Temporal Preparation Drives Statistical Learning of Regular Intervals in Complex Temporal Environments

Josh M. Salet[1*], Wouter Kruijne[1], Hedderik van Rijn[1], Eckart Zimmermann[2], and Nadine Schlichting[2]

[1]Department of Experimental Psychology, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, the Netherlands

[2]Institute for Experimental Psychology, Heinrich-Heine-University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany

Author Note

*corresponding author's email: saletjm@gmail.com

Funding: During this project, Josh M. Salet, Wouter Kruijne, and Hedderik van Rijn were supported by the research program "Interval Timing in the Real World: A functional, computational and neuroscience approach", project number 453-16-005, awarded to Hedderik van Rijn, financed by the Netherlands Organisation for Scientific Research (NWO). Nadine Schlichting and Eckart Zimmermann were supported by the European Research Council (project "moreSense", grant agreement 757184).

Abstract

The ubiquitous finding in statistical learning research is a performance benefit for regular versus irregular properties of the environment. This regularity benefit is commonly interpreted to emerge from a mechanism that detects and subsequently exploits the regularity as it would be impossible for irregular events. Here, for regular intervals in a complex temporal environment, we show that this interpretation overlooks the influence of learning various time-related factors that contribute to the regularity benefit. Instead of only learning the temporal regularity, participants exploited the temporal properties of targets that we initially considered as irregular and uninformative. Participants temporally prepared, in parallel, for different actions to perform and locations to attend, irrespective of target's regularity. In fact, this temporal preparation explained regularity benefits without assuming statistical learning of the regular target. Using a computational model, *f*MTP, we illustrate that such adaptation can arise from associative memory processes underlying temporal preparation.

Keywords: Statistical Learning; Temporal Preparation; Implicit Learning; Associative Memory; Perception Action Coupling; Serial Reaction Time Task

Temporal Preparation Drives Statistical Learning of Regular Intervals in Complex Temporal Environments

1 Introduction

Statistical learning (SL) is a major theme in the cognitive sciences and refers to the learning processes underlying the extraction of statistical regularities in the environment to guide behavior (e.g., Saffran et al., 1996; Sherman et al., 2020). Studies on SL's temporal aspects investigate whether the environment's temporal statistics are extracted and used to prepare for when events happen. The ubiquitous finding across a range of paradigms is that, while participants are unaware of any temporal regularity, their performance is better for regular versus irregular timed properties in the experiment (e.g., lower reaction time or higher accuracy in language learning, Hay and Saffran, 2012; Selchenkova et al., 2014; serial reaction time tasks, Olson and Chun, 2001; O'Reilly et al., 2008; Shin and Ivry, 2002; rapid serial visual presentation tasks, Sali et al., 2015; singleton tasks, Xu et al., 2022). These regularity benefits are commonly interpreted to emerge from the ability to extract regular intervals from the environment (e.g., Baker et al., 2004; Boettcher et al., 2022; Fiser & Aslin, 2001; Salet et al., 2021; Sherman et al., 2020; Zhao et al., 2013).

SL's temporal aspects have particularly been studied in temporal variants of serial reaction time (SRT) tasks (Gobel et al., 2011; Heideman et al., 2016, 2018; Olson & Chun, 2001; O'Reilly et al., 2008; Salidis, 2001; Sanchez et al., 2015; Shin & Ivry, 2002, 2003). In classic SRT tasks (Nissen & Bullemer, 1987), participants make a corresponding keypress to each item in a stimulus sequence (e.g., stimuli A-B-A responded to with index-middle-index finger). In temporal versions of such SRT tasks, a regularity benefit is found when the intervals between successive stimuli contain a regular pattern (e.g., the intervals between A-B-A always follow short-long). That is, although participants were unaware of any regularity, their responses were faster and more accurate to sequences with a regular interval pattern compared to randomly timed sequences. This regularity benefit is only found, however, if the temporal order of the sequences is ordinal (e.g., stimuli are

always presented in fixed order A-B-A, O'Reilly et al., 2008; Sanchez et al., 2015; Shin & Ivry, 2002).

Recently, studies have cast these findings with respect to the demands of nature's complex temporal structure (Boettcher et al., 2022; Salet et al., 2021; Salet, de Jong, & van Rijn, 2022). In contrast to SRT tasks, regular interval patterns in naturalistic environments almost never happen in isolation, but are continuously interrupted by other behaviorally relevant events. In such settings, SL can not operate through a mechanism directly extracting the regular intervals between ordinal events as in SRT tasks. Instead, to extract regular intervals, a mechanism must monitor them over and beyond other intervening events. To test this, recent timing studies developed paradigms in which responses to targets from the regular sequence are continuously interrupted by equally relevant targets from the irregular sequences (Boettcher et al., 2022; Salet et al., 2021; Shalev et al., 2022). Importantly, the overall ordinal sequence in these paradigms, that is, the sequence of any subsequent target regardless of their location, contained no regularities - neither in its temporal order nor in the intervals between targets. Nevertheless, these studies found responses to regular compared to irregular interval sequences to be faster and more accurate. A regularity benefit interpreted to show that SL findings from laboratory settings can generalize to more complex temporal environments.

In this study, we assess whether this regularity benefit found in complex temporal settings (Salet et al., 2021) results from learning which location to attend or what action to prepare at regular onsets. We set out to answer this question in two preregistered experiments (https://osf.io/3d6h8; https://osf.io/3b78x) in which we coupled the regular sequence to a unique location or action. However, in contrast with our preregistered hypotheses, we found regularity benefits not to result from SL of regular locations or actions, but from a phenomenon known as temporal preparation (see also Addendum to Salet et al., 2021: Salet, Schlichting, et al., 2022). The reoccurring finding in temporal preparation studies is that longer intervals allow more time for preparation, resulting in

faster responses (Los et al., 2014; Niemi & Näätänen, 1981; Nobre & Ede, 2018). Similarly, we found performance to improve as more time passed since target's previous onset. This was observed irrespective of sequence's regularity and for both location and action, suggesting that participants were capable of monitoring the timing of three sequences in these two dimensions in parallel.

There have been similar reports regarding the involvement of temporal preparation in SL (Heideman et al., 2016, 2018; Salidis, 2001). These studies found that as temporal preparation is not yet fully developed at short intervals, performance benefits more from SL at shorter than longer intervals. Here, we show that such temporal preparation fully accounts for the regularity benefit without assuming SL of the regular target. In contrast with the common SL interpretation that this benefit arises from only learning the temporal properties of the regularity, our findings suggest that the benefit arises from learning multiple time-related factors inherent to the experimental design, both regular and irregular. By using a computational model (*f*MTP, Salet, Kruijne, van Rijn, et al., 2022), we illustrate how such complex, adaptive behavior can arise from memory processes underlying temporal preparation.

2 Materials and Methods

2.1 Preregistration

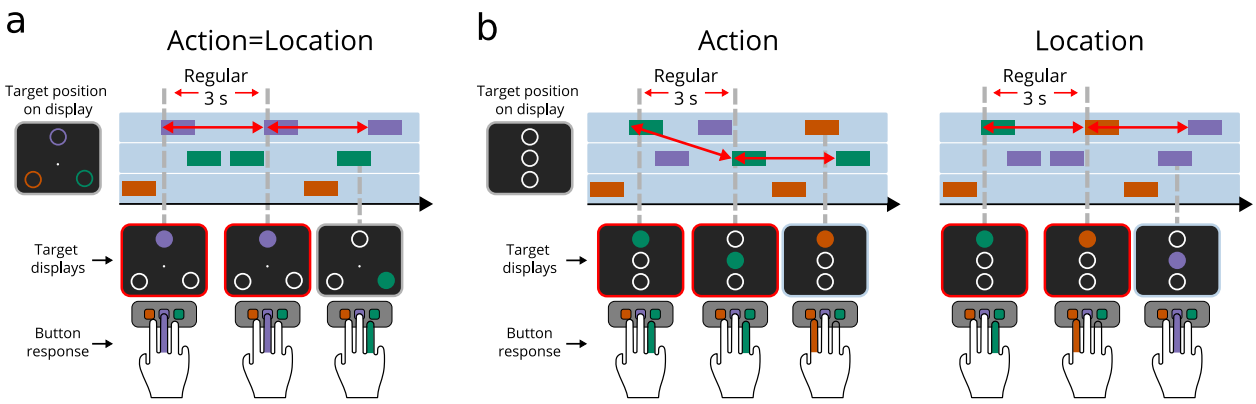
This study's planned sample size, included variables, hypotheses, and planned analyses were preregistered on Open Science Framework (OSF) prior to any data being collected (https://osf.io/3b78x).

2.2 Participants

59 First-year psychology students of the University of Groningen were recruited and received course credits for their participation. The experiment was approved by the Ethics committee of the Faculty of Behavioral and Social Sciences (PSY-1920-S-0360) and informed consent was given prior to participation. We continued data collection until we collected data from 54 participants meeting the inclusion criteria (see section 2.5). Sample

Figure 1

Experimental conditions.



Note. Segments of a series of target presentations in WAM. (a) In the 'action=location' condition, action and location were directly coupled and every 3000 ms a regular action was performed in a fixed location (red arrows). (b) In the 'action' and 'location' conditions, the direct mapping between location and action was removed. In the 'action' condition, the regular action was performed every 3000 ms in a random location. In contrast, in the 'location' condition, the regular location appeared every 3000 ms and was linked with a random action. The red arrows represent the regular interval, the other events appeared at pseudorandom time points. A video of the task is available at: https://osf.io/2ca8h/.

size rationale was based on Monte Carlo simulations for linear mixed models with the 'SIMR' package in R (Green & MacLeod, 2016), aiming for a 80% detection power of a regularity benefit (https://osf.io/3d6h8).

2.3 Stimuli and Task

Participants took part in an adapted version of the 'Whac-A-Mole' task (WAM, Salet et al., 2021), implemented in OpenSesame (Mathôt et al., 2012) using the PsychoPy backend (Peirce et al., 2019). Participants were seated in a dimly lit room at approximately 60 cm viewing distance from a 22" LCD monitor (100 Hz, Iiyama MA203DT) and aimed to score as many points as possible by reacting to sudden-onset targets with the correct button press. A video of the task is available at: https://osf.io/2ca8h/.

We introduced a color-response mapping in the WAM task (Figure 1). Three target positions were marked by unfilled circles (r = 1.5 cm). At target onset, the circles were filled for 1000 ms with one of three colors, spaced equidistantly (120°) in CIELab (L*a*b*: orange [54.0, 56.4, 47.2]; green [54.0, -31.4, 17.1]; purple [54.0, -43.1, -31.4]). Each color coded for a unique action: a button press with the index, middle, or ring finger. Targets in WAM were presented in a continuous sequence that was not divided into discrete trials, and multiple targets could be presented on screen at the same time (coincide by maximally 500 ms). The key manipulation was that, unbeknownst to the participants, one regular target consistently appeared every 3000 ms, while the other two irregular targets appeared at pseudorandom intervals. Each of the three targets was presented equally often (twenty times). The irregular sequence was designed such that the interval between two subsequent presentations was minimally 1250 ms (250 ms between offset and new onset), and that there was always at least one irregular target between two regular target presentations. Importantly, the sequences were designed such that there was no ordinal structure in the overall sequence of any subsequent target (see Appendix A for more details on the stimulus sequence).

To assess to what extent SL results from learning regular locations and/or actions, we designed three conditions that differed to what extent the regularity was linked to location or action. In the 'action=location' condition, the targets were placed at the three corners of an equilateral triangle (sides = 24 cm). Each target color appeared in a fixed location, requiring a fixed action. One of these was the regular target appearing every 3000 ms (Figure 1a, red arrows). In the 'action' and 'location' conditions, the direct mapping between location and action was removed (Figure 1b). To avoid Simon interference, the target positions were aligned vertically at the horizontal center of the screen. In the 'action' condition, the regular targets required a unique finger response (i.e., all regular targets had the same color), but appeared in a random location (Figure 1b, red arrows). Conversely, in the 'location' condition, the regular targets appeared in a fixed location, but

had a random color and were thus linked with random actions.

In all conditions, participants scored points by pressing the corresponding color-button on a conventional keyboard with their right index, middle and ring finger (colored stickers were placed on the left, right and down arrow keys). To indicate which button participants pressed, a rectangle in the associated color appeared around the three targets. Points were scored if the correct color-button was pressed within 1000 ms after target's onset. Each hit was rewarded with three points and accompanied by the sound of a dropping coin. If participants were too late (250 ms after target offset) or pressed an erroneous button, three points were deducted as a penalty.

2.4 Procedure

Each of the three conditions consisted of eight experimental blocks in which sixty targets were presented in total, twenty times at each of the three locations and in each of the three colors. A block lasted approximately one minute, and was followed by feedback on the total score, the number of hits and misses, and the high-score. Participants always began the experiment with the 'action=location' condition, preceded by one practice block with no regularity. The regular color and location in the 'action=location' condition was counterbalanced across participants.

After the 'action=location' condition, participants performed four practice blocks with vertically aligned targets where target locations and colors were no longer coupled, without any regularities. The order of the following 'action' and 'location' conditions was counterbalanced across participants. The features of the regular targets in these two following conditions changed with respect to the 'action=location' condition (e.g., if the regular targets were green in 'action=location' condition, they changed to orange or purple in the 'action' condition). We added the constraint that the middle finger response was never tied to the regularity in the 'action' condition, and that the regularity was never in the middle position in the 'location' condition. In this way, we aimed to prevent positive biases (i.e., faster and more accurate responses) that we found for the middle finger

response and middle target (Supplementary information). At the end of the experiment, participants answered seven questions to assess their awareness of the temporal regularity (reported in Appendix B).

2.5 Inclusion Criteria

We excluded and replaced the data from three participants that indicated to be aware of one of the regularities in the questionnaire. We additionally excluded data from blocks in which the hit rate for one of the three targets was less than 25%, suggesting that a participant might not have divided their attention across all targets. Two more participants were excluded because more than three out of eight blocks in one condition had to be discarded.

Response outliers were further identified in two steps: First, a global cutoff (across-participant) was set to remove extreme outliers (only including 150 ms $<$ RT $<$ 1500 ms). Second, we removed RTs that exceed the median -/+ three times the median absolute deviation (computed after the global cutoff) for each participant (within-participant cutoff). All RTs not identified as outliers, were defined as a hit.

2.6 Data Analyses

Analyses were conducted in Python 3.8.5 and R 3.6.1 and assessed participants' responses expressed by RT and HR. We analyzed log-transformed RT using linear mixed models (LMMs, Baayen et al., 2008; Bates et al., 2015) and HR by general (G)LMMs with a logistic link function.

For both RT and HR, we included 'participant' as a random intercept. The main interest of the study was whether regularity of targets (regular versus irregular) and their possible interaction with condition ('action=location', 'action', versus 'location') affected performance. As introduced later in section 3.2, we extended the statistical model with the predictor foreperiod (FP). These predictors are reported in the Results section. To investigate order effects, we additionally assessed whether these predictors interacted with the order of condition presentation ('action $\rightarrow$ location' versus 'location $\rightarrow$ action'; note

that condition 'action=location' was always presented first). Furthermore, we considered the effect of previous response (hit versus miss) and practice effects (time-on-task, indexed by block number). These additional analyses did not reveal any effects on RT or HR (Supplementary information).

We tested the contribution of each of the predictors by means of model comparisons, based on Bayes factors (BF) estimated from BIC scores (Wagenmakers, 2007). To quantify statistical evidence for or against a predictor term, we compare the best statistical model that includes this predictor against the best model without it. For readability, we relegated the details of statistical model selection to the Supplementary information.

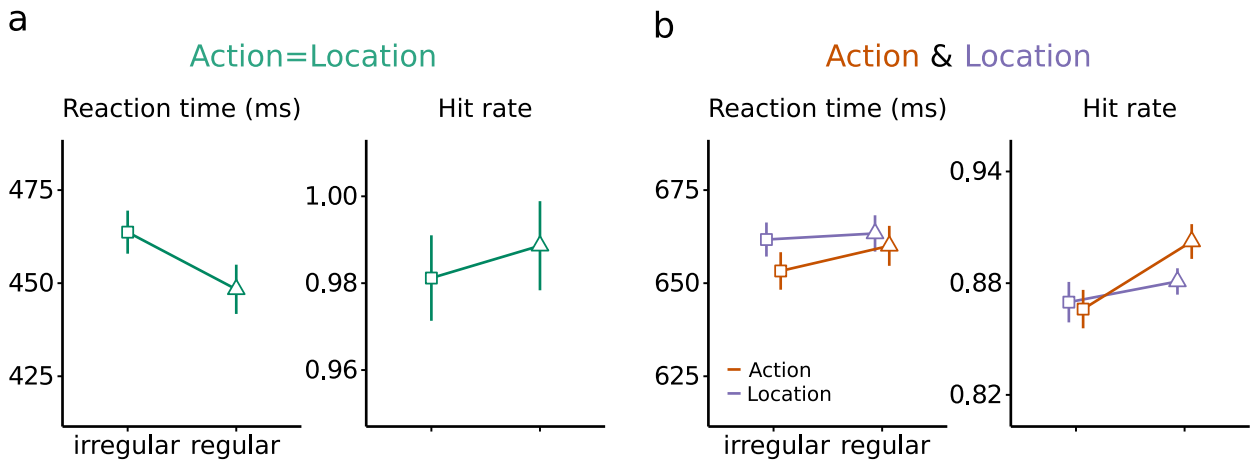
3 Results

3.1 Regularity Benefit

A questionnaire indicated that only three participants were aware of the regularity. Those were replaced by new, unaware participants. The findings of the 'action=location' condition (Figure 2a) replicate our previous results of the regularity benefit (Salet et al., 2021): RT was lower and HR higher for the regular compared to irregular targets ($\Delta BIC >$ 15.7, $BF > 1000$). In our earlier work, we attributed this benefit to a SL mechanism distilling the regular interval from the environment (Salet et al., 2021). For the 'action' and 'location' conditions (Figure 2b), the HR results align with these findings ($\Delta BIC > 28.0$, $BF > 1000$). However, we found no indication of faster responses to regular targets in the 'location' condition ($\Delta BIC = 8.9$, $BF = 83.6$) and inconclusive evidence indicating slower responses to the regular target in the 'action' condition ($\Delta BIC = 1.1$, $BF = 1.7$). This inconsistency between HR and RT is surprising: As faster responses lead to higher HRs, we had expected RT to be the strongest driver of regularity benefits (cf. Salet et al., 2021). In the following, we will show that a different explanation of the regularity benefit, rooted in temporal preparation, resolves this apparent inconsistency.

Figure 2

Performance as a function of regularity.



Note. Mean RT and HR split on regularity with 95% within-subject confidence intervals (Cousineau, 2005; Morey, 2008) for the (a) 'action=location' condition and (b) 'action' and 'location' conditions. Note that the y-scale of 'action=location' condition differs from the 'action' and 'location' conditions.

3.2 Temporal Preparation

[1]As the interval between irregular targets was highly variable (1250 – 14000 ms), we have thus far considered it unlikely that these intervals are monitored by participants and used to time responses. However, it has been suggested that participants still monitor these intervals to prepare for upcoming irregular onsets (see Salet, Schlichting, et al., 2022).

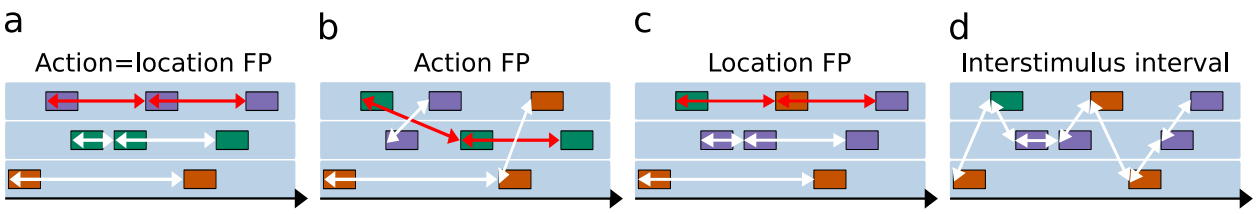
3.2.1 Preparing Irregular Targets

In this section, we test whether participants prepare for irregular targets, expressed by a decrease in RT and increase in HR as a function of the interval between each subsequent irregular onset (temporal preparation's signature: Los et al., 2014; Niemi & Näätänen, 1981; Nobre & Ede, 2018). From here on, we refer to this interval as the foreperiod (FP). Figure 3a illustrates the FP in the 'action=location' condition in which action was in correspondence with location. Note that the FP for the regular target was

[1] All following analyses were not preregistered.

Figure 3

A multitude of temporal properties in WAM.



Note. Illustrating the foreperiod (FP) in the (a) 'action=location', (b) 'action', and (c) 'location' condition. The red arrows represent the timing of the regular target (constant FP of 3000 ms), the white arrows the irregular targets (FPs ranging from 1250 - 14000 ms). (d) The time between any subsequent target presentation is referred to as the interstimulus interval (ranging from 500 - 2500 ms).

always 3000 ms by definition (Figure 3a, red arrows). In the 'action' and 'location' conditions, we define the FP as the time between two targets with the same action (e.g., ring-to-ring finger) or location (e.g., top-to-top location), respectively (Figure 3b and c).

Figure 3d illustrates another temporal property; the interstimulus interval (ISI). This is the time between any subsequent target presentation, and is defined on a much smaller range (500 – 2500 ms) than FPs (1250 - 14000 ms). As can be expected, we found RT to decrease and HR to increase as a function of ISI (Supplementary information). We found, however, that ISI does not modulate the effect of FP on regularity. To focus our report, we relegated these ISI analyses to the Supplementary information and continue with the critical predictors FP and regularity.

To test whether FP duration modulated performance, we extended the statistical model of the section 'Regularity Benefit' (predicting RT and HR from 'regularity' and 'condition') to also include an effect of FP (as a continuous predictor) and its interaction with 'condition'. Of note, we transformed FP into $\frac{1}{FP}$ to better capture its curvilinear relation with RT and HR (see Figure 4). This vastly improved the models of both HR and RT compared to a linear relation ($\Delta BIC > 1000$, $BF > 1000$). Figure 4 displays RT and

HR as a function of binned FP together with model predictions[2]. Model comparisons revealed decreasing RT and HR as a function of FP, suggestive of temporal preparation ($\Delta BIC > 482.7$, $BF > 1000$).

3.2.2 Is the Regularity Benefit an Illusion?

The inclusion of FP as predictor meant that the inclusion of the predictor 'regularity' was no longer warranted as fixed effect for both HR and RT ($\Delta BIC > 7.3$, $BF > 38.3$). Additionally, the model only including FP vastly outperformed the model including only regularity ($\Delta BIC > 488.8$, $BF > 1000$). In other words, what seemed to reflect SL can be better accounted for by temporal preparation. This is illustrated in Figure 4, which shows the average RT and HR as a function of regularity and FP. For irregular targets, short FPs (< 3000 ms) are characterized by high RTs and low HRs. Due to the strong asymptotic nature of preparation effects, long FPs (> 3000 ms) are not affected to the same extent. This asymmetry leads to an apparent regularity benefit when only considering the average RT and HR as a function of regularity (as in Figure 2). However, when observing performance as a function of FP, it can be seen that at FP = 3000 ms (Figure 4, insets), there is no apparent difference between regular and irregular targets.

3.2.3 Discussion

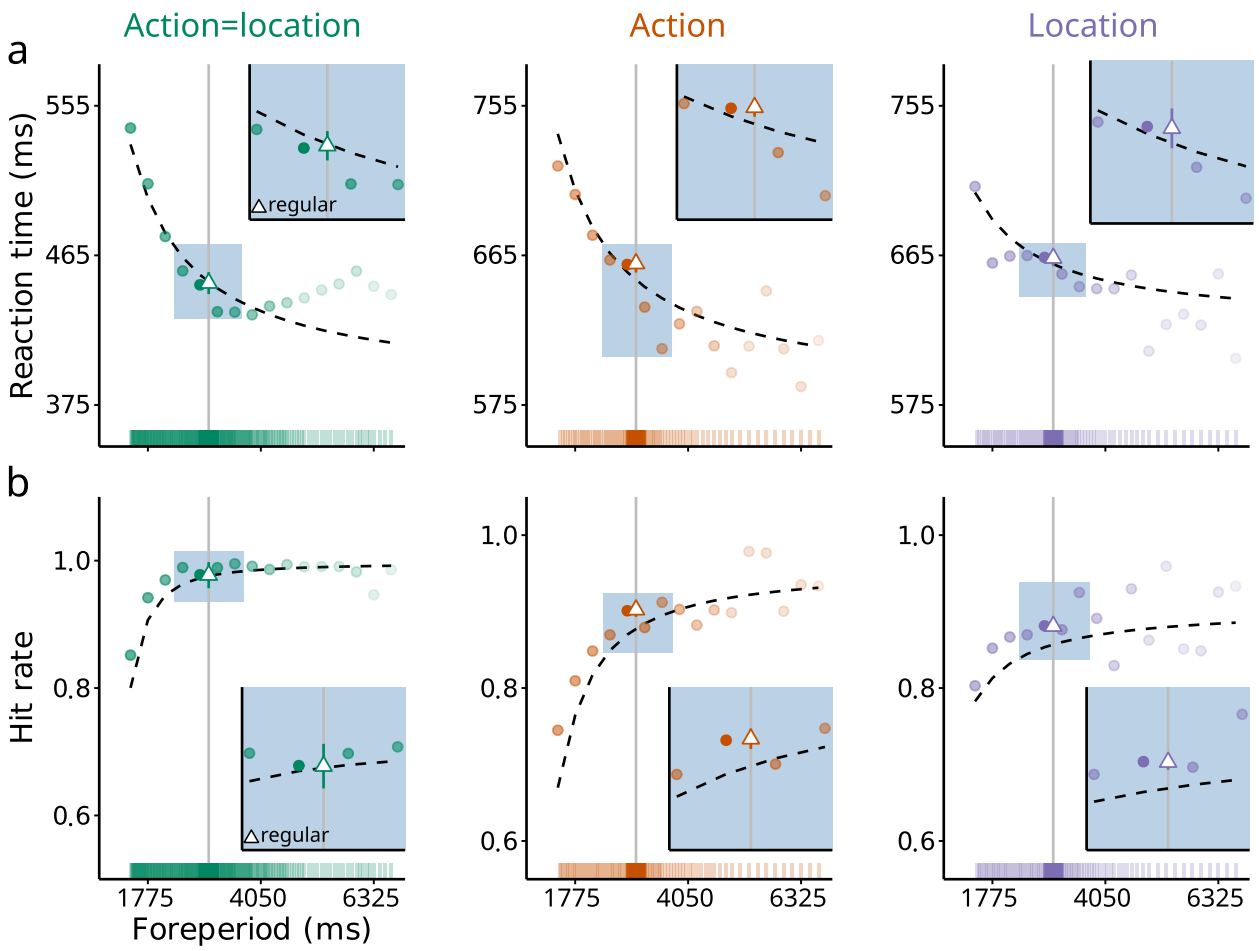
We found that participants temporally prepare for all upcoming targets, irrespective of their assigned regularity. Importantly, these findings go against our previous interpretation that participants only capitalize on the regularity embedded in WAM, giving rise to the regularity benefit (see also Addendum to Salet et al., 2021: Salet, Schlichting, et al., 2022). Instead, participants show to monitor both regular and irregular intervals to temporally prepare what action to perform and which location to attend.

---

[2] Note that statistical model predictions represent estimated marginal means (i.e., at the average level of other predictors) and not a least-squares fit to empirical data.

Figure 4

Temporal preparation.



Note. (a) Mean RT decreases and (b) HR increases as a function of foreperiod (FP), characteristic of temporal preparation effects for both regular and irregular targets. The points represent binned RT and HR (bin size = 350 ms), dashed lines represent statistical model fits. For illustration purposes, we omitted 1% of the tail of the data (FP > 7200 ms). Opacity of the colored bar above the x-axis and the data points illustrate the amount of data per FP bin. Most data is around an FP of 3000 ms (solid vertical line). The open triangle is the same average for regular targets as plotted in Figure 2. As displayed, there is no apparent difference between regular (triangle) and irregular targets (circles) around FP = 3000 ms. Note that for RT the y-axis scale is the same for the 'action' and 'location' conditions, but differs from the 'action=location' condition.

3.3 What Action and Which Location Comes Next

Thus far, we have defined the FP in the 'action' condition as the time between repeating actions, and in the 'location' condition as the time between repeating locations (Figure 3b and 3c). This allowed us to evaluate whether FP effects could explain the regularity benefit observed in these conditions. However, each target in the 'action' and 'location' conditions has both an 'action' and a 'location' FP (FP_A and FP_L; Figure 5a). It might be that participants exploit both FP types and prepare both what action (e.g., ring finger) and which location comes next (e.g., top location). If so, we should find preparation curves for both FP types in each condition. Additionally, as regularity does not affect performance and the conditions only differed in the assigned regularity (Figure 1b, action versus location), we expect the effect of FP type to be similar across the two conditions.

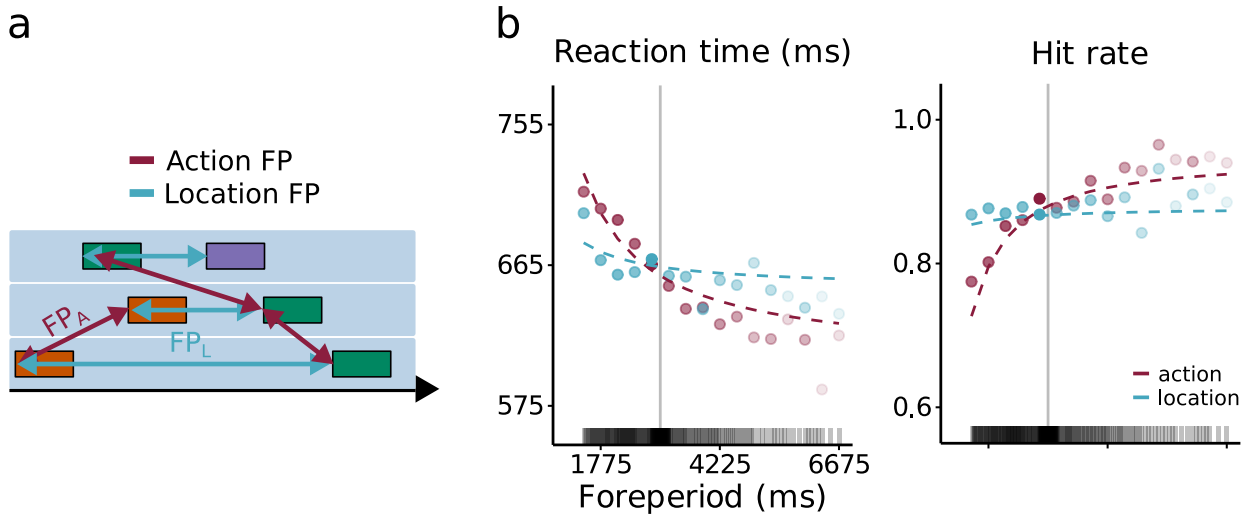
To test this, we constructed new statistical models including the predictors FP_A and FP_L, and tested whether they interacted with 'condition'. Figure 5b displays preparation effects on RT, split on FP type, together with the statistical model's predictions. Model comparisons revealed an effect of FP_A for both RT and HR ($\Delta BIC > 270.3$, $BF > 1000$). For FP_L, an effect was found on RT ($\Delta BIC = 26.0$, $BF > 1000$), but evidence of absence for an effect on HR ($\Delta BIC = 6.8$, $BF = 30.3$). The models' steeper RT curves as a function of FP_A versus FP_L (slope coefficient 4.1 times larger for FP_A) and the absence of an effect of FP_L for HR suggests that preparation effects were more pronounced for action than for location. Finally, FP effects did not differ between the 'action' and 'location' conditions ($\Delta BIC > 13.9$, $BF > 1000$). Indeed, this confirms that the conditions govern similar preparation dynamics and are unaffected by the modality of the regularity (action versus location).

3.3.1 Discussion

Here we show that in the 'action' and 'location' conditions, participants temporally prepare what action to perform and which location to attend in parallel. Furthermore, we found indications that the effect of 'action' preparation (FP_A) is more pronounced than

Figure 5

Where and what is next.



Note. (a) Illustrates the two FP types: action FP and location FP in the same stimulus segment. (b) Mean RT and HR as a function of FP (plotted as in Figure 4).

'location' preparation (FP_L). In the Supplementary information, we present results of a very similar experiment with 54 participants that replicate all findings presented so far.

3.4 *f*MTP: A Computational Account of Preparation

When considering all temporal information embedded in WAM (Figure 3), we find behavioral effects that reflect temporal preparation to both regular as irregular targets (Figure 4 and 5) and not a SL mechanism that only distills the regular interval. In this section, we use a model of temporal preparation (*f*MTP, Salet, Kruijne, van Rijn, et al., 2022) as a proof of concept to illustrate the cognitive mechanisms from which our empirical findings might arise. To do so, we focus on the 'action=location' condition and, in particular, on two key aspects of the data. First, does using WAM's stimulus sequence as input to the model lead to the same preparation curves as observed in the data? Second, does the average RT computed from these preparation effects give rise to an apparent regularity benefit?

3.4.1 Model Architecture

A detailed (mathematical) description of *f*MTP can be found in Salet, Kruijne, van Rijn, et al. (2022). In short, *f*MTP is a neural network consisting of a motor circuit and a timing circuit (Figure 6a). The motor circuit consists of an activation and inhibition node, and their interplay resembles the initiation of a motor plan (i.e., a button press corresponding to the target on screen). During the FP, activation is suppressed by inhibition (Davranche et al., 2007; Duque & Ivry, 2009; Greenhouse et al., 2015; Hasbroucq et al., 1999; Prut & Fetz, 1999; Sinclair & Hammond, 2008), enabling the staging of a motor plan without executing it, akin to an athlete awaiting the start signal before taking off. At target presentation, however, inhibition is released, and the activation node is triggered, initiating the response.

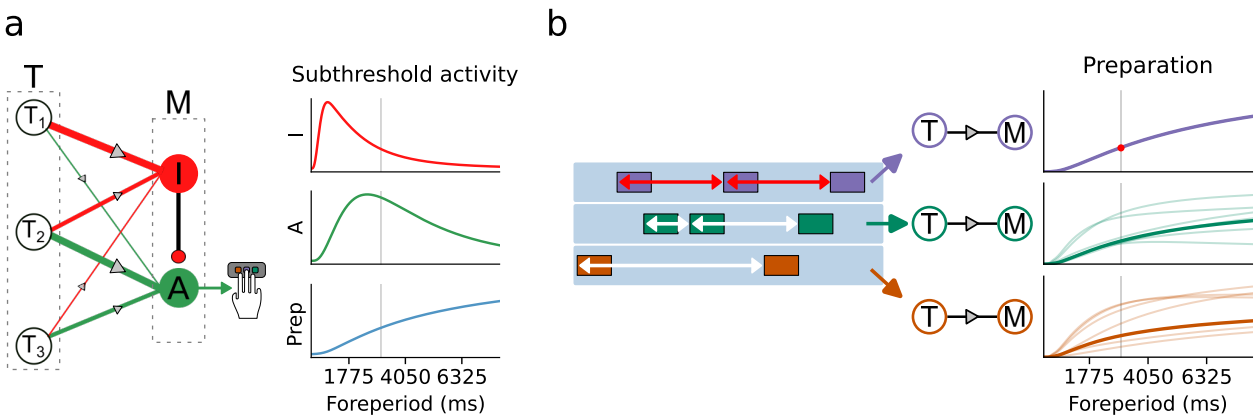
3.4.2 Trace Formation

This interplay between the inhibition and activation is driven by a timing circuit (Figure 6a, Shankar & Howard, 2012): An array of neuronal units that convey information about the passage of time during the FP (Eichenbaum, 2014; MacDonald et al., 2011; Pastalkova et al., 2008). Different time points during the FP are encoded by different nodes in the timing circuit. Each node in the timing circuit projects to the motor circuit, and their simultaneous firing leads to Hebbian learning (see Gallistel & Gibbon, 2000; Machado, 1997; Mauk & Ruiz, 1992, for similar rules used for learning to time). In this way, specific points in time get associated with either inhibition or activation.

The Hebbian associations between the timing and motor circuit are stored in 'memory traces' on each FP trial. A simplified example of this process is illustrated in the left panel of Figure 6a, depicting three example nodes in the timing circuit that encode time points before (T_1), around (T_2), and after (T_3) target onsets. Through Hebbian learning caused by simultaneous firing of the timing and motor circuit, T_1 primarily forms associations with inhibition (thick red line, it is too soon to respond), T_2 with activation (thick green line, it is time to respond), and T_3 neither with inhibition nor activation (thin

Figure 6

fMTP.



Note. Illustration of *f*MTP's underlying dynamics. (a) Trace formation. In the motor circuit (M), inhibition (I) controls activation (A) to prevent premature responses during the FP. Each node in the timing circuit (T; three example nodes are drawn) projects to both I and A. The thickness of the connections represents the association strength after a sequence of FPs. These associations determine the subthreshold fluctuations in I and A determining preparation (Prep) on the next trial. (b) In WAM, each target (unique location and action) triggers a unique preparation circuit. A constant FP (purple) leads to similar preparation on each trial. As the regular FP is always 3000 ms (solid vertical line), the model's predicted preparation is only observed at that specific FP (red dot). For the irregular targets (green and orange, varying FPs) preparation fluctuates. Transparent lines represent preparation of different target presentations, the non-transparent line represents average preparation across all presentations. Note that I and A are plotted in arbitrary units, but their ratio (preparation) directly scales to RT.

connection lines, target has already passed). These Hebbian associations define the 'memory trace' at each trial.

3.4.3 Memory

Critically, *f*MTP postulates that preparation is determined by the collective contribution of traces formed on previous trials, retrieved from memory. Due to forgetting, traces contribute according to their recency with traces of remote trials contributing less

strongly than of recent trials (Wixted, 2004; Wixted & Ebbesen, 1991). It is the contribution of memory traces by which the network's state tunes to the environment's temporal properties. The right panel of Figure 6a visualizes the result of this 'tuning' after a sequence of multiple FPs (range of 300 - 5500 ms). Early points in time become strongly associated with inhibition, whereas later points in time become more strongly associated with activation. These subthreshold fluctuations in the balance between activation and inhibition during the FP reflect preparation (Figure 6a, blue line). For later time points (FP $>$ 3000 ms), as activation starts to decrease later than inhibition, preparation gradually increases. This gives rise to the asymptotic nature of preparation effects.

3.5 *f*MTP and WAM

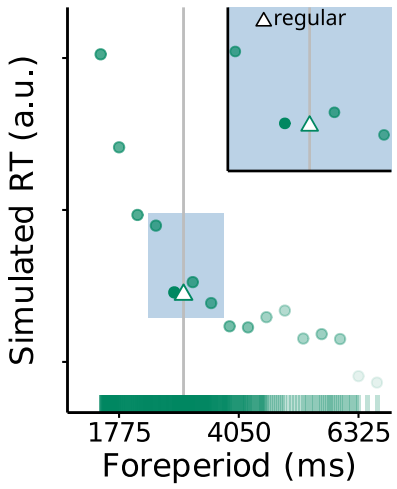
Originally, *f*MTP was developed and tested as a model of experiments in which FPs (range of 300 - 5500 ms) are presented in isolated trials (Salet, Kruijne, van Rijn, et al., 2022). To model the WAM task with three continuous target streams and a wide range of FPs (1250 – 14000 ms), we extend the model such that each target triggers a unique preparation circuit existing of timing and motor circuits (Figure 6b). With this, we assume specificity of preparation. Each target elicits its associated timing circuit that projects to a corresponding motor circuit, representing a specific response. Aside from this assumption (its rationale is discussed in detail in Appendix C), we keep all other aspects of *f*MTP the same and directly adopt the default parameter values reported in Salet, Kruijne, van Rijn, et al. (2022).

3.5.1 Simulation

To arrive at 'simulated' RT, we formalize the interplay of inhibition and activation in *f*MTP (i.e., preparation) as the I/A ratio. High levels of inhibition increase the I/A ratio and slow down responses, and vice versa for high levels of activation. Salet, Kruijne, van Rijn, et al. (2022) showed that this 'simulated RT' linearly relates to RT. Here, we focus the simulation on the goodness of the qualitative predictions of *f*MTP (i.e., 'simulated RT'). Of note, we do not simulate HR but hold the rationale that faster RT

Figure 7

fMTP: simulated RT.



Note. Mean simulated RT (arbitrary units, a.u.) as a function of FP (plotted as Figure 4 and 5).

leads to higher HR, a relation found across all empirical data (Figure 4 and 5).

Figure 7 shows that the model's predictions resemble the data. The RT curve for the irregular targets arises from *f*MTP's memory principles as earlier time points become, through Hebbian learning, primarily associated with inhibition, whereas for later time points activation is more pronounced. This changing balance, shifting in favor of activation as time passes, yields increasing preparation as a function of FP (see Figure 6), resulting in the downward sloping RT curve. Additionally, when observing simulated RT at FP = 3000 ms, there is no apparent difference between regular (always presented at a FP of 3000 ms) and irregular targets. Due to the asymmetric nature of preparation effects, averaging across regularity leads to an apparent regularity benefit, just as observed in the empirical data.

3.6 Discussion

Our simulations show that *f*MTP captures the two key aspects of the data. Guided by *f*MTP, we can make assumptions about the cognitive mechanisms from which our findings might arise. That is, responses in WAM are governed by associative memory processes between timing circuits that monitor the timing of the targets, and motor circuits

initiating specific responses. This target-by-target learning process drives preparation during the interval between targets, governing the readiness to respond. In sum, these simulations serve as a proof of concept to show that our results can arise from the memory processes of temporal preparation, without assuming that the processes underlying responses to regular targets are any different from responses to irregular targets.

4 General Discussion

SL research is signified by a behavioral benefit for regular versus irregular properties of the environment. This regularity benefit is commonly interpreted as arising from the ability to extract regularities from the environment (e.g., Baker et al., 2004; Boettcher et al., 2022; Fiser & Aslin, 2001; Salet et al., 2021; Sherman et al., 2020; Zhao et al., 2013). In line with this view, we theorized that the regularity benefit observed in WAM's complex temporal environment is a reflection of a mechanism extracting the temporal information of the regular target (Salet et al., 2021). Intriguingly, however, we found participants to be capable of extracting a range of temporal properties that we initially had considered to be irregular and uninformative. Participants monitored, in parallel, different intervals to temporally prepare what action to prepare in which location, irrespective of target's regularity. This temporal preparation even showed to fully account for the regularity benefit, leaving the experimental manipulation of regularity in WAM obsolete.

4.1 Mechanistic Explanations of Statistical Learning

SL research has convincingly shown the regularity benefit to be a robust finding across different domains (e.g., Fiser & Aslin, 2001; Olson & Chun, 2001; Saffran et al., 1996; Turk-Browne et al., 2009; Yu & Zhao, 2015; Zhao et al., 2013). The theoretical construct of SL, however, has been argued to be underspecified: mechanistic explanations of the cognitive processes that drive SL's regularity benefit lag behind the robust empirical phenomena (Frost et al., 2019; Thiessen, 2017).

4.1.1 Experimental Challenges of Finding Mechanistic Explanations

An experimental issue for finding mechanistic explanations of SL is that it is difficult to nail down what experimental factor drives the regularity benefit. Embedding a (statistical) regularity in the experiment can introduce other experimental factors that may contribute to the regularity benefit. An example of this issue has been highlighted for 'probability cueing' (Jiang, 2018; Shaw & Shaw, 1977), a visual search paradigm where regularity benefits can be observed for targets that appear in regular compared to irregular locations. It has been argued, however, that these benefits are for a large part driven by repetition priming, i.e., a phenomenon that refers to an improvement in performance when stimuli are repeatedly presented (Maljkovic & Nakayama, 1994; Rabbitt et al., 1979). By having targets appearing more regularly at certain locations (i.e., the manipulated statistical regularity), they almost invariably appear more recently at that location as well (i.e., repetition priming). It has been shown that repetition priming can fully account for regularity benefits, without assuming SL of the regular locations (Kabata & Matsumoto, 2012; Maljkovic & Martini, 2005; Walthew & Gilchrist, 2006).

Similarly, we show that temporal regularity benefits can be fully accounted for by temporal preparation, without assuming SL of regular intervals. Such temporal preparation might similarly play a role in recent work assessing the learning of regular interval patterns in a complex visual search task (Boettcher et al., 2022; Shalev et al., 2022). In this search task, regularity benefits were found and interpreted to arise from SL of the regular spatiotemporal pattern. However, the experiment's complex stimulus sequence embeds a multitude of other temporal properties just as in WAM (similar to Figure 3). Therefore, as in this study, the regularity benefit might be (partly) driven by temporal preparation to these temporal properties.

Temporal preparation and SL are, however, not necessarily mutually exclusive factors that drive regularity benefits. Just as later probability cueing studies showed that both repetition priming as SL contribute to regularity benefits (Goschy et al., 2014; Jones

& Kaschak, 2012), it is possible that this is also the case in complex temporal environments like in our WAM task and Boettcher et al. (2022)'s search task. For example, although we found no evidence for SL, it might be that in different settings of WAM (e.g., a speeded WAM version with lower accuracy scores), the higher predictability of the regular compared to the irregular targets is used to improve performance further.

Likewise, our findings do not imply that regularity benefits found in studies using simpler stimulus sequences, like temporal SRT tasks (O'Reilly et al., 2008; Shin & Ivry, 2002), can be fully explained by temporal preparation without assuming SL. They strongly urge, however, to consider the role temporal factors other than the manipulated statistical regularity play in learning behavior. Indeed, this has already been acknowledged in some temporal SRT studies (Heideman et al., 2016, 2018; Salidis, 2001). In these studies, a larger regularity benefit for shorter compared to longer intervals was interpreted to reflect an interaction between temporal preparation and SL: As temporal preparation is not yet fully developed at short intervals, SL has a stronger impact at shorter than longer intervals. All together, future experimental research could further address to what extent and under what circumstances temporal regularity benefits emerge from temporal preparation, SL, or both.

4.1.2 Statistical Learning as General Memory Processes

Despite these difficulties of dissociating SL of the regularity from learning other experimental factors, one line of theoretical work has put forward a unifying explanation of the learning principles underlying SL's regularity benefits. Various computational models developed under this view showed that regularity benefits can arise from general memory processes underlying both regular and irregular input (French et al., 2011; Kruijne & Meeter, 2015; Mareschal & French, 2017; McClelland & Elman, 1986; Perruchet & Pacton, 2006; Perruchet & Vinter, 1998; Schapiro et al., 2017; Thiessen & Pavlik Jr., 2013). Although these theories have their differences, conceptually they work from two main principles of memory: activation and a forgetting process modeled as decay or interference. In these memory models, each presented item (irrespective of its assigned regularity) is

automatically stored in a memory trace that decays over time. Trace activation increases as an item is repeated, while other items may cause interference and decrease trace activation. In this way, repeating items in the environment (i.e., regularities) gain dominance in memory through associative learning between incoming input and memory traces. The application of these models has provided support that such memory processes can lead to adaptation to the environment's statistical regularities. Our findings, demonstrating that regularity benefits can emerge from memory processes underlying temporal preparation, instantiate this perspective in the temporal domain. Similar to memory models of SL, we showed that *f*MTP's associative memory processes indirectly give rise to regularity benefits in WAM.

5 Conclusions

In conclusion, we found that temporal preparation instead of SL drives regularity benefits for regular intervals in complex temporal environments. Instead of only extracting a single regular interval through SL, we found participants to extract both regular and irregular intervals to temporally prepare for what action to perform and which location to attend. Guided by *f*MTP, we illustrated that such adaptive behavior can emerge from associative memory processes underlying temporal preparation. In a broad sense, we consider our findings as a warning of interpreting regularity benefits to only arise from learning the manipulated statistical regularity in the experiment. Especially in more complex, dynamic experimental tasks striving to reflect some demands of naturalistic settings (such as WAM), it is not unlikely that participants pick up on other experimental factors that contribute to or drive the observed regularity benefit.

6 Appendices

A Stimulus Sequences

For each participant and experimental block, a unique sequence of intervals for the irregular targets was generated. Note that the exact same sequences were used between conditions: sequences only differed in their assigned regularity between conditions, their

timing properties remained identical.

The generation of the sequences followed a stringent pseudo randomization procedure. The randomization of the timing of the irregular sequences was minimally constrained such that the interval between two subsequent targets in the same location was minimally 1250 ms (250 ms between offset and new onset), there was always at least one irregular target between two regular targets, and that target presentations at different locations could coincide by maximally 500 ms.

Additionally, the interval between the onset of a target and the onset of any next target (ISI, Figure 3d) was controlled for, such that the distributions of ISI's preceding regular and irregular targets was equated. Systematically longer ISI's in either condition could drive the regularity benefit effect in the same direction as SL. To minimize ISI differences between regular and irregular targets we assessed the ISI distributions for regular and irregular targets by means of the Kullback–Leibler (KL) divergence (a procedure developed in: Salet et al., 2021). For each participant's stimulus sequence we computed the KL divergence as a measure of divergence between the probability distribution of regular versus irregular sequences. We only used sequences with a KL-divergence lower than 0.001 (see section 'Interstimulus Interval' at https://osf.io/3d6h8 for more details).

B Questionnaire

At the end of the experiment, we assessed participant's awareness of the regular interval pattern using the following seven questions: (1) They were encouraged to report anything they might have found 'noteworthy' in the experiment; (2) After being informed about the temporal regularity in the 'action=location' condition, they were asked whether they had noticed this regularity (Yes/No); (3) They were presented with a forced choice question to indicate (or guess) which target location/color had been the regular one; (4) After having been informed of the existence of the temporal regularities in the 'action' and 'location' condition, they were presented with questions to indicate whether they had

noticed these (Yes/No); (5) They were asked a question to indicate in what order the regularities had been presented; Lastly, they were presented with two more forced choice questions to indicate (or guess) the color (6a) and location (6b) of the temporal regularity.

In the open question, none of the participants reported to be aware of the regularity. In the forced-choice yes/no questions, 5.5% ('action=location' condition) and 13.0% ('action' and 'location' conditions) indicated to have been aware of the regularity. However, only 30.0% ('action=location' condition) and 34.0% ('action' and 'location' conditions) of participants were able to correctly identify the regular color and/or location, closely in line with a guessing rate of 33.3%. Three participants were replaced because they both indicated to have been aware of the regularity, and had correctly identified the regular color and/or location.

C Specific versus nonspecific preparation

In our modelling (section 3.4), we characterized preparation in WAM as a specific process, meaning that each target presentation triggers its associated timing circuit that projects to a corresponding motor circuit representing a specific response (Figure 6b, main text). This is in contrast with earlier work in which fMTP characterized preparation as a general, 'non-specific' process independent of different response options (Salet, Kruijne, van Rijn, et al., 2022). A version of the model that incorporates this assumption, however, fails to capture preparation effects observed in WAM (Supplementary information). This mismatch motivated us to adopt the 'specific' preparation model presented in the main text. As shown in Figure 7 (main text), this model's predictions qualitatively align with the data. Although direct empirical evidence for this assumption is as of yet lacking, our results fit within a broad range of behavioral and neurophysiological studies which address the (non)specificity of motor preparation.

Traditionally, the finding that inhibition also applies to task-irrelevant effectors has led to the view that preparation comes with widespread non-specific inhibition of the entire motor system (Hasbroucq et al., 1999; Prut & Fetz, 1999). This is how we had

implemented *f*MTP (Salet, Kruijne, van Rijn, et al., 2022), and aligns with the observation that preparation effects are qualitatively highly similar in simple- versus choice RT tasks (Bertelson & Boons, 1960; Frith & Done, 1986; Los & van den Heuvel, 2001; Steinborn & Langner, 2012).

In contrast with this 'non-specific' view, however, studies have also found higher inhibition of the selected compared to the non-selected effector of the response, suggesting at least some specificity of inhibition (Davranche et al., 2007; Duque & Ivry, 2009; Greenhouse et al., 2015). This finding has led to the 'gain modulation hypothesis' that proposes that inhibition enhances the specificity of selected responses (Duque et al., 2017; Greenhouse et al., 2015). The specificity of inhibition can be illustrated by imagining an 'inhibition spotlight' centered around response representations (Greenhouse et al., 2015). A narrow focus, increasing specificity, solely inhibits the selected response, while a broad focus also inhibits non-selected responses. Interestingly, the focus – and thereby specificity – is hypothesized to be context dependent. In tasks with a single response option (e.g., simple RT task), the focus is more broad and therefore also inhibits non-selected responses. However, with multiple response options (e.g., choice RT task), its focus narrows to enhance the separation between the response representations; thereby decreasing the probability of selecting the incorrect response.

In WAM, participants had three response options, and continuously risked a penalty for incorrect choices. From the view of the gain modulation account, it might be that these task demands narrowed the spotlight of inhibition, thereby characterizing preparation as a specific process. A range of behavioral studies have similarly shown that preparation can, at least in some contexts, be specific for the selected response (Langner et al., 2018), target identity (Thomaschke et al., 2011, 2016; Wagener & Hoffmann, 2010), and target location (Wagener & Hoffmann, 2010). These findings align with the specificity assumption made in the main text (Figure 6b). However, to what extent and under which circumstances preparation reflects a (non)specific process remains a question that requires more research.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, *15*(7), 460–466. https://doi.org/10.1111/j.0956-7976.2004.00702.x

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bertelson, P., & Boons, J.-P. (1960). Time uncertainty and choice reaction time. *Nature*, *187*(4736), 531–532. https://doi.org/10.1038/187531a0

Boettcher, S. E. P., Shalev, N., Wolfe, J. M., & Nobre, A. C. (2022). Right place, right time: Spatiotemporal predictions guide attention in dynamic visual search. *Journal of Experimental Psychology: General*, *151*(2), 348–362. https://doi.org/10.1037/xge0000901

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

Davranche, K., Tandonnet, C., Burle, B., Meynier, C., Vidal, F., & Hasbroucq, T. (2007). The dual nature of time preparation: Neural activation and suppression revealed by transcranial magnetic stimulation of the motor cortex. *European Journal of Neuroscience*, *25*(12), 3766–3774. https://doi.org/10.1111/j.1460-9568.2007.05588.x

Duque, J., Greenhouse, I., Labruna, L., & Ivry, R. B. (2017). Physiological markers of motor inhibition during human behavior. *Trends in neurosciences*, *40*(4), 219–236. https://doi.org/10.1016/j.tins.2017.02.006

Duque, J., & Ivry, R. B. (2009). Role of corticospinal suppression during motor preparation. *Cerebral Cortex*, *19*(9), 2013–2024. https://doi.org/10.1093/cercor/bhn230

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. *Nature Reviews Neuroscience*, *15*(11), 732. https://doi.org/10.1038/nrn3827

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*(6), 499–504. https://doi.org/10.1111/1467-9280.00392

French, R. M., Addyman, C., & Mareschal, D. (2011). TRACX: A recognition-based connectionist framework for sequence segmentation and chunk extraction. *Psychological Review*, *118*(4), 614–636. https://doi.org/10.1037/a0025255

Frith, C. D., & Done, D. J. (1986). Routes to action in reaction time tasks. *Psychological Research*, *48*(3), 169–177. https://doi.org/10.1007/BF00309165

Frost, R., Armstrong, B. C., & Christiansen, M. H. (2019). Statistical learning research: A critical review and possible new directions. *Psychological Bulletin*, *145*(12), 1128–1153. https://doi.org/10.1037/bul0000210

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review*, *107*(2), 289–344. https://doi.org/10.1037/0033-295X.107.2.289

Gobel, E. W., Sanchez, D. J., & Reber, P. J. (2011). Integration of temporal and ordinal information during serial interception sequence learning [Publisher: American Psychological Association]. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 994–1000. https://doi.org/10.1037/a0022959

Goschy, H., Bakos, S., Müller, H. J., & Zehetleitner, M. (2014). Probability cueing of distractor locations: Both intertrial facilitation and statistical learning mediate interference reduction. *Frontiers in Psychology*, *5*, 1195. https://doi.org/10.3389/fpsyg.2014.01195

Green, P., & MacLeod, C. J. (2016). SIMR: An r package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Greenhouse, I., Sias, A., Labruna, L., & Ivry, R. B. (2015). Nonspecific inhibition of the motor system during response preparation. *Journal of Neuroscience*, *35*(30), 10675–10684. https://doi.org/10.1523/JNEUROSCI.1436-15.2015

Hasbroucq, T., Osman, A., Possamaï, C.-A., Burle, B., Carron, S., Dépy, D., Latour, S., & Mouret, I. (1999). Cortico-spinal inhibition reflects time but not event preparation: Neural mechanisms of preparation dissociated by transcranial magnetic stimulation. *Acta Psychologica*, *101*(2), 243–266. https://doi.org/10.1016/S0001-6918(99)00007-4

Hay, J. F., & Saffran, J. R. (2012). Rhythmic grouping biases constrain infant statistical learning. *Infancy*, *17*(6), 610–641. https://doi.org/10.1111/j.1532-7078.2011.00110.x

Heideman, S. G., Ede, F. v., & Nobre, A. C. (2018). Temporal alignment of anticipatory motor cortical beta lateralisation in hidden visual-motor sequences [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ejn.13700]. *European Journal of Neuroscience*, *48*(8), 2684–2695. https://doi.org/https://doi.org/10.1111/ejn.13700

Heideman, S. G., van Ede, F., & Nobre, A. C. (2016). Early behavioural facilitation by temporal expectations in complex visual-motor sequences. *Journal of Physiology-Paris*, *110*(4), 487–496. https://doi.org/10.1016/j.jphysparis.2017.03.003

Jiang, Y. V. (2018). Habitual versus goal-driven attention. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior*, *102*, 107–120. https://doi.org/10.1016/j.cortex.2017.06.018

Jones, J. L., & Kaschak, M. P. (2012). Global statistical learning in a visual search task [Publisher: American Psychological Association]. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(1), 152–160. https://doi.org/10.1037/a0026233

Kabata, T., & Matsumoto, E. (2012). Cueing effects of target location probability and repetition. *Vision Research*, *73*, 23–29. https://doi.org/10.1016/j.visres.2012.09.014

Kruijne, W., & Meeter, M. (2015). The long and the short of priming in visual search. *Attention, Perception, & Psychophysics*, *77*(5), 1558–1573. https://doi.org/10.3758/s13414-015-0860-2

Langner, R., Steinborn, M. B., Eickhoff, S. B., & Huestegge, L. (2018). When specific action biases meet nonspecific preparation: Event repetition modulates the variable-foreperiod effect. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(9), 1313–1323. https://doi.org/10.1037/xhp0000561

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., & 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

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron*, *71*(4), 737–749. https://doi.org/10.1016/j.neuron.2011.07.012

Machado, A. (1997). Learning the temporal dynamics of behavior. *Psychological Review*, *104*(2), 241–265. https://doi.org/10.1037/0033-295X.104.2.241

Maljkovic, V., & Martini, P. (2005). Implicit short-term memory and event frequency effects in visual search. *Vision Research*, *45*(21), 2831–2846. https://doi.org/10.1016/j.visres.2005.05.019

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. role of features. *Memory & Cognition*, *22*(6), 657–672. https://doi.org/10.3758/BF03209251

Mareschal, D., & French, R. M. (2017). TRACX2: A connectionist autoencoder using graded chunks to model infant visual statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160057. https://doi.org/10.1098/rstb.2016.0057

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

Mauk, M. D., & Ruiz, B. P. (1992). Learning-dependent timing of pavlovian eyelid responses: Differential conditioning using multiple interstimulus intervals. *Behavioral Neuroscience*, *106*(4), 666–681. https://doi.org/10.1037/0735-7044.106.4.666

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86. https://doi.org/10.1016/0010-0285(86)90015-0

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

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin*, *89*(1), 133–162. https://doi.org/10.1037/0033-2909.89.1.133

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1–32. https://doi.org/10.1016/0010-0285(87)90002-8

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

Olson, I. R., & Chun, M. M. (2001). Temporal contextual cuing of visual attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(5), 1299–1313. https://doi.org/10.1037/0278-7393.27.5.1299

O'Reilly, J. X., McCarthy, K. J., Capizzi, M., & Nobre, A. C. (2008). Acquisition of the temporal and ordinal structure of movement sequences in incidental learning.

Journal of Neurophysiology, *99*(5), 2731–2735. https://doi.org/10.1152/jn.01141.2007

Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsáki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science*, *321*(5894), 1322–1327. https://doi.org/10.1126/science.1159775

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238. https://doi.org/10.1016/j.tics.2006.03.006

Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language*, *39*(2), 246–263. https://doi.org/10.1006/jmla.1998.2576

Prut, Y., & Fetz, E. E. (1999). Primate spinal interneurons show pre-movement instructed delay activity. *Nature*, *401*(6753), 590. https://doi.org/10.1038/44145

Rabbitt, P., Cumming, G., & Vyas, S. (1979). Modulation of selective attention by sequential effects in visual search tasks. *Quarterly Journal of Experimental Psychology*, *31*(2), 305–317. https://doi.org/10.1080/14640747908400729

Saffran, J. R., Aslin, R. N., & work(s): E. L. N. R. (1996). Statistical learning by 8-month-old infants. *Science, New Series*, *274*(5294), 1926–1928. http://www.jstor.org/stable/2891705

Salet, J. M., Kruijne, W., & van Rijn, H. (2021). Implicit learning of temporal behavior in complex dynamic environments. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-020-01873-x

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*. https://doi.org/10.1037/rev0000356

Salet, J. M., Schlichting, N., Kruijne, W., & van Rijn, H. (2022). Addendum: Implicit learning of temporal behavior in complex dynamic environments. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-022-02194-x

Salet, J. M., de Jong, J., & van Rijn, H. (2022). Still stuck with the stopwatch [Publisher: American Psychological Association]. *Behavioral Neuroscience*, *136*(5), 453–466. https://doi.org/10.1037/bne0000527

Sali, A. W., Anderson, B. A., & Yantis, S. (2015). Learned states of preparatory attentional control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1790–1805. https://doi.org/10.1037/xlm0000146

Salidis, J. (2001). Nonconscious temporal cognition: Learning rhythms implicitly. *Memory & Cognition*, *29*(8), 1111–1119. https://doi.org/10.3758/BF03206380

Sanchez, D. J., Yarnik, E. N., & Reber, P. J. (2015). Quantifying transfer after perceptual-motor sequence learning: How inflexible is implicit learning? *Psychological Research*, *79*(2), 327–343. https://doi.org/10.1007/s00426-014-0561-9

Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: A neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160049. https://doi.org/10.1098/rstb.2016.0049

Selchenkova, T., Jones, M. R., & Tillmann, B. (2014). The influence of temporal regularities on the implicit learning of pitch structures. *Quarterly Journal of Experimental Psychology*, *67*(12), 2360–2380. https://doi.org/10.1080/17470218.2014.929155

Shalev, N., Boettcher, S., Wilkinson, H., Scerif, G., & Nobre, A. C. (2022). Be there on time: Spatial-temporal regularities guide young children's attention in dynamic

environments [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cdev.13770]. *Child Development*, *93*(5), 1414–1426. https://doi.org/10.1111/cdev.13770

Shankar, K. H., & Howard, M. W. (2012). A scale-invariant internal representation of time. *Neural Computation*, *24*(1), 134–193. https://doi.org/10.1162/NECO_a_00212

Shaw, M. L., & Shaw, P. (1977). Optimal allocation of cognitive resources to spatial locations. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(2), 201–211. https://doi.org/10.1037/0096-1523.3.2.201

Sherman, B. E., Graves, K. N., & Turk-Browne, N. B. (2020). The prevalence and importance of statistical learning in human cognition and behavior. *Current Opinion in Behavioral Sciences*, *32*, 15–20. https://doi.org/10.1016/j.cobeha.2020.01.015

Shin, J. C., & Ivry, R. B. (2002). Concurrent learning of temporal and spatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 445–457. https://doi.org/10.1037/0278-7393.28.3.445

Shin, J. C., & Ivry, R. B. (2003). Spatial and temporal sequence learning in patients with parkinson's disease or cerebellar lesions [Publisher: MIT Press]. *Journal of Cognitive Neuroscience*, *15*(8), 1232–1243. https://doi.org/10.1162/089892903322598175

Sinclair, C., & Hammond, G. R. (2008). Reduced intracortical inhibition during the foreperiod of a warned reaction time task. *Experimental Brain Research*, *186*(3), 385–392. https://doi.org/10.1007/s00221-007-1241-4

Steinborn, M. B., & Langner, R. (2012). Arousal modulates temporal preparation under increased time uncertainty: Evidence from higher-order sequential foreperiod effects. *Acta Psychologica*, *139*(1), 65–76. https://doi.org/10.1016/j.actpsy.2011.10.010

Thiessen, E. D. (2017). What's statistical about learning? insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1711), 20160056. https://doi.org/10.1098/rstb.2016.0056

Thiessen, E. D., & Pavlik Jr., P. I. (2013). iMinerva: A mathematical model of distributional statistical learning. *Cognitive Science*, *37*(2), 310–343. https://doi.org/10.1111/cogs.12011

Thomaschke, R., Hoffmann, J., Haering, C., & Kiesel, A. (2016). Time-based expectancy for task relevant stimulus features. *Timing & Time Perception*, *4*(3), 248–270. https://doi.org/10.1163/22134468-00002069

Thomaschke, R., Wagener, A., Kiesel, A., & Hoffmann, J. (2011). The specificity of temporal expectancy: Evidence from a variable foreperiod paradigm. *Quarterly Journal of Experimental Psychology*, *64*(12), 2289–2300. https://doi.org/10.1080/17470218.2011.616212

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, *21*(10), 1934–1945. https://doi.org/10.1162/jocn.2009.21131

Wagener, A., & Hoffmann, J. (2010). Temporal cueing of target-identity and target-location. *Experimental Psychology*, *57*(6), 436–445. https://doi.org/10.1027/1618-3169/a000054

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of p values. *Psychonomic Bulletin & Review*, *14*(5), 779–804. https://doi.org/10.3758/BF03194105

Walthew, C., & Gilchrist, I. D. (2006). Target location probability effects in visual search: An effect of sequential dependencies. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1294–1301. https://doi.org/10.1037/0096-1523.32.5.1294

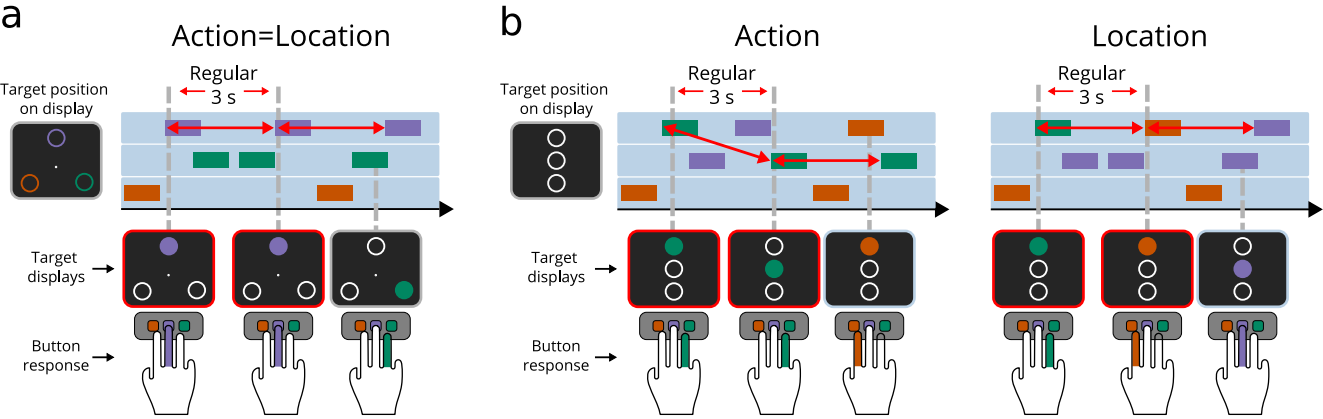
Wixted, J. T. (2004). On common ground: Jost's (1897) law of forgetting and ribot's (1881) law of retrograde amnesia. *Psychological Review*, *111*(4), 864–879. https://doi.org/10.1037/0033-295X.111.4.864

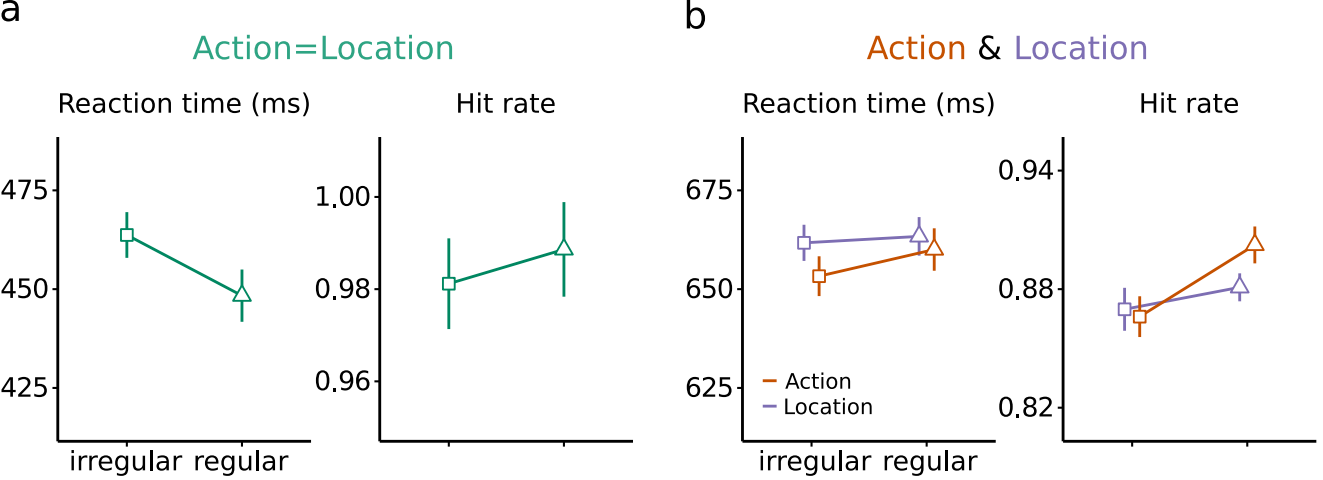
Wixted, J. T., & Ebbesen, E. B. (1991). On the form of forgetting. *Psychological Science*, *2*(6), 409–415. https://doi.org/10.1111/j.1467-9280.1991.tb00175.x

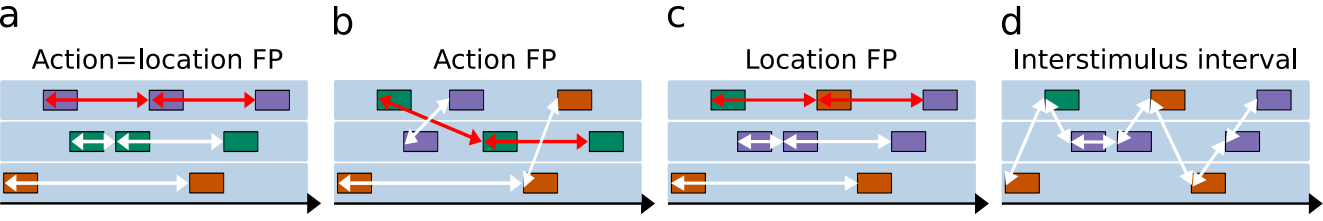
Xu, Z., Theeuwes, J., & Los, S. A. (2022). Statistical learning of spatiotemporal regularities dynamically guides visual attention across space. *Attention, Perception, & Psychophysics*. https://doi.org/10.3758/s13414-022-02573-5

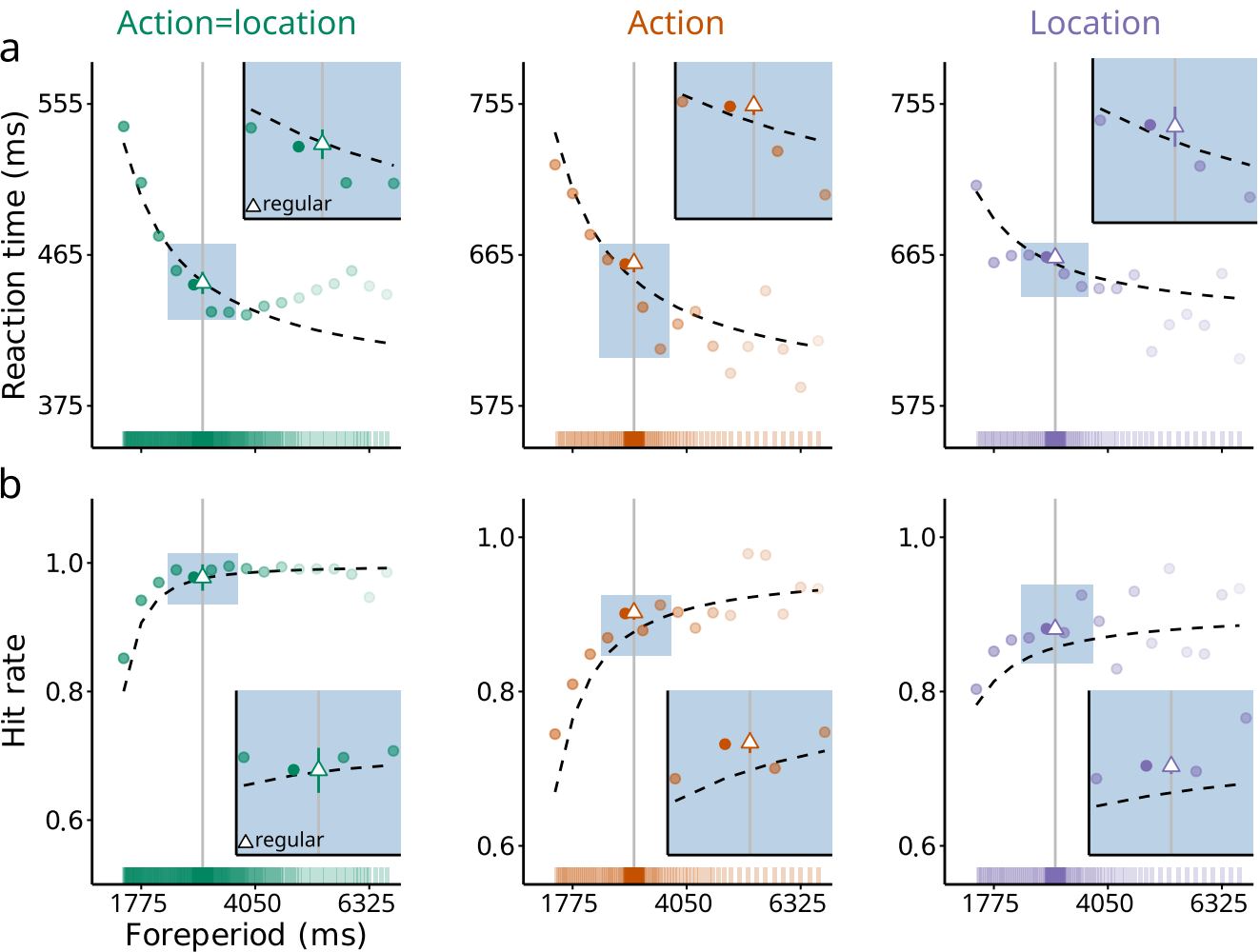
Yu, R. Q., & Zhao, J. (2015). The persistence of the attentional bias to regularities in a changing environment. *Attention, Perception, & Psychophysics*, *77*(7), 2217–2228. https://doi.org/10.3758/s13414-015-0930-5

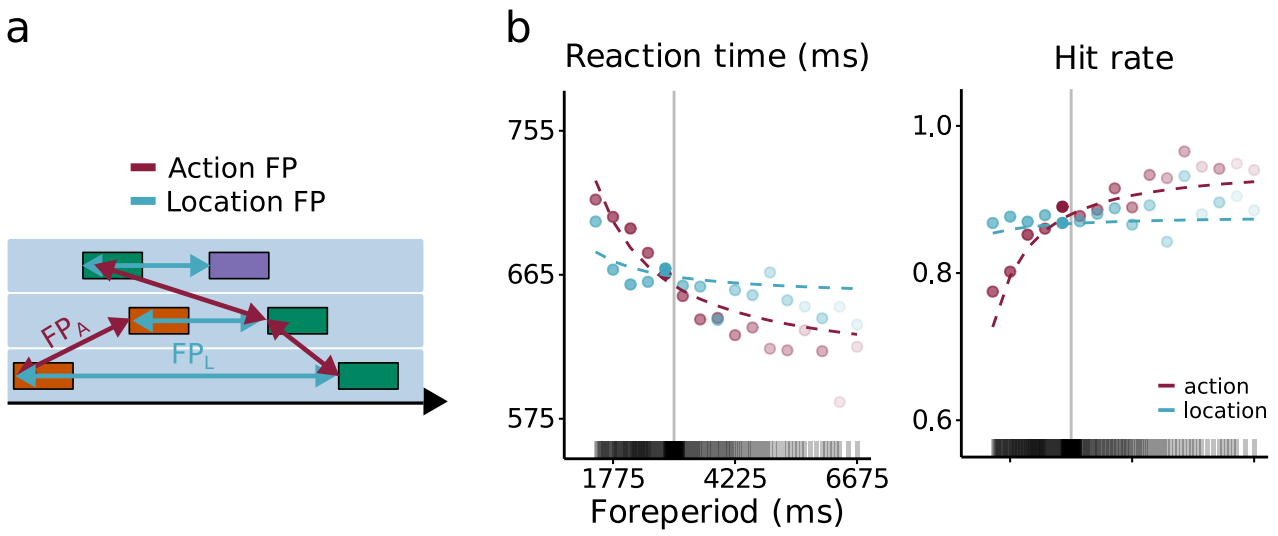
Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (2013). Attention is spontaneously biased toward regularities. *Psychological science*, *24*(5), 667–677. https://doi.org/10.1177/0956797612460407

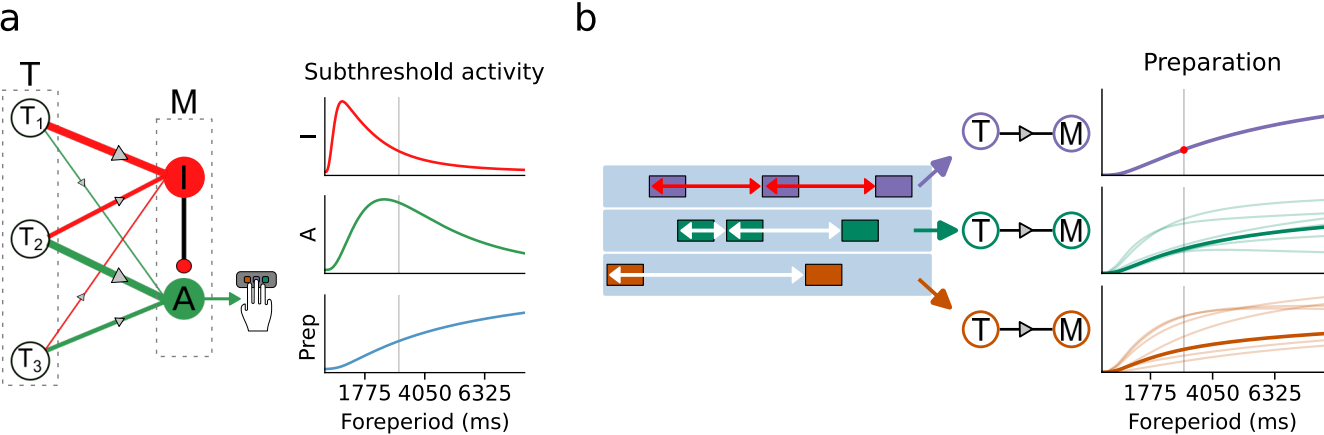

a
Action=Location
Regular
3 s
Target position on display
Target displays
Button response
b
Action
Regular
3 s
Target position on display
Target displays
Button response
Location
Regular
3 s

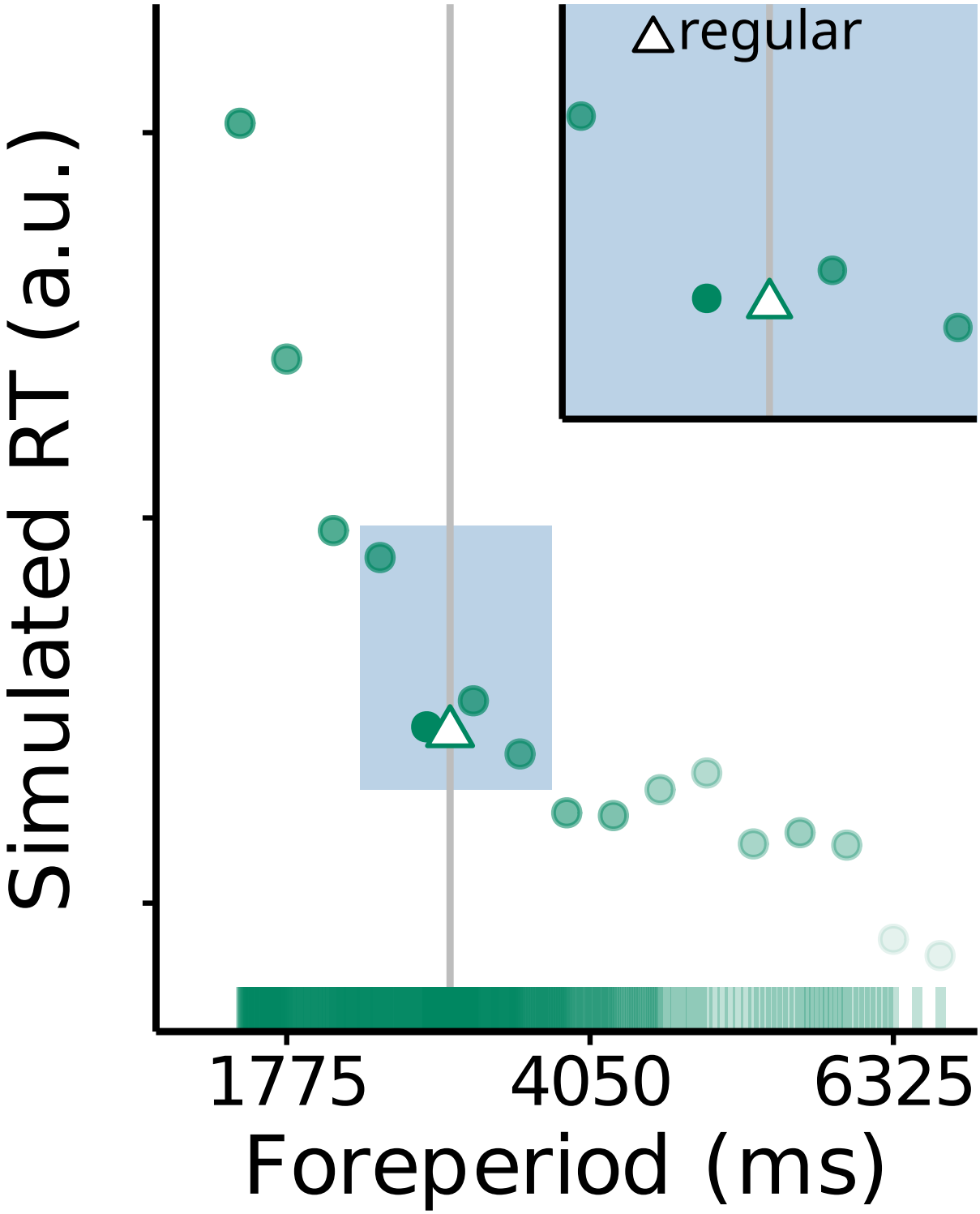

a
Action=Location
Reaction time (ms)
Hit rate
475
450
425
1.00
0.98
0.96
irregular
regular
b
Action & Location
Reaction time (ms)
Hit rate
675
650
625
0.94
0.88
0.82
Action
Location
irregular
regular

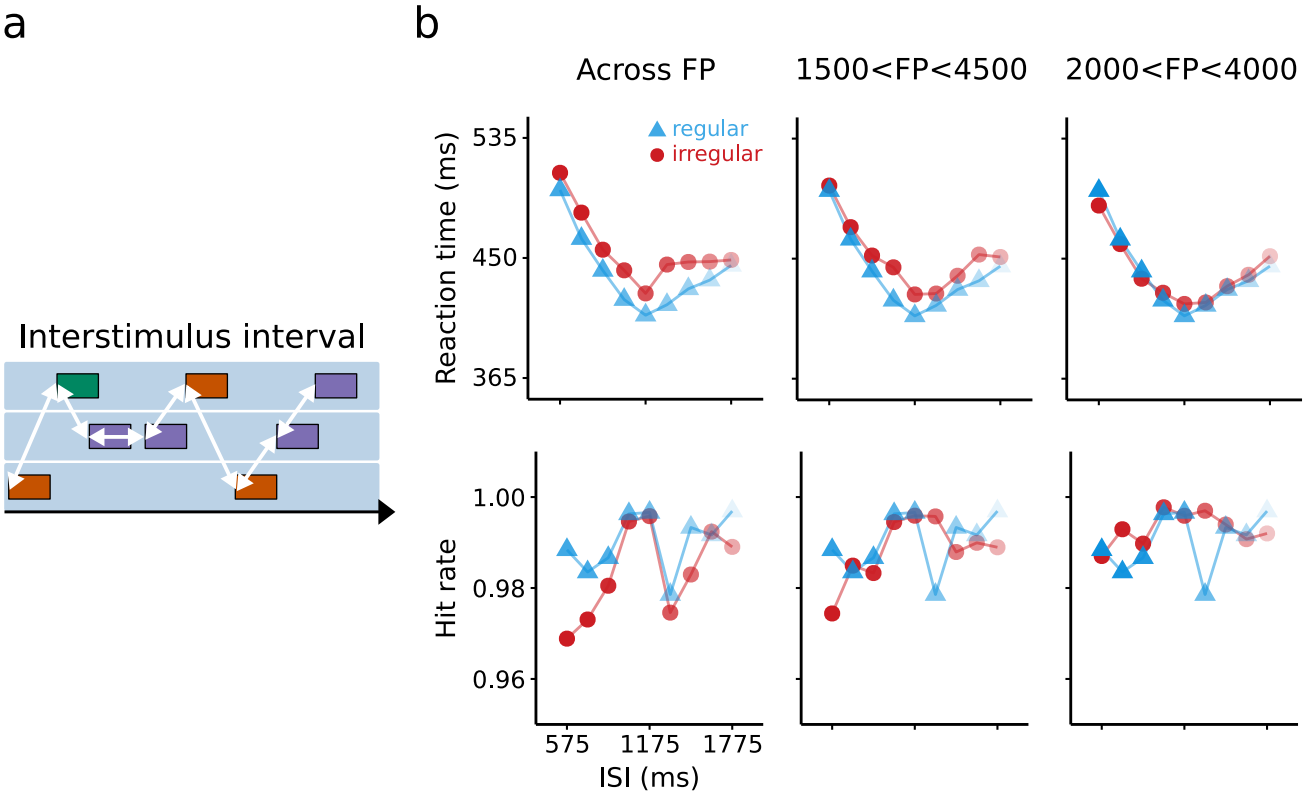

a
Action=location FP
b
Action FP
c
Location FP
d
Interstimulus interval

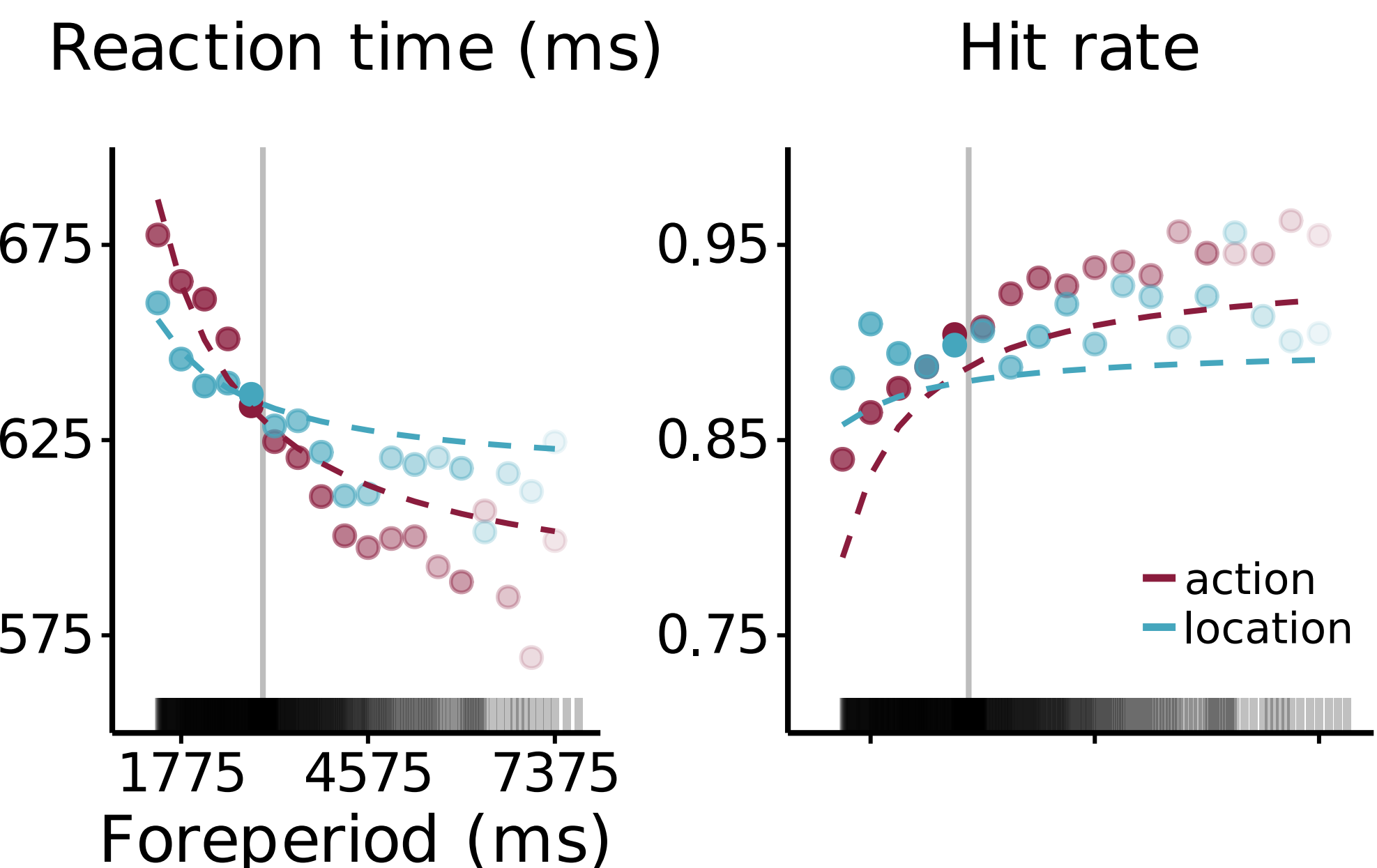

Action=location
Action
Location
a
b
Reaction time (ms)
555
465
375
755
665
575
755
665
575
Hit rate
1.0
0.8
0.6
1.0
0.8
0.6
1.0
0.8
0.6
regular
regular
1775
4050
6325
1775
4050
6325
1775
4050
6325
Foreperiod (ms)

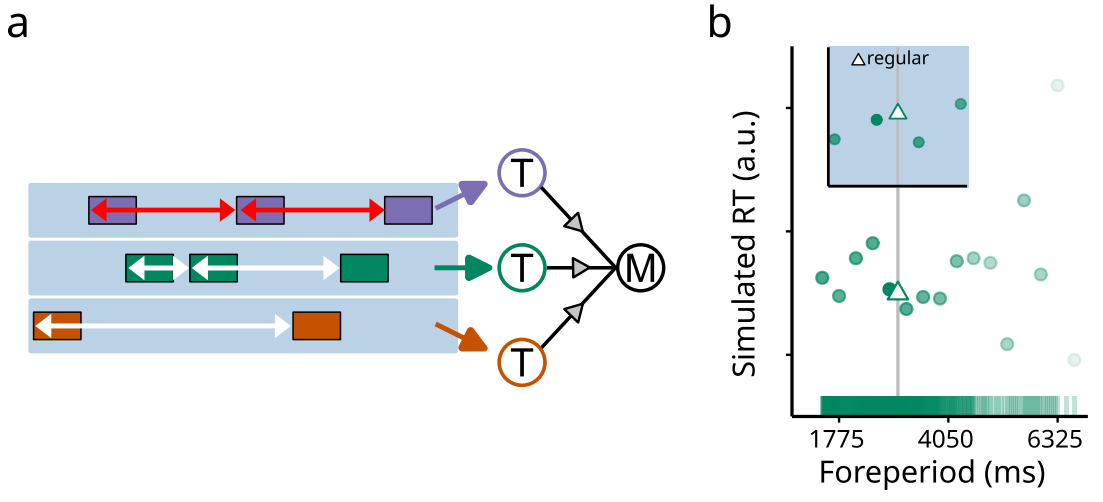

a
Action FP
Location FP
FP_A
FP_L
b
Reaction time (ms)
755
665
575
1775
4225
6675
Foreperiod (ms)
Hit rate
1.0
0.8
0.6
action
location


a
T
M
T1
T2
T3
I
A
Subthreshold activity
I
A
Prep
1775 4050 6325
Foreperiod (ms)
b
Preparation
T
M
1775 4050 6325
Foreperiod (ms)


△regular
Simulated RT (a.u.)
1775
4050
6325
Foreperiod (ms)


a
Interstimulus interval
b
Across FP
1500<FP<4500
2000<FP<4000
regular
irregular
Reaction time (ms)
535
450
365
Hit rate
1.00
0.98
0.96
575
1175
1775
ISI (ms)


Reaction time (ms)
675
625
575
1775
4575
7375
Foreperiod (ms)
Hit rate
0.95
0.85
0.75
action
location


a
T
T
T
M
b
regular
Simulated RT (a.u.)
1775
4050
6325
Foreperiod (ms)