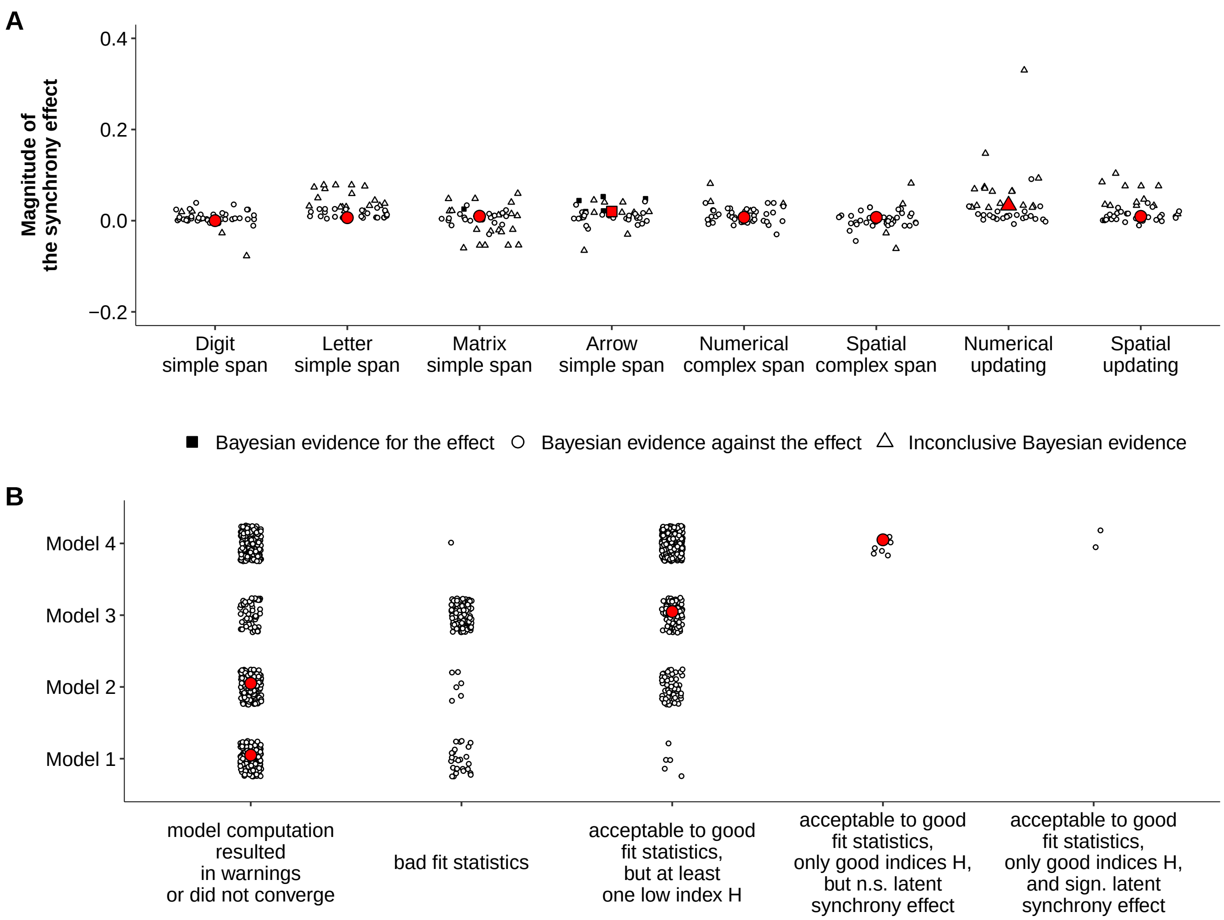
Latent-Change Model 3 and Latent-Change Model 4



Note. Panel A: Model 3. Panel B: Model 4. See Figure 2 note for details. The numbers next to the straight, single-headed arrows are the standardized factor loadings (interpretable as standardized regression coefficients). For the sake of clarity, the loadings are aligned to each measure (i.e., next to the measures for the general factor(s) and next to the factors for the specific factor(s)). The outward numbers adjacent to each measure are the error variances, attributable to idiosyncratic task requirements and measurement error. The numbers adjacent to the curved, double-headed arrows next to each measure are the correlations between the error variances. Parameters specific to the latent change are unstandardized. For all values, boldface type indicates $p < .05$. Num. = numerical; spat. = spatial; Δ_{AC} = latent-change factor for attentional control.

Figure 4

Synchrony Effect: Overview of the Results from the Multiverse-Analysis Approach



Note. Panel A: Magnitude of the synchrony effect across short-term and working-memory tasks. Panel B: Results from the latent-change models estimating the synchrony effect for short-term maintenance and attentional control. A model fit was considered as good if the Bentler's comparative fit index (CFI) was larger than .95 and the standardized root-mean-square residual (SRMR) was smaller than .08. A model fit was considered as acceptable if the CFI ranged from .90 to .95 and the SRMR was smaller than .08. Otherwise, the model fit was considered as bad. The index *H* was considered as low if it was smaller than .70. Otherwise, it was considered as good. For both panels, each datapoint represents the results combining the different data transformations, participants' selections, trimming procedures, and analyses listed in Table S7. The results reported in the text are presented in red. N.s. = not significant; sign. = significant.

Appendix

Time-of-day effects

We investigated the time-of-day effect, that is, the difference in performance between the morning and evening sessions. Thus, irrespective of the chronotype, performance in the morning was compared to performance in the evening. Consistent with the analyses on the synchrony effect, the analyses were first performed at the individual-task level and then at the latent-variable level.

At the individual-task level, the descriptive results are displayed in Figure A1, and the results from the null hypothesis significance testing (NHST) and Bayesian approach are described in Table A1. For most tasks, no time-of-day effect was observed. Only in the digit simple span and in the numerical updating task, Bayesian evidence was inconclusive regarding the presence or the absence of the effect. However, the effect sizes were small for these tasks.

At the latent-variable level, the results are presented in Figure A2. Consistent with the results on the synchrony effect, the results showed a difficulty of measuring maintenance and attentional control at the latent-variable level. When these were modeled satisfactorily (i.e., with acceptable to good fit statistics and good indices H), the time-of-day effect was in most cases not significant. When, however, it was significant, the latent change was small (maximum value for the unstandardized intercept of the latent-change factor = 0.02).

Table A1

Time-Of-Day Effect: Inferential Statistical Values for the t*-Tests Comparing Morning and Evening Sessions for Each Short-Term Memory and Working-Memory Task Separately, and Bayes Factors (BF) from the Bayesian* t*-Tests.*

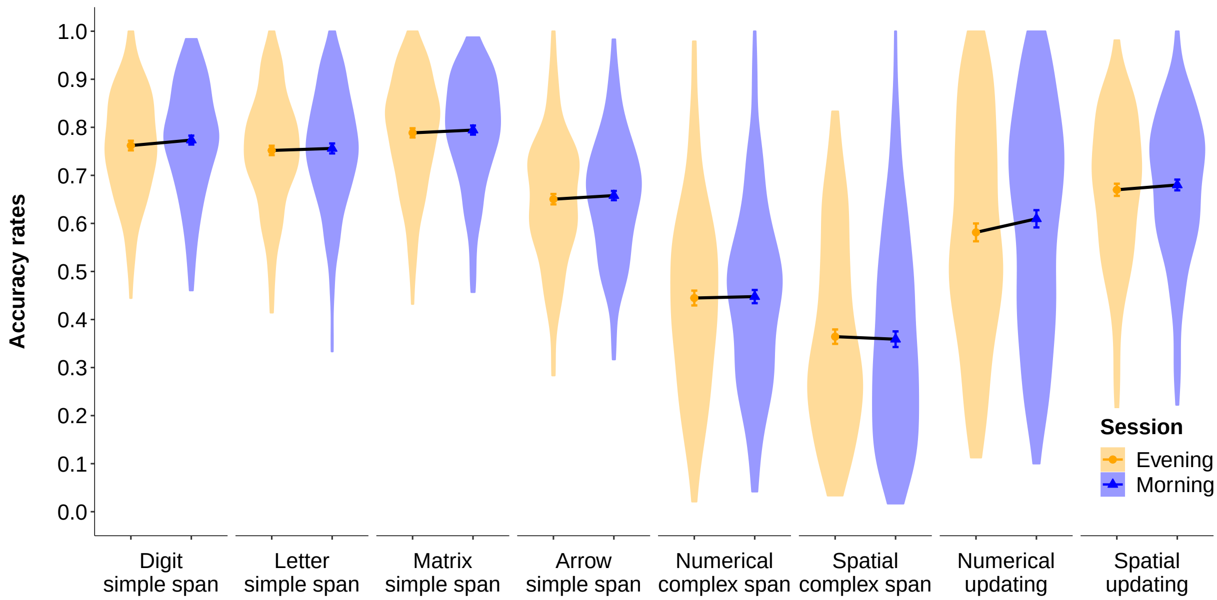
Task	$t(190)$	p	Cohen's d	BF_{10}	BF_{01}
Digit simple span	1.77	.079	0.13	0.37	2.70
Letter simple span	*0.60*	*.546*	*0.04*	*0.10*	*10.34*
Matrix simple span	*0.90*	*.368*	*0.07*	*0.12*	*8.29*
Arrow simple span	*1.01*	*.313*	*0.07*	*0.13*	*7.49*
Numerical complex span	*0.27*	*.789*	*0.02*	*0.08*	*11.94*
Spatial complex span	*-0.48*	*.632*	*0.03*	*0.09*	*11.05*
Numerical updating	2.00	.047*	0.14	0.57	1.76
Spatial updating	*1.22*	*.226*	*0.09*	*0.17*	*5.99*

Note. For the sake of clarity, results with a BF_{10} larger than 3 are presented in bold, whereas results with a BF_{01} larger than 3 are presented in italics.

* $p < .05$.

Figure A1

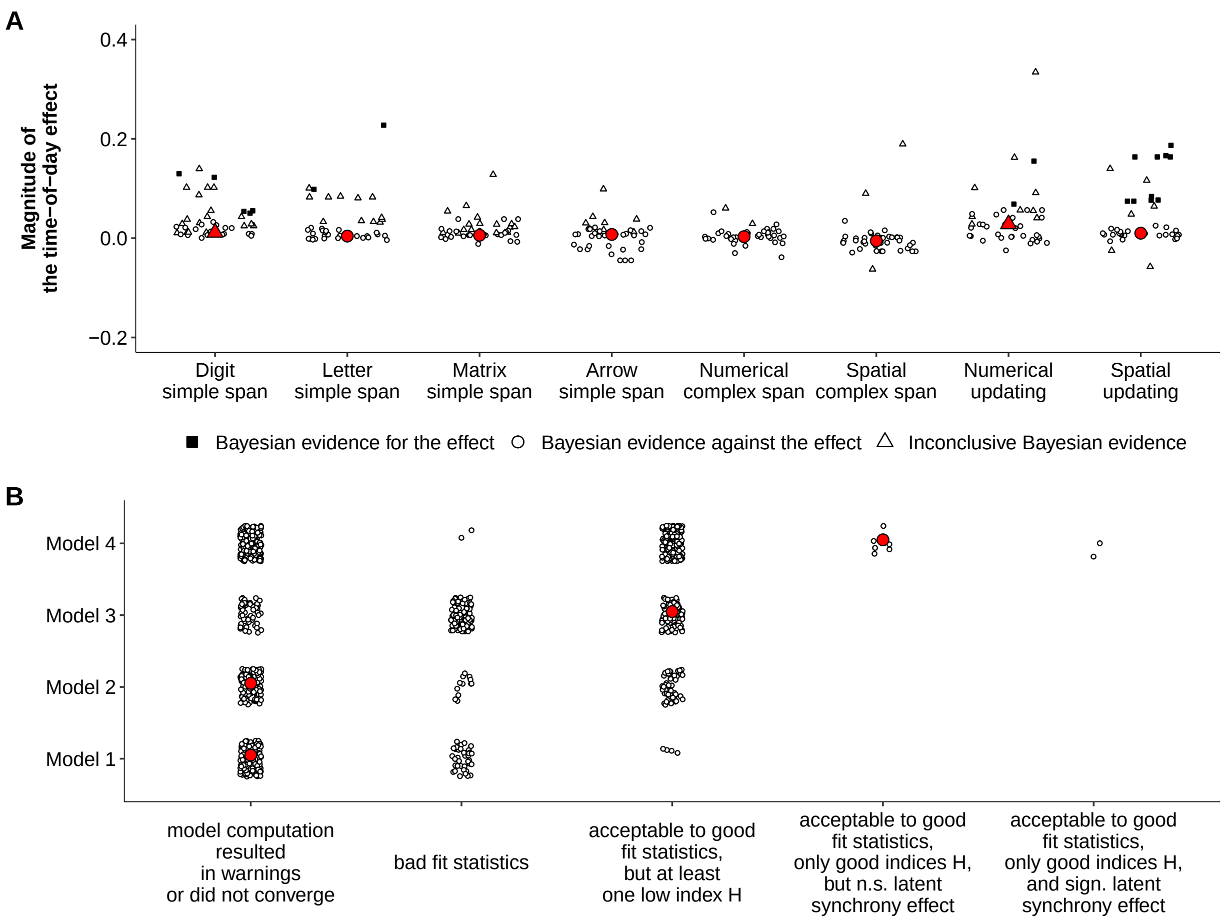
Time-Of-Day Effect (i.e., Mean Accuracy Difference between Morning and Evening Sessions) for Each Short-Term Memory and Working-Memory Task



Note. Error bars represent within-subject confidence intervals (Cousineau, 2005; Morey, 2008). Shaded areas display data density plots.

Figure A2

Time-Of-Day Effect: Overview of the Results from the Multiverse-Analysis Approach



Note. Panel A: Magnitude of the time-of-day effect across short-term memory and working-memory tasks. Panel B: Results from the latent-change models estimating the synchrony effect for short-term maintenance and attentional control. A model fit was considered as good if the Bentler's comparative fit index (CFI) was larger than .95 and the standardized root-mean-square residual (SRMR) was smaller than .08. A model fit was considered as acceptable if the CFI ranged from .90 to .95 and the SRMR was smaller than .08. Otherwise, the model fit was considered as bad. The index H was considered as low if it was smaller than .70. Otherwise, it was considered as good. For both panels, each datapoint represents the results combining the different data transformations, participants' selections, trimming procedures, and analyses listed in Table S7. The results from the procedure reported in the text are presented in red. N.s. = not significant; sign. = significant.