

NOTICE: This is the authors' version of a work that was **accepted for publication in Behavior Research Methods**, and can be used for scholarly non-commercial purposes. Changes resulting from the publishing process, such as editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Definitive version will be subsequently published by the journal.

Unifying temporal preparation: The Temporal Preparation Task (TEP-Task)

Mariagrazia Capizzi^{1,2,*,#}, Lucie Attout^{3,*}, Giovanna Mioni⁴, Pom Charras^{3,#}

¹Mind, Brain and Behavior Research Center (CIMCYC), University of Granada; ²Department of Experimental Psychology, University of Granada, Granada, Spain

³Univ Paul Valéry Montpellier 3, EPSYLON EA 4556, F34000, Montpellier, France

⁴Department of General Psychology, University of Padova, Padova, Italy

* Joint first authors

Corresponding authors

Mariagrazia Capizzi

Mind, Brain and Behavior Research Center (CIMCYC), University of Granada, Campus of Cartuja, 18011, Granada, Spain

Department of Experimental Psychology, University of Granada, Granada, Spain

mgcapizzi@ugr.es

Pom Charras

Epsilon, Université de Montpellier Paul-Valéry, Site de Saint-Charles

Rue du Professeur Henri Serre

34000 - Montpellier

pom.charras@univ-montp3.fr

Abstract

The dynamic nature of our environment allows us to anticipate the onset of relevant events, enhancing our responses to them. Temporal preparation can be assessed in the laboratory using various tasks, including foreperiod tasks, temporal orienting tasks, and rhythmic tasks. However, the existing literature lacks a unified task to measure the most common temporal preparation effects (i.e., foreperiod, sequential, temporal orienting, and rhythmic effects) in a single session. The main goal of the present study was to fill this gap by devising the temporal preparation task (TEP-Task) to measure temporal preparation effects in a single 35-min testing session. Besides its utility in single-session assessments, the TEP-Task may also serve for future research across diverse populations and experimental demands.

Keywords: foreperiod, sequential effects, temporal orienting, auditory rhythm, visual rhythm.

Introduction

Anticipating and preparing for upcoming events, commonly referred to as temporal preparation, is crucial for optimizing response readiness. The ubiquity of temporal information in our environment offers numerous opportunities for this process. For example, the longer a slice of bread is toasted, the higher the expectation that it will pop out of the toaster, allowing for a more prepared response. Likewise, a traffic light changing from red to green or the regular rhythm of a countdown (3...2...1...) before a race can prompt faster reactions. In experimental settings, these real-world temporal scenarios – passage of time, symbolic cues, and rhythmic patterns – are captured using well-established tasks: foreperiod tasks, temporal orienting tasks, and rhythmic tasks, respectively. These tasks are widely employed to investigate temporal preparation and to promote the development of a more integrative theoretical framework, given the current absence of a unified account in the field. Indeed, while a consistent finding across such tasks is that stimuli presented at expected intervals elicit faster responses than those presented at unexpected times, it remains an open question whether they reflect common or distinct temporal preparation processes (see Denison, 2024; Nobre & Van Ede, 2018; Seibold et al., 2023, for recent reviews).

In a typical foreperiod task, participants are presented with a warning signal followed by a target stimulus that demands a quick response. The time interval between warning and target is referred to as the “foreperiod”. When various foreperiods are randomly intermixed across trials, each with an equal a-priori probability of appearing, a common finding is that reaction times (RTs) decrease at longer foreperiod trials¹ (Capizzi & Correa, 2018; Coull, 2009; Niemi & Näätänen, 1981; Visalli et al., 2023). This phenomenon, termed the *foreperiod effect*, is formalized by the hazard function, i.e., the conditional probability that an event will occur

¹ Unlike in the variable foreperiod task, in a fixed foreperiod task, the foreperiod remains constant within a block but varies between blocks. In such a case, RTs are typically shorter in blocks with shorter foreperiods (Bausenhardt et al., 2008; Rolke et al., 2016). This advantage is attributed both to enhanced temporal preparation for short and predictable foreperiod blocks and to the fact that time estimation becomes less precise as interval durations increase, consistent with Weber’s law (Gibbon, 1977).

given that it has not occurred yet (Herbst, Fiedler, & Obleser, 2018; Janssen & Shadlen, 2005; Visalli et al., 2019; 2021).

Besides the foreperiod effect, another temporal phenomenon observed in the variable foreperiod task is represented by *sequential effects*. Sequential effects reflect the finding of shorter RTs at the current short foreperiod when the previous foreperiod was short rather than long. By contrast, at the current long foreperiod, there is usually little or no difference in RTs between a previous short and a previous long foreperiod (Capizzi et al., 2015; Han & Proctor, 2022; Los et al., 2014; Steinborn et al., 2008; Vallesi et al., 2013). Sequential effects underscore our ability to leverage implicit temporal information to shape expectations and predict when a specific event might occur. Thus, both foreperiod and sequential effects highlight the incidental use of time in contexts that do not explicitly emphasize temporal aspects in their overt task goals (Capizzi et al., 2022; Coull & Nobre, 2008). However, it is debated whether they are explained by a single temporal preparation process or they result from different processes (Salet et al., 2022; Vallesi, 2010).

Unlike the foreperiod task, in the temporal orienting task an informative cue signals when the target is most likely to appear (Correa et al., 2006; Coull & Nobre, 1998; Kingstone, 1992). *Temporal orienting effects* are usually measured by comparing performance in valid trials, where the cue predicts the timing of target onset with high probability (~75-80%), to invalid trials, where the cue and target timing do not align (i.e., a late cue is paired with an early target onset, and vice versa). Another approach is to compare valid trials, where the cue is 100% accurate, to neutral trials, where the cue provides no information about the target timing, with 50% of trials having an early target onset and 50% a late one. In both approaches, temporal orienting effects are stronger at shorter foreperiods but tend to diminish or disappear at longer foreperiods. If the target does not appear after a short foreperiod in the early-cue/late-target invalid condition (or in the neutral condition), participants may still

reorient to the longer foreperiod, thereby reducing the negative impact of an unmet early expectation regarding target onset (Capizzi et al., 2012; Correa et al., 2004). The choice between a valid-invalid and a valid-neutral design depends on the study's goal, with the latter approach used when the focus is on the facilitatory effects of temporal cues (Capizzi et al., 2023; Mento & Tarantino, 2015).

While symbolic cues can deliberately orient attention to time, regular rhythmic patterns, such as a sequence of stimuli presented at a constant predictable pace, can foster temporal attention more automatically (Jones et al., 2002; Large & Jones, 1999). This allocation of attention leads to *rhythmic effects*, with faster RTs for targets following a regular rhythm compared to an irregular one (Cravo et al., 2013; Rohenkohl et al., 2012; Román-Caballero et al., 2024; Sanabria et al., 2011). Relative to temporal orienting effects, rhythmic effects survive dual-task interference (Cutanda et al., 2015; 2019; de la Rosa et al., 2012), remain effective when participants are instructed to ignore them (Rohenkohl et al., 2011), and can function even when their use could potentially hinder the task at hand (Breska & Deouell, 2014).

Based on the above, temporal preparation is not a unitary construct as it can be shaped by different sources of information across various tasks. Converging evidence from behavioral, neuropsychological, and non-invasive brain stimulation studies support the notion that temporal preparation effects are mediated by dissociable processes and neural mechanisms (Correa et al., 2014; Breska & Ivry, 2018; de la Rosa et al., 2012; Triviño et al., 2010; 2011; Vallesi et al., 2007a, 2007b, 2014). However, these dissociations have mostly emerged from comparisons of only a subset of temporal preparation effects such as temporal orienting vs. rhythmic effects (Breska & Ivry, 2018; Capizzi et al., 2013), or foreperiod vs. sequential effects (Vallesi et al., 2014). Furthermore, studies have varied widely in stimuli, protocols, and foreperiod durations, and it is known that the processes underlying each effect may be

conflated by the chosen manipulations and task parameters (see Lawrence & Klein, 2013; Olk, 2014).

Only a few studies have systematically examined multiple temporal preparation effects within a single session. Correa and colleagues (2010) devised a temporal orienting task to simultaneously measure temporal orienting, foreperiod, and sequential effects. Temporal orienting effects were indexed by subtracting RTs in valid trials from those in invalid trials. The foreperiod effect was measured by comparing RTs in short and long foreperiod trials within the invalid condition. Sequential effects were calculated as the difference between previous-long and previous-short foreperiods, regardless of cue validity. Despite the simplicity of using a single temporal orienting task to measure different temporal preparation effects, it has been shown that these effects might influence one another. For example, sequential effects are larger when attention is not explicitly directed to a particular moment in time (i.e., in invalid temporal orienting conditions), suggesting that strategic processes can modulate their magnitude (Los & van den Heuvel, 2001). Supporting this, sequential effects are more pronounced in blocked manipulations of temporal cues compared to a within-trial design (Correa et al., 2004). For this reason, foreperiod and sequential effects are usually measured in neutral blocks (rather than valid ones) to minimize confounds from cue-related expectancy effects (Capizzi et al., 2023; Johnson et al., 2016).

Another limitation of using a single temporal orienting procedure to measure different temporal preparation effects is that it typically precludes the inclusion of rhythmic cues. Although some studies comparing static and rhythmic cues have manipulated the pace of rhythmic sequences (e.g., fast vs. slow) as a cue for target onset timing (Johnson et al., 2016; Triviño et al., 2011), this approach does not rely solely on rhythmic entrainment. Participants must extract temporal information from the sequence pace and apply it to the forthcoming foreperiod, thus reducing the purely exogenous influence of rhythm.

Overall, while previous research has highlighted the importance of measuring multiple temporal preparation effects, capturing all four within a single block remains challenging. To address this, we designed the Temporal Preparation Task (TEP-Task), which uses separate blocks to target each temporal preparation effect within a simple RT paradigm. A simple task design was preferred over a choice-based or more perceptually demanding task to avoid additional complexity and longer testing duration (see also the Discussion section).

In sum, the TEP-Task was developed to address two main gaps. First, as outlined above, most studies have examined temporal preparation effects in isolation, which limits direct and statically robust comparisons across effects. Although a few studies have compared multiple temporal preparation effects within the same sample, no existing task has incorporated all of them. Second, the TEP-Task allows for the combined assessment of different temporal preparation effects within a relatively short timeframe and with sufficient reliability (see below). By design, however, the TEP-Task is not intended to disentangle the mechanisms underlying these effects across varying task demands, as all measurements are obtained within a simple RT context.

Method

Participants

One hundred and fourteen participants took part in the study in exchange for course credits. All of them were enrolled in an introductory psychology course at the Montpellier Paul Valery University (France). All participants reported having normal or corrected-to-normal vision and normal audition. Four participants were excluded due to the presence of diagnosed developmental disorders such as dyspraxia, specific language impairment, or Attention Deficit Hyperactivity Disorder (ADHD). Data from one additional participant were discarded for excessive slower RTs with respect to the sample mean (i.e., >3 SD). A final sample of 109

participants (mean age= 20.88 years, range= 17–54 years, SD= 5.31 years, 93 females, 17 left-handed) was considered for the analyses. Because of the difficulty in collecting data from the same group of participants, only a subgroup of the initial sample (N=44, mean age= 20.55 years, range= 17–54 years, SD= 5.47 years, 35 females, 7 left-handed) repeated the task one week later to evaluate test-retest reliability.

Monte Carlo simulations and the Bayesian Factor Design Analysis (BFDA; <http://shinyapps.org/apps/BFDA/>; Schönbrodt & Wagenmakers, 2018) assessed the sensitivity of our statistical design in providing evidence for the significant presence of temporal preparation effects. For the sensitivity analysis, we assumed a small to medium effect size of Cohen's $d=0.30$, according to previous literature (Capizzi et al., 2015; de la Rosa et al., 2012; Steinborn et al., 2008). If the effects of interest exist, the median sample size required to achieve a high level of evidence ($BF_{10}>10$) in support of the effect should be $N=112$.

The study was conducted in accordance with a protocol approved by the Paul Valéry Ethical Committee (2020-02), and all the participants signed an approved informed consent form prior to participation.

Apparatus and stimuli

Participants completed the study at the Platform of research and analysis of human behaviour dedicated to student training courses (PEACH platform, University of Montpellier Paul Valéry). The experiment was run in a group testing room equipped with a series of personal computers connected to 15-in monitors. Opensesame 3.3.5 (Mathôt et al., 2012) was used to program and run the experiment. The viewing distance was approximately 60 cm.

The background color of the monitor was set to black throughout the whole experiment, while the instructions, feedback and stimuli were displayed in white. Stimuli included a fixation cross (size $0.5^\circ \times 0.5^\circ$ of visual angle), a series of visual and auditory cues, and a target. The

visual cues comprised a central stimulus formed by a largest (outer) circle and a smallest (inner) circle (diameter: 2.4° and 0.95° , respectively). The auditory cues were a series of 80-dBA pure tones of 700 Hz. The target stimulus was an asterisk (*), with a size of $0.8^\circ \times 0.8^\circ$.

Procedure and task

Figure 1 illustrates the timing and sequence of events for all temporal manipulations, each implemented in a separate block. Specifically, there was a temporal orienting block, a neutral block (serving as a measure for temporal orienting and sequential effects), a foreperiod block (assessing the foreperiod effect), a visual rhythmic block, and an auditory rhythmic block². A valid-neutral design was used for temporal orienting to measure the facilitatory effects of temporal cues. A valid-invalid design would have also lengthened task duration, as it requires allocating a sizable proportion of trials to invalid conditions (typically ~20-25%). The overall block structure was consistent across manipulations, except for the specific differences detailed below.

In the temporal orienting and neutral blocks, each trial began with a fixation cross followed by a precue (a central stimulus with inner and outer circles); both were shown for the same randomly selected duration between 500 and 1500 ms (uniform distribution). The precue was intentionally designed for its use in future studies with more vulnerable populations. Instead of using a cue with an abrupt onset, we opted for a precue that allowed participants to focus their attention on the stimulus before the brightening of the inner or outer circle, thereby facilitating the interpretation of the cue's meaning. Following the precue, either the inner or

² Only for the rhythmic task, both visual and auditory versions were included, given prior evidence suggesting that rhythmic effects may be sensitive to presentation modality. Findings on the advantage of auditory over visual rhythmic sequences remain inconsistent, with some studies reporting no auditory advantage (Cutanda et al., 2019; Jones, 2015) and others showing superior rhythmic effects in the auditory modality (Attout et al., 2024; Ball et al, 2022). Although modality effects were not the primary focus of this study, we included an auditory rhythmic block as an additional exploratory component to assess potential differences with the visual rhythmic block. Given that the TEP-Task is customizable, this auditory block can be incorporated into future studies or omitted if deemed unnecessary, also in light of our findings directly comparing between auditory and visual rhythms.

outer circle (in the temporal orienting block) or both circles (in the neutral block) brightened for 200 ms, transitioning from a thinner to a thicker border. In the temporal orienting block, the brightening of the inner circle indicated with 100% validity that the target would appear after a short foreperiod (400 ms), whereas the brightening of the outer circle signalled a long foreperiod (1600 ms). The foreperiod started at the cue offset³. Participants were explicitly instructed to use the temporal cues to anticipate target onset. In the neutral block, the brightening of both circles served as a neutral warning signal, providing no information about the foreperiod duration (either 400 or 1600 ms) preceding target onset. After the cue disappeared, the screen remained blank for a variable duration based on the foreperiod assigned to that trial. At the end of the foreperiod, the target appeared for 150 ms. In half of the trials, the target could appear to the right of fixation (0.5°), whereas in the remaining trials, it could appear to the left of fixation (-0.5°). This slightly lateralized target presentation was aimed to improve discriminability from the preceding centrally presented stimuli, aiding future comparisons with studies involving more vulnerable populations, such as older adults and patients. Participants had to respond to the target onset by pressing the left mouse button (not visible on the screen) with their right hand as quickly as possible while avoiding anticipatory responses. If a response was initiated before target onset, a visual message warning the participant to wait for the target replaced target presentation and the trial was repeated (up to a maximum of six times). Otherwise, a blank screen lasting 2000 ms was displayed for response collection. Each trial ended with a 750 ms blank inter-trial interval.

³ Because the temporal cues carried symbolic meaning that participants were required to interpret, it is reasonable to assume that preparation primarily began at cue offset rather than at cue onset (see Los & Schut, 2008, for discussion). Although the cue carried no meaning in the neutral and foreperiod blocks, even assuming that preparation began at cue onset in these blocks would not change the key findings and their interpretation, namely, that the passage of time confers a performance advantage, as evidenced by shorter RTs on long foreperiod trials.

The foreperiod block was identical to the neutral block, except that the target could appear after one of four possible foreperiods (400, 800, 1200, or 1600 ms), each occurring with equal probability, following a commonly used aging distribution.

The rhythmic blocks differed from the temporal orienting, neutral and foreperiod blocks in the following aspects. In the visual rhythmic block, the fixation cross was followed by six blank screens interleaved with five central circles (the same concentric circles used in the neutral and foreperiod blocks), each presented for 100 ms. A shorter 100-ms stimulus presentation, compared to the 200-ms temporal cues, was used to emphasize the exogenous nature of the rhythmic sequence. The alternation of blank screens and visual stimuli created a flickering effect that could be either regular or irregular. In the regular sequence, each blank screen, except for the last one, had a fixed duration of 500 ms. Conversely, in the irregular sequence, the duration of the blank screens varied randomly across different values (100, 300, 500, 700, or 900 ms), resulting in a total duration of 2500 ms, matching that of the regular rhythm. The final blank screen preceding the target (i.e., the foreperiod) had a fixed duration of 1000 ms for both the regular and irregular sequences (see Charras et al., 2023, for a similar procedure). In the auditory rhythmic block, five 100-ms sounds were delivered through headphones, interleaved with silent intervals arranged in either regular or irregular pattern. A blank screen was shown throughout the sound sequence. Following the visual or auditory rhythms, the target appeared, with all other aspects remaining consistent with the temporal orienting, neutral, and foreperiod blocks. Participants were instructed to detect the target and informed that the preceding rhythmic sequence was task-irrelevant.

All temporal blocks were randomly presented across participants. The temporal orienting block comprised 16 short-cued and 16 long-cued foreperiod trials, randomly intermixed. The neutral block comprised 64 trials, with 32 trials per foreperiod. Of these, 16 trials were preceded by a short foreperiod ($n-1_{\text{short}}$) and 16 trials by a long foreperiod ($n-1_{\text{long}}$). For this

block only, the trial order was pseudo-randomized to ensure an equal number of trials for each previous foreperiod. The foreperiod block included a total of 48 trials, with 12 trials per foreperiod. Finally, the visual and auditory rhythmic blocks each included 32 trials, equally divided between regular and irregular conditions. A short, self-paced break was allowed between blocks.

Prior to the experimental blocks, participants completed two brief practice blocks to familiarize themselves with both the sequential and static cue types used in the task. The first block comprised four regular and four irregular auditory rhythmic trials, whereas the second consisted of four short-cued and four long-cued foreperiod temporal orienting trials, illustrating the association between cue type and foreperiod duration. Participants were then informed about the neutral and visual rhythmic cues to avoid novelty effects from these stimuli during the main task. Participants received feedback on correct detections or missed responses (displayed for 1000 ms) after each practice trial. As noted above, during the experimental trials, participants were only informed about anticipatory responses. Accuracy was monitored during each practice block, and if it fell below 75% across the 8 presented trials, an additional 8 trials were administered within the same block. If performance remained below criterion after these extra trials (either in one or both types of practice), a warning screen was programmed to appear, indicating that the minimum accuracy requirement had not been met. This feature was included for future testing with vulnerable populations. However, it was not triggered for any of the younger participants in the present study. The experimenter remained in the room throughout the entire session to ensure a quiet environment to perform the task.

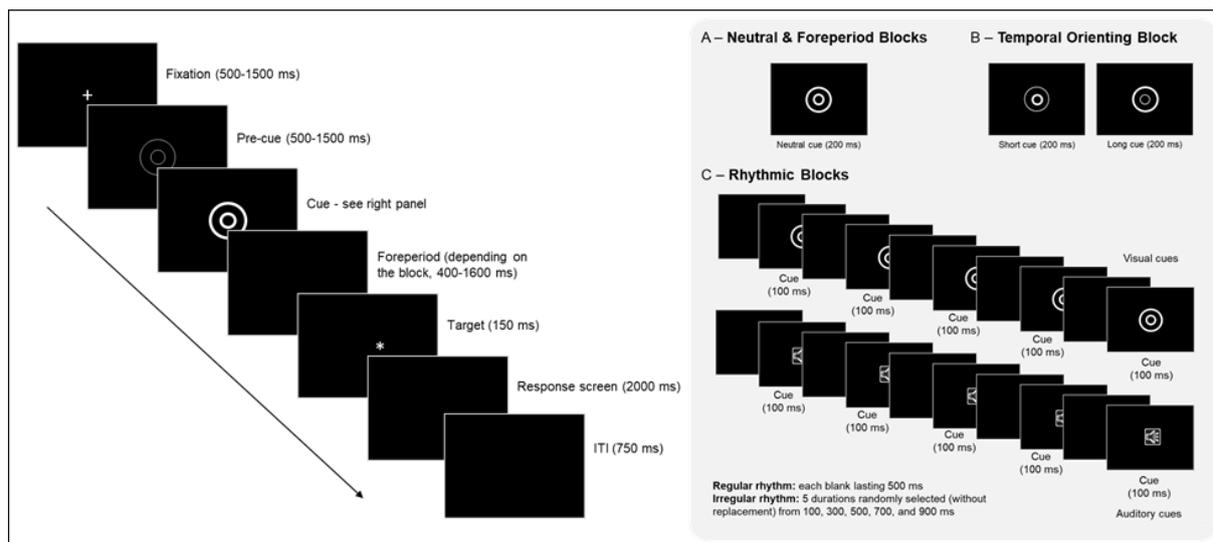


Figure 1. Trial structure across (A) Neutral and Foreperiod blocks, (B) Temporal orienting block, and (C) Rhythmic blocks. In the neutral, foreperiod, and temporal orienting blocks, each trial started with the fixation cross, followed by a precue stimulus. Subsequently, the two circles (for neutral and foreperiod blocks) or either the inner or outer circle (for the temporal orienting block) transitioned from a thinner to a thicker border, respectively, providing neutral or valid information about the timing of the target onset. In the rhythmic blocks, the fixation cross was followed by a sequence of blank screens interleaved with either five visual or five auditory stimuli presented at a regular or irregular pace. Please refer to the main text for further details. ITI stands for inter-trial interval.

Data Analysis

Descriptive statistics and Internal Consistency Analysis

Data from practice trials were discarded before data analysis. Trials with RTs below 150 ms (0.81% of trials), indicative of premature responses, and trials with missed responses (1.75%) were excluded. Next, for each participant and temporal condition, trials were removed if RTs deviated more than $\pm 2.5 SD$ from their individual block means (0.54% of trials).

Statistical analyses were performed with a Bayesian approach, which relies on a model comparison rationale and adopts a model selection strategy to quantify the strength of

evidence both for and against each model, in contrast to the traditional frequentist framework (Morey & Rouder, 2011). We report results using the Bayes factor (BF), a relative measure of evidence that quantifies how much more likely the data are under one model compared to another (e. g., H_1 vs. H_0). For instance, a BF_{10} of 15 indicates that the data are 15 times more likely under H_1 than H_0 . Conversely, evidence in favor of H_0 is expressed as $BF_{01}=1/BF_{10}$, representing how much more likely the data are under H_0 compared to H_1 . Although the Bayesian framework does not rely on fixed statistical thresholds, we followed the indicative guidelines proposed by Jeffreys (1961) for interpreting BF: $BF < 1$ indicates no evidence, $1 < BF < 3$ suggests anecdotal evidence, $3 < BF < 10$ corresponds to moderate evidence, $10 < BF < 30$ reflects strong evidence, $30 < BF < 100$ represents very strong evidence, and $BF > 100$ indicates extreme or decisive evidence for the presence of a given effect. Bayesian analyses were conducted using JASP software (version 0.10.2.0; JASP Team, 2017; jasp-stats.org), using default settings for the Cauchy prior distribution (location = 0, scale = 0.707). This prior is advantageous as it assumes that small effect sizes are generally more likely than large ones (Berkhout et al., 2024; Morey & Rouder, 2018). For correlation analyses, we instead used the default stretched beta prior with a width of 1.0. Partial Eta Squared (η^2_p) and Cohen's d (Cohen, 1977) were used as measures of effect size for ANOVAs and t-tests, respectively.

To assess participants' processing speed, we first examined Bayesian correlations between RTs across the different blocks, to ensure that any variations in RTs could not be attributed to general task context or to instructions that were consistent across all conditions. To assess internal consistency, we conducted split-half reliability analyses within each block. For instance, in the temporal orienting block, trials for each condition (valid-short and valid-long) were randomly split into two halves, the correlations between halves were computed, and the Spearman-Brown correction was applied. We also assessed the reliability of the temporal preparation effects, calculated as described below. For example, to estimate the reliability of

the temporal orienting effect, we randomly split short trials from the neutral block and valid-short trials from the temporal orienting block for each participant into two equally sized subgroups, calculated the RT difference between neutral and valid trials within each subgroup, and then calculated the Spearman-Brown corrected reliability index. All reliability estimates were based on 1000 random permutations using a custom R script.

Temporal Preparation Effects Analysis

We conducted separate analyses for each temporal manipulation. Regarding the foreperiod block, mean RTs were submitted to a Bayesian one-way ANOVA, with foreperiod duration (400, 800, 1200, and 1600 ms) as within-participant factor. To compute sequential effects and the temporal orienting effect, analyses were restricted to targets appearing at the short foreperiod in order to control for the influence of the foreperiod effect observed at the long foreperiod (Coull & Nobre, 1998). The full analyses with long foreperiods can be found as Supplementary material⁴. For sequential effects, a Bayesian paired t-test compared RTs between trials preceded by a short foreperiod and those preceded by a long foreperiod within the neutral block. For temporal orienting, a Bayesian paired t-test compared RTs from valid trials in the temporal orienting block with those from the neutral block. Finally, for the auditory and visual rhythmic blocks, separate Bayesian paired t-tests compared the RTs between regular and irregular trials.

Magnitude of Temporal Preparation Effects Analysis

To assess the sensitivity of individual RTs to temporal predictability, we computed a normalized delta score as follows: the difference between the condition with supposedly

⁴ As reported in the Supplementary file, the global analyses including the long foreperiod replicated well-established findings in the literature: significant temporal orienting effects were observed at the short foreperiod but not at the long foreperiod, and sequential effects were asymmetrical, with a larger and significant difference between previous short and long foreperiods at the current short foreperiod, but not at the current long foreperiod.

slower RTs (temporally unexpected trials) and the condition with faster RTs (temporally expected trials) divided by the sum of both RTs (unexpected RTs – expected RTs) / (unexpected RTs + expected RTs). For each block, we defined the temporally unexpected and expected trials as previously described (for sequential effects, trials preceded by a long foreperiod and a short foreperiod, respectively; for temporal orienting, neutral and valid trials; and for auditory and visual rhythms, irregular and regular trials). For the foreperiod effect, we contrasted the foreperiods of 400 and 1600 ms⁵.

The use of a normalized score instead of a simple differential score was to compare the effect itself rather than just a raw RT difference, to account for the task's dependency on the overall speed of RTs. Positive delta scores would reflect the presence of the temporal preparation effect, with faster RTs in the condition associated with better temporal preparation. A score of zero would indicate no difference, while a negative delta score would reflect slower RTs in the condition that should benefit from temporal preparation.

A series of Bayesian one-sample t-tests on delta scores were conducted to test whether temporal preparation effects were present at the group-level (different from zero). Moreover, a Bayesian repeated-measure ANOVA on the different delta scores was used to directly contrast the magnitude of temporal preparation effects.

Correlation Analysis between Temporal Preparation Effects

A series of Bayesian correlations were performed to examine the relationship between the different temporal preparation effects (measured as delta scores).

Test-Retest Reliability Analysis

⁵ We also computed a regression slope across the four foreperiod durations. In line with the delta score results (see below), we found strong evidence for a negative slope in the foreperiod effect (BF₁₀= 1.60E+17, Cohen's d=1.08), indicating a robust group-level effect. However, for consistency with the other temporal preparation effects, where a positive delta score was expected, we opted to report the delta score rather than the regression slope for the foreperiod effect as well.

Test-retest reliability was assessed using Bayesian analyses for each temporal preparation effect (foreperiod, temporal orienting, sequential, auditory, and visual rhythms). Separate Bayesian repeated-measure ANOVAs were conducted with the factors Session (session 1 vs. session 2) and Temporal Preparation Condition (temporally expected vs. temporally unexpected trials, or varying foreperiods) to determine whether temporal preparation effects were replicated in the second session.

In addition, following Hedge et al. (2018) and Pennington et al. (2025), we calculated the Intraclass Correlation Coefficient (ICC) using a two-way mixed-effects model for absolute agreement, which accounts for various sources of variance separately.

$$ICC = \frac{\text{Variance between participants}}{\text{Variance between participants} + \text{Error variance} + \text{Variance between sessions}}$$

We interpreted ICC according to Koo and Li's (2016) guidelines: excellent (> 0.90), good (0.75–0.90), moderate (0.50–0.75), and poor (< 0.50) reliability. All estimates were computed using R. For each measure, we also report the Spearman's rho correlation, as a further estimate of test-retest reliability (Pennington et al., 2025). Table 3 reports the test-retest reliabilities for each temporal manipulation and its respective measures (e.g., the valid and neutral conditions for the temporal orienting effect, as well as the delta score calculated as described above). To be consistent with the literature, which typically uses a raw RT difference, we also included this measure in the test-retest reliability analysis.

Finally, following recent recommendations to compare traditional frequentist and Bayesian approaches for estimating test-retest reliability (Snijder et al., 2023), we employed a hierarchical Bayesian model (HBM) to assess the reliability of RT difference scores. Traditional approaches, such as ICCs, typically treat summary measures as representative indicators of performance and do not account for trial-to-trial variability. In contrast, hierarchical methods incorporate this variability by modeling performance at the trial level

(Rouder & Haaf, 2019). The HBM framework was applied to RT difference scores derived from each temporal preparation effect (Table 3). Test-retest reliability was quantified as a Bayesian correlation between the test and retest phase estimates $r(\Delta 1, \Delta 2)$. All model parameters (three chains of 3000 iterations after 1000 warm-up iterations) were estimated with Stan (Stan Development Team, 2020b) through an interface in R, called RStan (Stan Development Team, 2020a), following the procedure of Snijder et al. (2023). As commonly accepted guidelines for interpreting Bayesian test-retest correlations are not available, the thresholds proposed by Koo and Li's (2016) for ICCs were used.

Results

Descriptive results and Internal Consistency

The median completion time for the tasks was 32 min (SD = 1.85 min, range = 26-38 min). Descriptive statistics and internal consistency estimates are provided in Table 1⁶. As anticipated for a simple detection task, accuracy was very high (mean = 0.96, SD = 0.03). Strong positive correlations between RTs across the different blocks (r : from 0.72 to 0.85, BF_{10} : from 7.91E+15 to 1.66 E+28) confirmed high consistency in participants' overall processing speed across task contexts. Regarding split-half reliability analyses, Spearman-Brown corrected estimates indicated good to excellent internal consistency (>0.80-0.90) for each individual condition. In contrast, the reliability of the RT difference scores used to quantify temporal preparation effects was generally lower than that of their constituent measures, being good for the temporal orienting effect (0.77), moderate for the foreperiod effect (0.54) and the auditory rhythmic effect (0.52), and poor for the visual rhythmic effect (0.41) and sequential effects (0.20).

⁶ Table 1 also reports the skewness and kurtosis coefficients of the RT distributions. Of note, the Bayesian analyses of temporal preparation effects did not change after applying a log-transformation to improve normality.

Table 1. Descriptive statistics and internal consistency estimates assessed using split-half reliability indexes corrected with the Spearman-Brown formula. Measures marked with an asterisk indicate those used to compute foreperiod, sequential, and temporal orienting effects, calculated as a Reaction Time (RT) difference score between conditions. Note that in the temporal orienting block, only valid trials (400 and 1600 ms) were included; however, short (400 ms) and long (1600 ms) trials from the neutral block are reported for completeness, as short-neutral trials were used to compute the temporal orienting effect (RT difference between short-neutral and short-valid trials). CI stands for confidence interval.

Temporal block		Mean	SD	Skewness	Kurtosis	Min	Max	Split-half corrected reliability [95% CI]	
Foreperiod	400 ms*	464.89	63.17	0.84	0.41	364.63	679.22	0.88 [0.82, 0.92]	
	800 ms	433.94	53.18	1.03	2.06	348.27	660.41	0.82 [0.74, 0.89]	
	1200 ms	422.61	56.09	1.28	1.65	349.83	636.41	0.86 [0.80, 0.90]	
	1600 ms*	414.84	50.66	1.11	1.84	331.81	601.75	0.83 [0.76, 0.88]	
	RT diff	50.066	46.42	0.36	0.97	-76.17	188.06	0.54 [0.37, 0.67]	
Neutral	n _{400 ms}	n-1 _{400ms} *	445.51	58.36	0.98	0.69	347.60	635.62	0.90 [0.86, 0.93]
	n-1 _{1600ms} *	466.77	57.73	0.85	0.86	362.53	674.20	0.87 [0.83, 0.91]	
	n _{1600 ms}	n-1 _{400ms}	422.14	58.95	1.33	2.36	339.56	654.28	0.90 [0.86, 0.93]
	n-1 _{1600ms}	416.64	53.71	1.20	1.65	337.40	610.85	0.92 [0.89, 0.95]	
	RT diff	21.26	28.36	0.46	1.42	-69.81	118.06	0.20 [0, 0.30]	
Temporal orienting	400 ms	Valid*	400.16	75.39	0.93	1.51	256.09	685.15	0.92 [0.89, 0.94]
		Neutral*	456.15	56.29	0.90	0.47	360.25	628.25	0.94 [0.91, 0.95]
	1600 ms	Valid	417.50	62.85	1.21	2.06	320.12	658.12	0.91 [0.85, 0.93]
		Neutral	419.88	53.66	1.17	1.86	347.36	629.55	0.95 [0.93, 0.97]
	RT diff	55.99	52.40	0.37	0.04	-67.95	207.98	0.77 [0.67, 0.83]	
Auditory rhythm	Regular	397.35	49.24	0.56	0.88	268.46	555.56	0.90 [0.86, 0.94]	
	Irregular	408.18	49.72	0.94	2.02	283.20	585.62	0.91 [0.88, 0.94]	
	RT diff	10.83	30.98	-0.18	2.01	-109.19	95.61	0.52 [0.34, 0.65]	
Visual rhythm	Regular	421.02	67.06	0.87	1.09	254.21	625.18	0.93 [0.90, 0.95]	
	Irregular	444.79	66.77	0.45	0.17	299.85	638.86	0.90 [0.87, 0.93]	
	RT diff	23.77	36.74	-0.68	2.09	-122.85	102.22	0.41 [0.22, 0.57]	

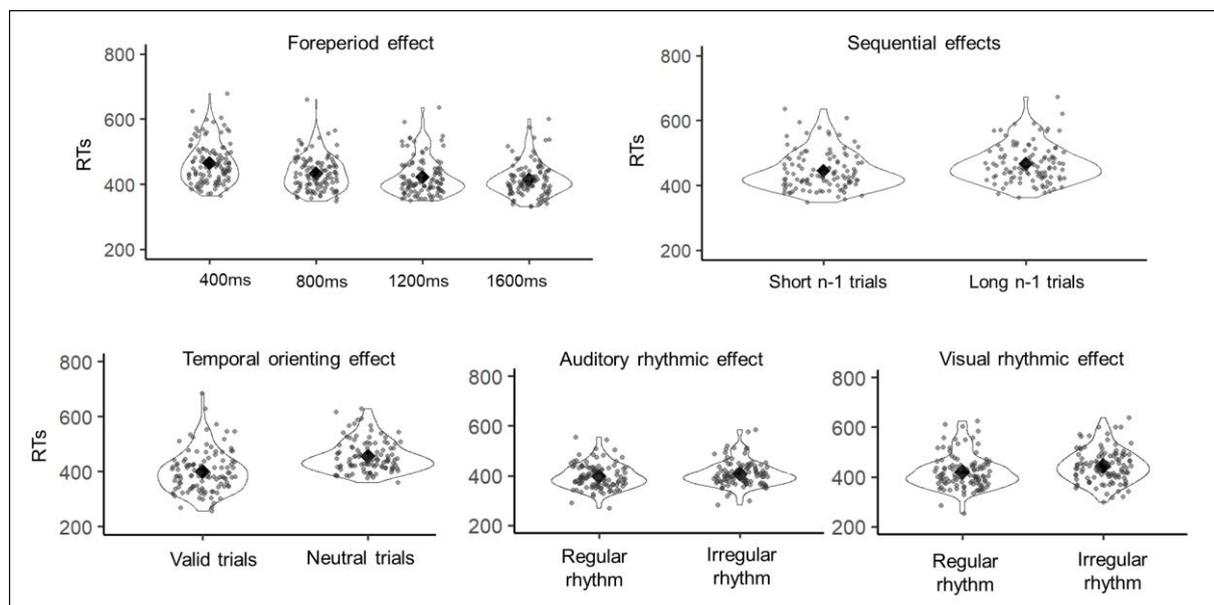
Note. The SD of Skewness for 109 participants corresponds to 0.23 and the SD of Kurtosis to 0.46.

Temporal preparation effects

Temporal preparation effects are illustrated in Figure 2. For the foreperiod effect, the analysis showed strong evidence ($BF_{10} = 2.41E+27$, $n^2p = 0.36$) supporting progressively faster RTs

with longer foreperiods ($BF_{10} > 5.85$, Cohen's $d > 0.67$), except between the two last foreperiods (1200 and 1600 ms; $BF_{10} = 1.94$, Cohen's $d = 0.24$). For the sequential effects, strong evidence was found for faster RTs on trials preceded by a short foreperiod compared to those preceded by a long foreperiod ($BF_{10} = 2.22E+9$, Cohen's $d = 0.75$). Similarly, temporal orienting was indexed by faster RTs for valid versus neutral trials ($BF_{10} = 4.66E+16$, Cohen's $d = 1.07$). Finally, we observed strong evidence supporting rhythmic effects, with faster RTs for regular compared to irregular trials in both the auditory ($BF_{10} = 48.92$, Cohen's $d = 0.35$) and visual modalities ($BF_{10} = 1.30E+7$, Cohen's $d = 0.65$).

Figure 2. Violin plots depicting the distribution of Reaction times (RTs) for each temporal preparation effect. Inside each plot, the central black diamond represents the mean and the dots represent individual data points.



Magnitude of temporal preparation effects

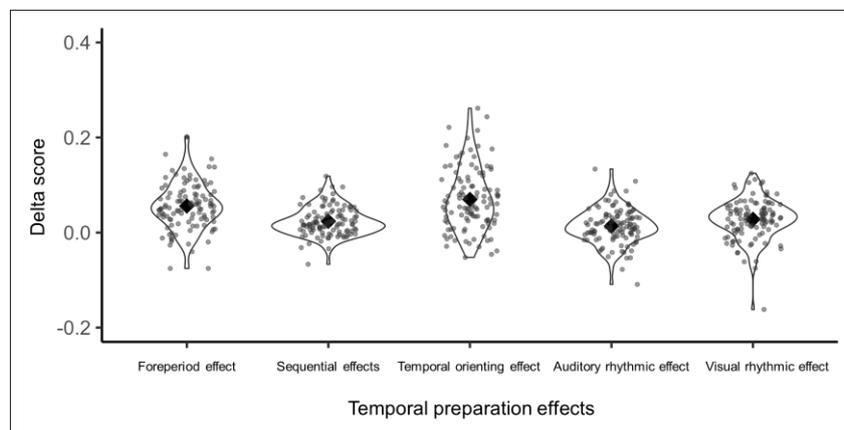
In terms of individual sensitivity of RTs to variations in temporal predictability, we found strong evidence for the foreperiod effect ($BF_{10} = 1.60E+17$, Cohen's $d = 1.08$), indicating a robust group-level effect, with 88% of participants exhibiting a positive delta score. Decisive

evidence also emerged for sequential effects ($BF_{10}= 3.24E+10$, Cohen's $d=0.79$) and temporal orienting ($BF_{10}= 5.99E+16$, Cohen's $d=1.06$), with 79% and 86% of participants showing positive delta scores, respectively. Regarding rhythmic effects, we observed very strong evidence in both the auditory ($BF_{10}= 180.58$, Cohen's $d=0.37$) and visual modalities ($BF_{10}= 4.57E+7$, Cohen's $d=0.66$), with 67% and 79% of the participants showing positive deltas, respectively. Overall, 54% of participants exhibited positive effects across all temporal preparation measures; 39% of participants showed positive effects in three of them; 7% showed positive effects in only two; and no participant displayed positive effects in just one measure. Among those showing two effects, no systematic pattern emerged that would indicate a strong relationship between specific temporal preparation pairings.

There was compelling evidence for a differential magnitude of temporal preparation effects ($BF_{10}=1.54E+19$, $n^2p=0.20$; see Figure 3). Planned post-hoc comparisons showed strong evidence for a larger temporal orienting effect compared to all other effects (sequential effects: $BF_{10}= 3.77E+6$, Cohen's $d=0.62$; auditory rhythmic effect: $BF_{10}= 2.53E+9$, Cohen's $d=0.75$; visual rhythmic effect: $BF_{10}= 1.21E+5$, Cohen's $d=0.55$), except for the foreperiod effect ($BF_{10}= 0.48$, Cohen's $d=0.17$). Additionally, we found decisive evidence that the foreperiod effect was larger than sequential effects ($BF_{10}= 7.34E+4$, Cohen's $d=0.54$), the auditory rhythmic effect ($BF_{10}= 1.89E+7$, Cohen's $d=0.66$), and the visual rhythmic effect ($BF_{10}= 694.35$, Cohen's $d=0.42$), while only anecdotal evidence was observed for its difference with the temporal orienting effect. We also found anecdotal evidence for differences between sequential and rhythmic effects (auditory: $BF_{10}= 1.70$, Cohen's $d=0.23$; visual: $BF_{10}= 0.15$, Cohen's $d=0.08$), as well as between the auditory and visual rhythmic effects ($BF_{10}= 1.86$, Cohen's $d=0.24$).

Figure 3. Violin plots depicting the distribution of the delta scores for each temporal preparation effect. Inside each plot, the central black diamond represents the mean, and the

dots indicate individual data points. A positive delta score reflects a temporal preparation effect, with faster RT in the condition associated with better temporal preparation. A score of zero indicates no difference, while a negative delta score shows slower RTs in the condition expected to benefit from temporal preparation.



Correlations between temporal preparation effects

We found no evidence for correlations between the different temporal preparation effects, including between visual and auditory rhythmic effects (Table 2). All correlation coefficients (r) were below 0.18, and all BF_{10} values were less than 0.67, indicating moderate evidence in favor of the null hypothesis (i.e., lack of correlation) for most of the correlations. In a few cases, the evidence was anecdotal, suggesting only weak support for a lack of correlation. Note that the correlation results remained unchanged when using RT differences instead of delta scores.

Table 2. Correlations between temporal preparation effects calculated as delta scores.

Temporal preparation effects	1	2	3	4
1. Foreperiod effect	--			
2. Sequential effects	$r = -.09$ (0.18)	--		
3. Temporal orienting effect	$r = .01$ (0.12)	$r = -.07$ (0.16)	--	
4. Auditory rhythmic effect	$r = -.09$ (0.18)	$r = .18$ (0.67)	$r = .03$ (0.13)	--
5. Visual rhythmic effect	$r = -.03$ (0.12)	$r = -.09$ (0.19)	$r = .05$ (0.14)	$r = -.16$ (0.48)

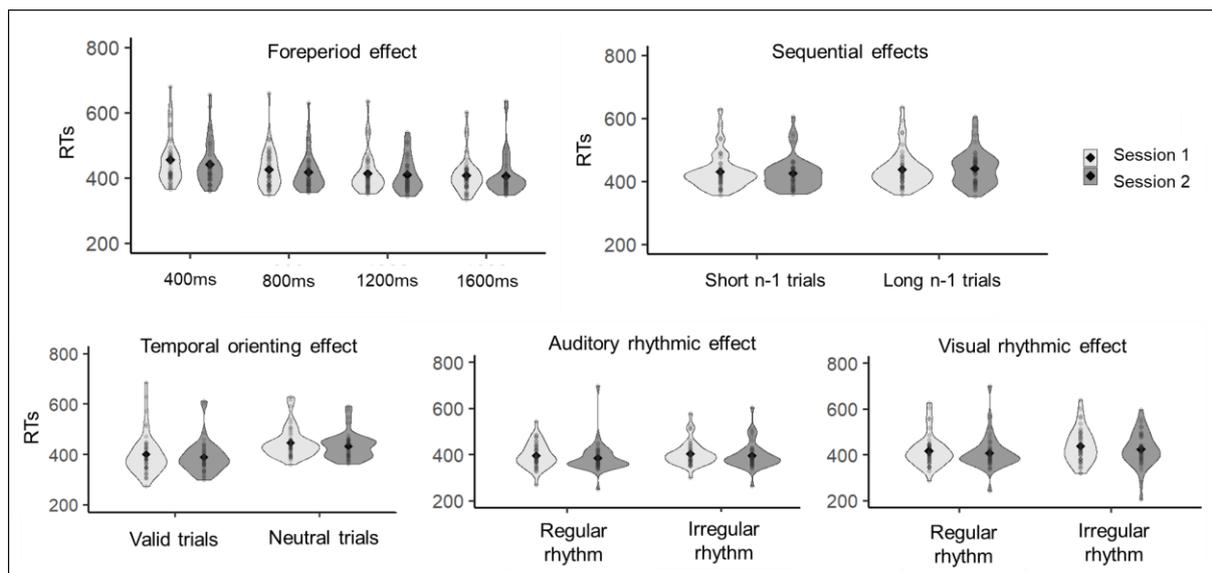
Note. Correlation coefficient (r) and Bayes Factor (BF_{10}) in brackets.

Test-retest reliability

Before conducting the test-retest analyses, we assessed the internal consistency of each session using the same procedure applied to the full sample (see Supplementary Table 1). In line with those results, we observed good to excellent internal consistency for each individual condition in both sessions, whereas the reliability of the RT difference scores ranged from good to poor.

For each temporal preparation effect (Figure 4), the Bayesian repeated-measure ANOVA showed no evidence supporting a main effect of Session (foreperiod effect: $BF_{10}=0.36$, $n^2p=0.02$; temporal orienting effect: $BF_{10}=1.56$, $n^2p=0.10$; auditory rhythmic effect: $BF_{10}=1.01$, $n^2p=0.07$; visual rhythmic effect: $BF_{10}=0.45$, $n^2p=0.03$), except for moderate evidence in the case of sequential effects ($BF_{10}=7.62$, $n^2p=0.49$), with faster RTS observed in the second session. Strong evidence in favor of a main effect of the Temporal manipulation was observed for all effects (foreperiod effect: $BF_{10}=3.73E+12$, $n^2p=0.44$; sequential effects: $BF_{10}=1.80E+4$, $n^2p=0.39$; temporal orienting effect: $BF_{10}=9.84E+6$, $n^2p=0.58$; visual rhythmic effect: $BF_{10}=866.41$, $n^2p=0.36$), except for the auditory rhythmic effect, which showed only moderate evidence ($BF_{10}=2.54$, $n^2p=0.13$). Finally, anecdotal evidence was found for an interaction between Session and Temporal Preparation effects (foreperiod effect: $BF_{10}=0.10$, $n^2p=0.02$; sequential effects: $BF_{10}=0.29$, $n^2p=0.01$; temporal orienting effect: $BF_{10}=0.21$, $n^2p=0.01$; auditory rhythmic effect: $BF_{10}=0.28$, $n^2p=0.01$; visual rhythmic effect: $BF_{10}=0.33$, $n^2p=0.01$).

Figure 4. Violin plots depicting the distribution of Reaction times (RTs) for each temporal preparation effect for both sessions. Inside each violin plot, the central black diamond represents the mean and the dots represent individual data.



As shown in Table 3, the single RT measures generally demonstrated moderate to poor test-retest reliability, whereas delta scores and raw RT differences showed poor reliability, as indicated by their low ICC values. The HBM analysis improved test-retest estimates for the foreperiod ($r= 0.65$), temporal orienting ($r= 0.64$), and auditory rhythmic effects ($r= 0.60$). Instead, reliability remained poor for sequential and visual rhythmic effects (< 0.50). In contrast, Spearman's rho correlations indicated generally moderate consistency in participants' rank ordering across sessions. Overall, this pattern suggests that, despite limited absolute reliability, individuals tended to maintain their relative performance levels between sessions.

Table 3. Test-retest analyses. ICC=Intraclass Correlation Coefficients. HBM= Hierarchical Bayesian Model. Significant correlations calculated with Spearman's rho are outlined in bold.

Temporal block	Trial Type	ICC (95% CI)	Spearman rho (95% CI)	HBM Pearson's correlation (BF ₁₀)
Foreperiod	400 ms	0.32 [0.15; 0.89]	0.60 [0.31; 0.79]	
	800 ms	0.45 [0.22; 0.90]	0.66 [0.41; 0.83]	
	1200 ms	0.65 [0.32; 0.95]	0.64 [0.39; 0.82]	
	1600 ms	0.47 [0.22; 0.91]	0.61 [0.36; 0.78]	
	RT diff	0.07 [0.06; 0.74]	0.23 [-0.11; 0.51]	0.65 (4643.48)
	Delta score	0.07 [0.06; 0.73]	0.20 [-0.13; 0.48]	
Neutral	N-1 _{400 ms}	0.40 [0.25; 0.92]	0.61 [0.35; 0.80]	
	N-1 _{1600 ms}	0.34 [0.25; 0.91]	0.72 [0.53; 0.85]	
	RT diff	0.22 [0.13; 0.80]	0.24 [-0.08; 0.51]	0.17 (0.35)
	Delta score	0.17 [0.12; 0.75]	0.21 [-0.11; 0.48]	
Temporal orienting	Valid	0.52 [0.28; 0.92]	0.61 [0.37; 0.79]	
	Neutral	0.37 [0.27; 0.92]	0.71 [0.50; 0.83]	
	RT diff	0.39 [0.18; 0.82]	0.49 [0.21; 0.69]	0.64 (3054.93)
	Delta score	0.33 [0.15; 0.82]	0.47 [0.18; 0.67]	
Auditory rhythm	Regular	0.28 [0.13; 0.82]	0.73 [0.54; 0.85]	
	Irregular	0.53 [0.31; 0.92]	0.72 [0.50; 0.86]	
	RT diff	0.14 [0.10; 0.76]	0.13 [-0.18; 0.43]	0.60 (1590.91)
	Delta score	0.11 [0.09; 0.73]	0.10 [-0.21; 0.40]	
Visual rhythm	Regular	0.56 [0.33; 0.93]	0.77 [0.57; 0.88]	
	Irregular	0.67 [0.32; 0.92]	0.69 [0.47; 0.85]	
	RT diff	0.19 [0.11; 0.76]	0.25 [-0.06; 0.50]	0.16 (0.31)
	Delta score	0.16 [0.11; 0.76]	0.20 [-0.12; 0.49]	

Discussion

It is well-established that target detection is faster when the timing of target onset can be anticipated. The benefit of temporal preparation on response readiness has been demonstrated across various experimental protocols, including foreperiod tasks, temporal orienting tasks, and rhythmic tasks (Capizzi & Correa, 2018). As highlighted in several recent reviews (Denison, 2024; Nobre & van Ede, 2018; Seibold et al., 2023), despite these tasks sharing a similar basic structure – where either a single warning signal/cue or a sequence of events helps anticipate the timing of target onset – temporal preparation is not a unitary construct. Developing a task that can simultaneously and independently assess temporal preparation effects within a single paradigm is crucial for advancing our understanding of temporal preparation. Successfully achieving this goal, we derived a robust measure of each temporal preparation effect within a single 35-min task. This encompassed behavioural advantages for longer compared to shorter foreperiods (the foreperiod effect), previous short foreperiods

compared to previous long foreperiods (sequential effects), predictive valid cues compared to neutral cues (temporal orienting effects), and regular sequences compared to irregular sequences (rhythmic effects).

In terms of magnitude of temporal preparation effects, we observed stronger temporal orienting and foreperiod effects, with no differences found between sequential and rhythmic effects, or between auditory and visual rhythmic effects. Moreover, our data showed that not all participants exhibited temporal preparation effects. Specifically, only 54% and 39% of participants demonstrated all or three temporal preparation effects, respectively, while a minority (7%) exhibited only two. These findings support the notion that temporal preparation is a multifaceted ability that may be deployed through different mechanisms. Future research could explore whether these patterns reflect individual preferences or strategic differences.

Focusing on individual differences, Spearman correlations indicated some consistency in performance between the two sessions. However, in line with the “reliability paradox” (Haines et al., 2020; Hedge et al., 2018; Rouder & Haaf, 2019), Intraclass Correlation Coefficients (ICC) revealed poor test–retest reliability for temporal preparation effects expressed as RT differences across conditions, while reliability was comparatively higher for raw RT values. Notably, within a single session, internal consistency was good to excellent for individual conditions, ranging from good to poor for RT difference scores. This discrepancy, where internal consistency estimates (e.g., split-half indices) exceeds test-retest estimates (e.g., ICC) is expected, since the former are calculated on a single timepoint, while the latter require stability across sessions. Accordingly, internal consistency and test-retest reliability should not be treated as interchangeable but rather reported together (see Snijder et al., 2023). To further examine test-retest reliability, we applied a hierarchical Bayesian model (HBM), which has been shown to improve traditional test-retest reliability estimates in the domain of cognitive control (Snijder et al., 2023). For RT difference measures of temporal

preparation, HBMs yielded higher test–retest estimates for foreperiod, temporal orienting, and auditory rhythmic effects, but only negligible improvements for sequential and visual rhythmic effects. Taken together, the results point to generally poor to moderate test-retest reliability of temporal preparation effects derived from RT differences, consistent with previous reports on other cognitive difference measures, such as the Flanker effect (Haines et al., 2020). As noted in previous studies (Haines et al., 2020; Hedge et al., 2018; Snijder et al., 2023), the reliability of difference scores is limited by the design of experimental tasks, which are typically intended to maximize within-subject variance across conditions while minimizing between-subject variance. As a result, subtracting scores across conditions can reduce between-participant variability, thus leading to lower reliability. In contrast to the literature on cognitive control, the field of temporal preparation lacks a comprehensive psychometric evaluation against which to compare our results. More broadly, reliability is rarely reported in cognitive experimental research (Parson et al., 2019; see also Garre-Frutos et al., 2024), such that our task contributes to addressing this gap. Although a larger sample size in the second session would strengthen the assessment of test–retest reliability, the present study already represents a comprehensive initial step by reporting indices that are rarely included in experimental research (i.e., internal consistency estimates, traditional ICCs, and Bayesian approaches to test-retest reliability), as recommended in recent work (Snijder et al., 2023). Importantly, the weak test-retest reliability of temporal preparation effects expressed as difference scores does not imply that the TEP-Task is unsuitable for between-group designs. Rather, it suggests that, consistent with other well-established cognitive tasks, the TEP-Task may not reliably capture stable individual differences within a population. At the same time, the HBM results indicate that certain temporal preparation effects, such as foreperiod, temporal orienting, and auditory rhythmic effects, may show comparatively higher reliability. However, this conclusion should be interpreted with caution and examined further

in future research to clarify both the commonalities and the differences across temporal preparation effects.

Another finding of our study was the absence of significant correlations between temporal preparation effects. At first, this may seem counterintuitive, considering prior evidence that some temporal preparation effects could be mediated by shared processes. For instance, it has been proposed that while temporal orienting and foreperiod effects depend on more strategic and controlled processes reliant on frontal regions (Triviño et al., 2010; 2011), sequential and rhythmic effects are more automatic and dependent on parietal and subcortical structures (Vallesi et al., 2007a). This view, suggesting a controlled–automatic continuum for temporal preparation effects, is further supported by behavioural studies showing that only temporal orienting and foreperiod effects, but not sequential or rhythmic effects, are sensitive to concurrent working memory demands (Capizzi et al., 2012; 2013; Cutanda et al., 2015; de la Rosa et al., 2012; Vallesi et al., 2014). Therefore, one might expect significant correlations between the more controlled temporal orienting and foreperiod effects, on the one hand, and between the more automatic sequential and rhythmic effects, or even between auditory and visual rhythmic tasks (see also Attout et al., 2024).

As highlighted in the Introduction, it is possible that previous task manipulations may have favored the relationships between temporal preparation effects, rather than stressing their independence (Correa et al., 2010; Triviño et al., 2010). Another possibility is that the weak reliability of RT difference scores may have obscured true underlying correlations. Although this concern is mitigated by our sample size, which should provide sufficient statistical power for correlational analyses (e.g., Hedge et al., 2018), it is important to note that the observed Bayes factors did not provide conclusive evidence for the absence of a relationship across all of our measures.

At a preliminary level, our correlational findings challenge temporal preparation models assuming that foreperiod and sequential effects stem from a common implicit learning mechanism (Los et al., 2014; see also Salet et al., 2022, for computational modeling). According to such models, both effects would emerge from the joint activation of memory traces formed on a trial-by-trial basis. This view contrasts with dual-process models (Vallesi et al. 2007b), which propose that foreperiod and sequential effects rely on distinct mechanisms. Overall, our findings are more in line with the latter framework, supporting the idea that foreperiod and sequential effects may arise from separable processes. Regarding rhythms, our correlation results also suggest independent and modality-specific effects of rhythms on temporal preparation. This finding replicates our previous results, which showed no correlation between visual and auditory rhythms (Attout et al., 2024).

Deepening our understanding of whether temporal preparation effects are mediated by shared or distinct mechanisms is essential for developing more precise cognitive and computational models of temporal preparation. Our task may serve as a valuable starting point for this endeavor. For example, future studies could expand its utility by exploring alternative temporal manipulations, such as replacing neutral trials with invalid trials in the temporal orienting condition, introducing different foreperiod distributions (e.g., non-aging) or designs (e.g., fixed foreperiods), or incorporating response selection (e.g., choice RT tasks) or increased perceptual demands. Indeed, one potential limitation of our study is the use of a simple RT task, which might have limited our ability to detect more nuanced relationships between temporal preparation effects. While we acknowledge this point, we deliberately began with a simple design to establish a reliable and replicable task that could later be adapted to more demanding cognitive contexts. A more complex design would also have increased task duration, which was another key aspect in our design. Moreover, for temporal orienting effects, there is evidence that these can be attenuated in choice RT tasks (Correa et

al., 2004) and that more complex perceptual tasks make it harder to predict how different temporal preparation effects relate (Seibold et al., 2023).

A final potential limitation of our task is the relatively low number of trials, which may reduce its sensitivity particularly in future clinical studies, where a larger number of trials might be necessary. Mitigating this concern, split-half reliability analyses for each block indicated good to excellent internal consistency across individual conditions. From a practical standpoint, the limited number of trials reflects the deliberate goal of designing a brief and user-friendly task. However, given the absence of practice effects across sessions, it may be feasible to repeat the task to increase the number of trials when needed. Moreover, since the task is openly available on OSF, researchers can readily adapt it to include more trials within a single session.

In conclusion, the TEP-Task enables the simultaneous assessment of the most common temporal preparation effects within a single session and can be easily customized for diverse populations and task demands.

Open practices statement

All code, data, and materials mentioned in this article are publicly available at <https://osf.io/rnc4v>. The study was not preregistered.

Ethics declarations

Conflicts of interest

The authors declare no conflict of interest in connection with this work.

Ethics approval

The study was approved by the ethics committee of the University Paul Valéry 3 of Montpellier.

Consent to participate

Informed consent was obtained from all individual participants included in the study.

Consent for publication

All authors gave their consent to publish this manuscript.

Author contribution

M.C.: Conceptualization, Methodology, Software, Formal Analysis, Supervision, Funding acquisition, Writing – original draft, Writing – review & editing. L.A.: Investigation, Methodology, Software, Formal Analysis, Visualization, Writing – original draft; Writing – review & editing. G.M.: Methodology, Software, Writing – review & editing. P.C.: Conceptualization, Methodology, Software, Formal Analysis, Visualization, Supervision, Resources, Funding acquisition, Writing – review & editing.

Acknowledgements

We thank Sara Errigo (University of Padova) for her assistance in programming the task.

Funding information

This work was supported by a grant (PID2021-128696NA-I00) funded by MICIU/AEI/10.13039/501100011033 and ERDF/EU to M.C., and by a grant from the Agence National de Recherche (ANR-18-CE28-0009-01) to P.C. It also received support from grant CEX2023-001312-M, funded by MICIU/AEI/10.13039/501100011033 and grant UCE-PP2023-11 funded by University of Granada. M.C. also acknowledges support of a María Zambrano Fellowship at the University of Granada from the Spanish Ministry of Universities and the European Union NextGeneration.

References

- Attout, L., Capizzi, M., & Charras, P. (2024). Enhancing rhythmic temporal expectations: The dominance of auditory modality under spatial uncertainty. *Attention, Perception & Psychophysics*, 86(5), 1681–1693. <https://doi.org/10.3758/s13414-024-02898-3>
- Ball, F., Andreca, J., & Noesselt, T. (2022). Context dependency of time-based event-related expectations for different modalities. *Psychological Research*, 86(4), 1239–1251. <https://doi.org/10.1007/s00426-021-01564-9>
- Bausenhart, K. M., Rolke, B., & Ulrich, R. (2008). Temporal preparation improves temporal resolution: evidence from constant foreperiods. *Perception & Psychophysics*, 70(8), 1504–1514. <https://doi.org/10.3758/PP.70.8.1504>
- Berkhout, S.W., Haaf, J.M., Gronau, Q.F. et al. A tutorial on Bayesian model-averaged meta-analysis in JASP. *Behavior Research Methods*, 56, 1260–1282 (2024). <https://doi.org/10.3758/s13428-023-02093-6>
- Breska, A., & Deouell, L. Y. (2014). Automatic bias of temporal expectations following temporally regular input independently of high-level temporal expectation. *Journal of Cognitive Neuroscience*, 26(7), 1555–1571. https://doi.org/10.1162/jocn_a_00564
- Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease. *Proceedings of the National Academy of Sciences of the United States of America*, 115(48), 12283–12288. <https://doi.org/10.1073/pnas.1810596115>
- Capizzi, M., & Correa, Á. (2018). Measuring temporal preparation. In A. Vatakis, F. Balci, M. Di Luca, & Á. Correa (Eds.). *Timing and time perception: Procedures, measures, and applications*. Ed. Brill, Leiden Boston, pp. 216–232.
- Capizzi, M., Correa, A., & Sanabria, D. (2013). Temporal orienting of attention is interfered by concurrent working memory updating. *Neuropsychologia*, 51(2), 326–339. <https://doi.org/10.1016/j.neuropsychologia.2012.10.005>
- Capizzi, M., Correa, A., Wojtowicz, A., & Rafal, R. D. (2015). Foreperiod priming in temporal preparation: testing current models of sequential effects. *Cognition*, 134, 39–49. <https://doi.org/10.1016/j.cognition.2014.09.002>
- Capizzi, M., Martín-Signes, M., Coull, J. T., Chica, A. B., & Charras, P. (2023). A transcranial magnetic stimulation study on the role of the left intraparietal sulcus in temporal orienting of attention. *Neuropsychologia*, 184, 108561. <https://doi.org/10.1016/j.neuropsychologia.2023.108561>
- Capizzi, M., Sanabria, D., & Correa, Á. (2012). Dissociating controlled from automatic processing in temporal preparation. *Cognition*, 123(2), 293–302. <https://doi.org/10.1016/j.cognition.2012.02.005>

Capizzi, M., Visalli, A., Faralli, A., & Mioni, G. (2022). Explicit and implicit timing in older adults: Dissociable associations with age and cognitive decline. *PLoS ONE*, *17*(3), e0264999. <https://doi.org/10.1371/journal.pone.0264999>

Charras, P., Chica, A. B., & Capizzi, M. (2023). On the relationship between rhythm-based temporal expectations and endogenous spatial attention in simple reaction-time tasks. *Frontiers in Cognition*, *2*, 1191595. <https://doi.org/10.3389/fcogn.2023.1191595>

Cohen, J., 1977. In: rev (Ed.), *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ.

Correa, A., Cona, G., Arbula, S., Vallesi, A., & Bisiacchi, P. (2014). Neural dissociation of automatic and controlled temporal preparation by transcranial magnetic stimulation. *Neuropsychologia*, *65*, 131–136. <https://doi.org/10.1016/j.neuropsychologia.2014.10.023>

Correa, A., Lupiáñez, J., Madrid, E., & Tudela, P. (2006). Temporal attention enhances early visual processing: a review and new evidence from event-related potentials. *Brain Research*, *1076*(1), 116–128. <https://doi.org/10.1016/j.brainres.2005.11.074>

Correa, A., Lupiáñez, J., Milliken, B., & Tudela, P. (2004). Endogenous temporal orienting of attention in detection and discrimination tasks. *Perception & Psychophysics*, *66*(2), 264–278. <https://doi.org/10.3758/bf03194878>

Correa, A., Triviño, M., Pérez-Dueñas, C., Acosta, A., & Lupiáñez, J. (2010). Temporal preparation, response inhibition and impulsivity. *Brain and cognition*, *73*(3), 222–228. <https://doi.org/10.1016/j.bandc.2010.05.006>

Coull J. T. (2009). Neural substrates of mounting temporal expectation. *PLoS Biology*, *7*(8), e1000166. <https://doi.org/10.1371/journal.pbio.1000166>

Coull, J., & Nobre, A. (2008). Dissociating explicit timing from temporal expectation with fMRI. *Current Opinion in Neurobiology*, *18*(2), 137–144. <https://doi.org/10.1016/j.conb.2008.07.011>

Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *The Journal of Neuroscience*, *18*(18), 7426–7435. <https://doi.org/10.1523/JNEUROSCI.18-18-07426.1998>

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *The Journal of Neuroscience*, *33*(9), 4002–4010. <https://doi.org/10.1523/JNEUROSCI.4675-12.2013>

Cutanda, D., Correa, Á., & Sanabria, D. (2015). Auditory temporal preparation induced by rhythmic cues during concurrent auditory working memory tasks. *Journal of Experimental Psychology. Human Perception and Performance*, *41*(3), 790–797. <https://doi.org/10.1037/a0039167>

Cutanda, D., Sanabria, D., & Correa, Á. (2019). Cognitive entrainment to isochronous rhythms is independent of both sensory modality and top-down attention. *Psicológica*, 40, 62-84. <https://doi.org/10.2478/psicolj-2019-0005>

de la Rosa, M. D., Sanabria, D., Capizzi, M., & Correa, A. (2012). Temporal preparation driven by rhythms is resistant to working memory interference. *Frontiers in Psychology*, 3, 308. <https://doi.org/10.3389/fpsyg.2012.00308>

Denison, R.N. (2024). Visual temporal attention from perception to computation. *Nature Review Psychology*, 3, 261–274. <https://doi.org/10.1038/s44159-024-00294-0>

Garre-Frutos, F., Vadillo, M. A., González, F., & Lupiáñez, J. (2024). On the reliability of value-modulated attentional capture: An online replication and multiverse analysis. *Behavior Research Methods*, 56(6), 5986–6003. <https://doi.org/10.3758/s13428-023-02329-5>

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, 84, 279-385.

Haines N., Kvam P.D., Irving L.H., Smith C., Beauchaine T.P., Pitt M.A., et al. (2020). Learning from the reliability paradox: How theoretically informed generative models can advance the social, behavioral, and brain sciences. *PsyArXiv*. <https://doi.org/10.31234/osf.io/xr7y3>

Han, T., & Proctor, R. W. (2022). Change of variable-foreperiod effects within an experiment: A Bayesian modeling approach. *Journal of Cognition*, 5(1), 40. <https://doi.org/10.5334/joc.235>

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, 50(3), 1166–1186. <https://doi.org/10.3758/s13428-017-0935-1>

Herbst, S. K., Fiedler, L., & Obleser, J. (2018). Tracking temporal hazard in the human electroencephalogram using a forward encoding model. *eNeuro*, 5(2), ENEURO.0017-18.2018. <https://doi.org/10.1523/ENEURO.0017-18.2018>

Janssen, P., & Shadlen, M. N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nature Neuroscience*, 8(2), 234–241. <https://doi.org/10.1038/nn1386>

Jeffreys, H. (1961). *Theory of probability*, 3rd Edn. Oxford: Oxford University Press.

Johnson, K. A., Bryan, M., Polonowita, K., Decroupet, D., & Coull, J. T. (2016). Isochronous sequential presentation helps children orient their attention in time. *Frontiers in Psychology*, 7, 1417. <https://doi.org/10.3389/fpsyg.2016.01417>

Jones A. (2015). Independent effects of bottom-up temporal expectancy and top-down spatial attention. An audiovisual study using rhythmic cueing. *Frontiers in Integrative Neuroscience*, 8, 96. <https://doi.org/10.3389/fnint.2014.00096>

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science*, *13*(4), 313–319. <https://doi.org/10.1111/1467-9280.00458>

Kingstone, A. (1992). Combining expectancies. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *44A*(1), 69–104. <https://doi.org/10.1080/14640749208401284>

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting Intraclass Correlation Coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review*, *106*(1), 119–159. <https://doi.org/10.1037/0033-295X.106.1.119>

Lawrence, M. A., & Klein, R. M. (2013). Isolating exogenous and endogenous modes of temporal attention. *Journal of Experimental Psychology: General*, *142*(2), 560–572. <https://doi.org/10.1037/a0029023>

Los, S. A., Kruijne, W., & Meeter, M. (2014). Outlines of a multiple trace theory of temporal preparation. *Frontiers in Psychology*, *5*, 1058. <https://doi.org/10.3389/fpsyg.2014.01058>

Los, S. A., & Schut, M. L. (2008). The effective time course of preparation. *Cognitive Psychology*, *57*(1), 20–55. <https://doi.org/10.1016/j.cogpsych.2007.11.001>

Los, S. A., & van den Heuvel, C. E. (2001). Intentional and unintentional contributions to nonspecific preparation during reaction time foreperiods. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(2), 370–386. <https://doi.org/10.1037//0096-1523.27.2.370>

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. <https://doi.org/10.3758/s13428-011-0168-7>

Mento, G., & Tarantino, V. (2015). Developmental trajectories of internally and externally driven temporal prediction. *PLoS ONE*, *10*(8), e0135098. <https://doi.org/10.1371/journal.pone.0135098>

Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods*, *16*(4), 406–419. <https://doi.org/10.1037/a0024377>

Morey, R. D., & Rouder, J. N. (2018). BayesFactor 0.9.12-4.2. Comprehensive R Archive Network. Retrieved from, <http://cran.r-project.org/web/packages/BayesFactor/index.html>.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin*, *89*(1), 133–162.

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: temporal structure in attention. *Nature Reviews. Neuroscience*, *19*(1), 34–48. <https://doi.org/10.1038/nrn.2017.141>

Olk, B. (2014). Effects of spatial, temporal and spatiotemporal cueing are alike when attention is directed voluntarily. *Experimental Brain Research*, 232(11), 3623–3633. <https://doi.org/10.1007/s00221-014-4033-7>

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>

Pennington, C. R., Birch-Hurst, K., Ploszajski, M., Clark, K., Hedge, C., & Shaw, D. J. (2025). Are we capturing individual differences? Evaluating the test-retest reliability of experimental tasks used to measure social cognitive abilities. *Behavior Research Methods*, 57(2), 82. <https://doi.org/10.3758/s13428-025-02606-5>

Rohenkohl, G., Coull, J. T., & Nobre, A. C. (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. *PLoS ONE*, 6(1), e14620. <https://doi.org/10.1371/journal.pone.0014620>

Rohenkohl, G., Cravo, A. M., Wyart, V., & Nobre, A. C. (2012). Temporal expectation improves the quality of sensory information. *The Journal of Neuroscience*, 32(24), 8424–8428. <https://doi.org/10.1523/JNEUROSCI.0804-12.2012>

Rolke, B., Festl, F., & Seibold, V. C. (2016). Toward the influence of temporal attention on the selection of targets in a visual search task: An ERP study. *Psychophysiology*, 53(11), 1690–1701. <https://doi.org/10.1111/psyp.12734>

Román-Caballero, R., Martín-Arévalo, E., Martín-Sánchez, P. D. C., Lupiáñez, J., & Capizzi, M. (2024). Influence of rhythmic contexts on perception: No behavioral and eye-tracker evidence for rhythmic entrainment. *Consciousness and Cognition*, 126, 103789. <https://doi.org/10.1016/j.concog.2024.103789>

Rouder, J. N., & Haaf, J. M. (2019). A psychometrics of individual differences in experimental tasks. *Psychonomic Bulletin and Review*, 26(2), 452–467. <https://doi.org/10.3758/s13423-018-1558-y>

Salet, J. M., Kruijne, W., van Rijn, H., Los, S. A., & Meeter, M. (2022). FMTP: A unifying computational framework of temporal preparation across time scales. *Psychological Review*, 129(5), 911–948. <https://doi.org/10.1037/rev0000356>

Sanabria, D., Capizzi, M., & Correa, A. (2011). Rhythms that speed you up. *Journal of Experimental Psychology. Human Perception and Performance*, 37(1), 236–244. <https://doi.org/10.1037/a0019956>

Schönbrodt, F. D., & Wagenmakers, E. J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin and Review*, 25(1), 128–142. <https://doi.org/10.3758/s13423-017-1230-y>

Seibold, V. C., Balke, J., & Rolke, B. (2023). Temporal attention. *Frontiers in Cognition*, 2:1168320. <https://doi.org/10.3389/fcogn.2023.1168320>

- Snijder, J. P., Tang, R., Bugg, J. M., Conway, A. R. A., & Braver, T. S. (2023). On the psychometric evaluation of cognitive control tasks: An Investigation with the Dual Mechanisms of Cognitive Control (DMCC) battery. *Behavior Research Methods*, *56*(3), 1604–1639. <https://doi.org/10.3758/s13428-023-02111-7>
- Steinborn, M. B., Rolke, B., Bratzke, D., & Ulrich, R. (2008). Sequential effects within a short foreperiod context: evidence for the conditioning account of temporal preparation. *Acta Psychologica*, *129*(2), 297–307. <https://doi.org/10.1016/j.actpsy.2008.08.005>
- Triviño, M., Arnedo, M., Lupiáñez, J., Chirivella, J., & Correa, A. (2011). Rhythms can overcome temporal orienting deficit after right frontal damage. *Neuropsychologia*, *49*(14), 3917–3930. <https://doi.org/10.1016/j.neuropsychologia.2011.10.009>
- Triviño, M., Correa, A., Arnedo, M., & Lupiáñez, J. (2010). Temporal orienting deficit after prefrontal damage. *Brain*, *133*(Pt 4), 1173–1185. <https://doi.org/10.1093/brain/awp346>
- Vallesi, A. (2010). Neuro-anatomical substrates of foreperiod effects. In: Nobre, A. C. & Coull, J.T (Eds.). *Attention and time*. Oxford University Press, pp. 303–316.
- Vallesi, A., Arbula, S., & Bernardis, P. (2014). Functional dissociations in temporal preparation: evidence from dual-task performance. *Cognition*, *130*(2), 141–151. <https://doi.org/10.1016/j.cognition.2013.10.006>
- Vallesi, A., Lozano, V. N., & Correa, A. (2013). Dissociating temporal preparation processes as a function of the inter-trial interval duration. *Cognition*, *127*(1), 22–30. <https://doi.org/10.1016/j.cognition.2012.11.011>
- Vallesi, A., Mussoni, A., Mondani, M., Budai, R., Skrap, M., & Shallice, T. (2007a). The neural basis of temporal preparation: Insights from brain tumor patients. *Neuropsychologia*, *45*(12), 2755–2763. <https://doi.org/10.1016/j.neuropsychologia.2007.04.017>
- Vallesi, A., Shallice, T., & Walsh, V. (2007b). Role of the prefrontal cortex in the foreperiod effect: TMS evidence for dual mechanisms in temporal preparation. *Cerebral Cortex*, *17*(2), 466–474. <https://doi.org/10.1093/cercor/bhj163>
- Visalli, A., Capizzi, M., Ambrosini, E., Kopp, B., & Vallesi, A. (2021). Electroencephalographic correlates of temporal Bayesian belief updating and surprise. *NeuroImage*, *231*, 117867. <https://doi.org/10.1016/j.neuroimage.2021.117867>
- Visalli, A., Capizzi, M., Ambrosini, E., Kopp, B., & Vallesi, A. (2023). P3-like signatures of temporal predictions: a computational EEG study. *Experimental Brain Research*, *241*(7), 1919–1930. <https://doi.org/10.1007/s00221-023-06656-z>
- Visalli, A., Capizzi, M., Ambrosini, E., Mazzonetto, I., & Vallesi, A. (2019). Bayesian modeling of temporal expectations in the human brain. *NeuroImage*, *202*, 116097. <https://doi.org/10.1016/j.neuroimage.2019.116097>