Anticipation of Time Spans: New Data from the Foreperiod Paradigm and the Adaptation of a Computational Model

Johannes Lohmann, Oliver Herbort, Annika Wagener, and Andrea Kiesel

University of Würzburg
Department of Psychology
Roentgenring 11
97070 Würzburg
johannes.lohmann@stud-mail.uni-wuerzburg.de,
{oliver.herbort,wagener,kiesel}@psychologie.uni-wuerzburg.de

Abstract. To act successfully, it is necessary to adjust the timing of one's behavior to events in the environment. One way to examine human timing is the foreperiod paradigm. It requires experimental participants to react to events that occur at more or less unpredictable time points after a warning stimulus (foreperiod). In the current article, we first review the empirical and theoretical literature on the foreperiod paradigm briefly. Second, we examine how behavior depends on either a uniform or peaked (at 500ms) probability distribution of many (15) possible foreperiods. We report adaptation to different probability distribution with a pronounced adaptation for the peaked (more predictable) distribution. Third, we show that Los and colleagues' [1] computational model accounts for our results. A discussion of specific findings and general implications concludes the paper.

1 Introduction

To act successfully it is not only necessary to behave in a skillful way, but also the timing of the behavior is crucial. On a macroscopic timescale, timing of buys and sells on the real estate market can make a considerable difference. On a smaller timescale, waving down a bus or catching a ball requires us to initiate movements in time. On an ever smaller timescale, in sports, like tennis or baseball, the precise timing of a stroke is of paramount importance. And finally, even the eye blink reflex may be adjusted by mere milliseconds.

Acting successfully is even more complex because the events on which we need to react are not always fully predictable. Often, we have to learn which warning signals precede critical events and the duration of the time interval between both. Learning helps to anticipate the onset of critical events and to prepare or adjust behavior to react quickly and adequately. Thus, to understand how humans excel at a broad range of tasks and skills, we need to understand how humans adapt their behavior in time, when events occur more or less predictable.

G. Pezzulo et al. (Eds.): ABiALS 2008, LNAI 5499, pp. 170–187, 2009.

1.1 Investigation of Behavior in Time

In the lab, behavior in time has been systematically studied with the foreperiod paradigm [2,3], for a review see [4]. In these experiments, a human participant has to react as quickly as possible to a target stimulus, like the onset of a visual stimulus or a sound. To enable the formation of temporal anticipations, a warning stimulus (WS) appears at a certain point in time before the target stimulus. The interval between the WS and the target stimulus is called the foreperiod (FP).

If the FP stays constant for a while, participants are able to form expectations about the time of the appearance of the target stimulus and thus react faster upon it. Interestingly, they react the faster the shorter the FP, as the temporal resolution is higher for shorter FPs [5,6]. Only if FPs are very short, between 0 and 100ms, RTs rise again [2].

If the FP varies unpredictably from trial to trial, it is not possible to anticipate the exact time point of stimulus onset. Nevertheless, participants still have some information to prepare their response. For example, as time passes by and the target stimulus has not yet occured, the time window in which the target stimulus might appear shrinks. Thus, in the case of unpredictable FPs, participants react the faster the longer the FP [7,2,8,9].

In addition, the data reveal strong sequential effects. Compared to repetition trials (subsequent trials with identical FPs) reaction times (RT) increase if the preceding FP was longer than the current FP. The effect seems to exist only in one direction, because the preceding FP does not affect RTs if it was shorter than the current FP. Hence, RT increase if the current FP is unexpectedly shorter while RTs are unaffected if the current FP is longer than expected [10,11,12] (but see [13] for contradictory results).

The FP paradigm is well established and there is a substantial body of data that describes how humans anticipate upcoming events and adapt to the predictability of those events [14,15,16,17,18,19,20]. The aim of the current article is threefold. First, we review current theories of timing. Second, we provide new experimental data that reveals how human behavior adapts to different probability distributions of FPs. Third, we test if the data can be explained by a current model of timing [1].

The remainder of the article is structured as follows. The next section reviews different theories of timing. Then, the behavioral experiment and the novel empirical data will be described. After that, a mathematical formulation of a computational model of timing will be given and it will be compared to empirical data. A short discussion concludes the paper.

2 Theories of Timing

By now, mainly two theories of timing emerged to account for behavior in the FP paradigm: Gibbon's "scalar expectancy theory" (SET) [21] and the "behavioral theory of timing" (BeT) by Killen and Fetterman [22]. The SET is a cognitive

approach that explains temporal regularities of learned behavior by a number of information processing devices. An internal pacemaker generates variable pulses with a high frequency [23]. An accumulator accumulates the pulses up to a critical event. The number of accumulated pulses is stored in longterm–memory. To recall or reproduce a certain duration, the memorized number of pulses is compared to the currently accumulated number of pulses. The relative discrepancy of these two values determines behavior, mediated by adjustable thresholds.

In contrast, the BeT conceives the organism as moving across an invariant series of "behavioral classes" between a WS and a target stimulus. Like in SET an internal pacemaker generates pulses that cause the organism to cycle through the behavioral classes. When the target stimulus appears, the currently active class is reinforced. When the organism perceives a WS later on, it again starts to cycle through the behavioral classes. As its behavioral intensity is partially determined by the activity of the currently active behavioral class, it is then able to adjust its behavior to the experienced FP. The development of BeT as well as SET stimulated the quantitative description of behavior in time.

However, crucial aspects of timing remained unexplained. The adjustable thresholds in SET and the discriminative function of behavioral classes in BeT imply a learning process, but both theories fail to specify one. Hence, Machado [24] reformulated the BeT as a mathematical model, specifying two mechanisms of adaptation: reinforcement and exinction (for other approaches see [25,26,27]). Finally, Los and colleagues [1,28] reinterpreted the output of the model to account for human RTs. Their formulation of the BeT makes four assumptions:

1. Peaks of activation develop around the possible moments, at which the target stimulus may appear. The more FPs are used in a experimental design, the more peaks can be expected.
2. However, the temporal resolution is limited and degrades for longer intervals. Activation peaks become broader and flatter as they are more remote from the WS.
3. Reinforcement only occurs if the peak coincides with the time point of the occurrence of the target stimulus.
4. Extinction occurs at any peak that is associated with a moment prior to the relevant moment. Peaks of time points after the appearance of the target stimulus remain unchanged.

The model conceives RTs as inverse proportional to the activation at the moment the target stimulus occurs. The four assumptions explain most of the observed effects. If FPs are predictable a single activation peak is reinforced and will quickly reach its maximal amplitude. This results in faster reactions if the target stimulus appears at the expected time point. If FPs vary unpredictably from trial to trial, all FPs are reinforced or subject to extinction from time to time. Because peaks associated to shorter FPs are activated more frequently, they are also more often discounted than reinforced. This results in higher RTs for shorter FPs.

Assumption three and four explain sequential effects on RTs. All peaks that are associated with time points in the FP of one trial are subject to extinction and only the peak associated to the actual FP is reinforced in the respective trial. If the subsequent trial requires a reaction at one of these early time points, the RT tends to be higher because the respective peaks have just recently been decreased. If the FP of a subsequent trial is longer then RTs are only slightly affected.

To conclude, our understanding of human timing is based on FP experiments and models that assume that different time intervals are represented by discrete peaks of activity. This raises further questions regarding the experimental methodology and the theoretical models. First, in many FP experiments, FPs were either completely predictable or unpredictable (e.g. [1,18]). However, in everyday life, timing intervals are usually distributed around a specific duration. In most cases an interval has a certain duration but also shorter or longer intervals might be experienced from time to time. Thus, we are interested in how humans adapt their anticipation to a single-peaked probability distribution of possible FPs. Second, in most experiments, participants are exposed to a limited number of more or less distinguishable FPs. However, in many situations not only some but a continuum of FPs may be expected. Thus, we examine if the same effects can be observed if 15 possible FPs are applied in an experiment. Finally, we want to test if the computational model of Los and Agter [18] is also valid for peaked, quasi-continous distribution of FPs, or if the model needs to be further refined.

In the next section, the experimental protocol and results are described. We then give a mathematical description of the computational model we used and test the model on our data.

3 Foreperiod Experiment

In the following section, we report the protocol and results of two experiments. According to the FP paradigm, participants had to press a key upon the appearance of a a target stimulus. A WS preceded the target stimulus at random and thus unpredictable FPs, which ranged between 100ms and 1500ms. Both experiments differed in the probability distribution of the different possible FPs. In the first experiment, each FP had the same probability (uniform distribution), while in the second experiment, the 500ms FP was much more frequent than any other (peaked distribution). To measure the degree to which participants adapted their behavior, we recorded RTs, with shorter RTs indicating better timing. To evaluate the general adaptation to the different probability distributions, we compared RTs at different FPs. To further evaluate the short-term adaptation based on single trials, we analyzed sequential effects, that is, we compared RTs dependent on the FP of the current and the preceding trial.

3.1 Experimental Method

Participants. In each experiment, ten participants (uniform: 8 women and 2 men, age 19-22; peaked: 7 women and 3 men, age 19-25) volunteered to either

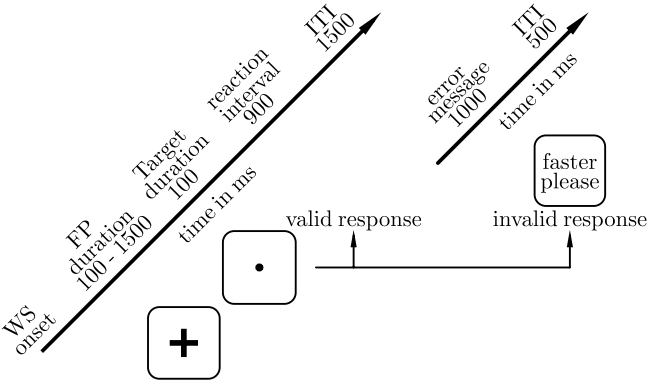

Fig. 1. The schematic time course of one trial (ITI: intertrial interval)

satisfy course requirements or in exchange for pay. All participants reported having normal or corrected-to-normal vision and were not familiar with the purpose of the experiment.

Apparatus and Stimuli. Stimuli were displayed on a 17 inch CRT monitor and RTs were recorded with an IBM-compatible computer (Pentium IV with 2.6 GHz) running E-Prime [29]. Figure 1 illustrates the trial procedure. Each trial started with the presentation of a fixation cross. The onset of the fixation cross also marked the onset of the FP. We used fifteen different FPs: 100ms, 200ms, ... , and 1500ms. After the FP, a circle (approximately 2 cm x 2 cm) was displayed as target stimulus for 100ms, followed by a blank interval of 900ms. All stimuli appeared in white on dark-grey background. Participants had to press a key with the right index finger upon appearance of the target. If participants responded within 1000ms to the target stimulus, the screen stayed blank for another 1500ms, then the next trial began. If participants failed to respond, the German words "bitte schneller" (faster, please) were displayed in red letters for 1000 ms and the next trial was initiated another 500ms later. Both experiments consisted of ten blocks with 120 trials each. Table 1 lists the distribution of foreperiods in each block for both experiments. In the uniform distribution experiment, all FP appeared with the same probability. In contrast, in the peak distribution experiment, the FP of 500ms was 46 times as likely as any other FP. The presentation order of all trials were randomized for each block and were thus unpredictable for the participants.

3.2 Experimental Results

Uniform Distribution Experiment. We aggregated data from every three FPs, resulting in the five different FP ranges (see left panel of Table 1) to enable

Table 1. Frequencies of FPs, FP ranges, and data points in each FP range (Freq) in each block of the peak and the uniform distribution experiment

	Uniform Distribution			Peaked Distribution		
FP (ms)	Freq_{FP}	FP_{Range}	$\text{Freq}_{\text{FPrange}}$	Freq_{FP}	FP_{Range}	$\text{Freq}_{\text{FPrange}}$
100	8	100 - 300		2	100 - 400	
200	8	↕	24	2		
300	8	100 - 300		2	↕	8
400	8	400 - 600		2	100 - 400	
500	8	↕	24	92	500	92
600	8	400 - 600		2	600 - 1000	
700	8	700 - 900		2		
800	8	↕	24	2	↕	10
900	8	700 - 900		2		
1000	8	1000 - 1200		2	600 - 1000	
1100	8	↕	24	2	1100 - 1500	
1200	8	1000 - 1200		2		
1300	8	1300 - 1500		2	↕	10
1400	8	↕	24	2		
1500	8	1300 - 1500		2	1100 - 1500	

better illustration and statistical analysis. We analyzed the RT and response validity[1] data using ANOVAs[2] with two within-subject factors: the FP range of the trial (FP range_{trial}) and the FP range of the preceding trial (FP $\text{range}_{trial-1}$).

Figure 2 shows RT (A, B) and response validity data (C, D). The single means for the different FP ranges are typical for unpredictable FPs: the shorter the FP, the higher the RTs, $F(4, 36) = 57.9,\ p < .001$. Also, the FP range of the preceding trial (FP $\text{range}_{trial-1}$) has a significant influence on RT, resulting in generally shorter RTs in a trial if the preceding trial also had a short FP, $F(4, 36) = 21.7,\ p < .001$. Both factors are not independent but interact: The impact of FP $\text{range}_{trial-1}$ on RT data decreases with increasing FP range_{trial}, $F(16, 144) = 4.2,\ p < .01$.

Response validity data roughly follows the same pattern (Fig. 2C, D). Error rates increase with the increasing FP range_{trial}, $F(4, 36) = 4.8,\ p < .05$. Additional, there is a significant influence of FP $\text{range}_{trial-1}$, $F(4, 36) = 0.1,\ p < .01$, but the interaction failed to reach significane, $F(16, 144) = 2.2,\ p = 0.11$.

In general, the results fit nicely to existing data and model assumptions. Participants respond the faster, the longer the FP in each trial. In addition, the data reveals asymmetric sequential effects. If the current FP is short (FP

[1] A trial response was invalid if the participant pressed the key either before the onset of the target stimulus, more than 1000ms after its onset, or not at all. As most invalid responses were due to premature key presses, invalid responses mostly reflect a higher behavioral activation.

[2] We applied the Greenhouse–Geisser correction because the assumption of sphericity was violated in our data. For clarity, we report F-values with unadjusted degrees of freedom.

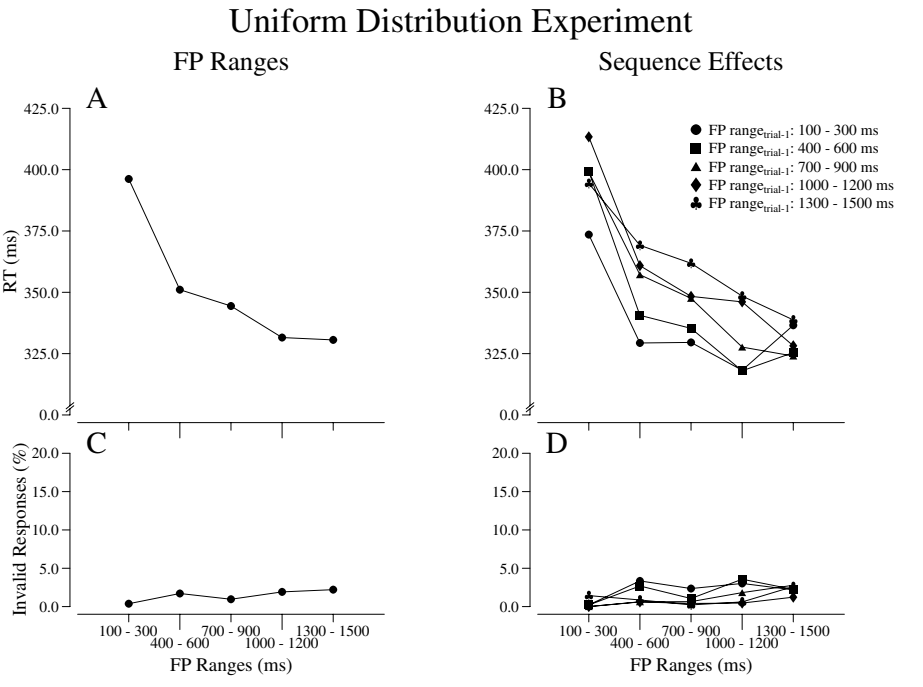

Fig. 2. The charts show the results of the uniform distribution experiment: the impact of the FP range on RT (A) and response validity (C) and the impact of specific sequences of FPs on RT (B) and response validity (D)

range 100-300) the impact of the previous FP seems to be much more severe than when the current FP is long. However, our data does not show an ordered influence of the preceding FP, as would be theoretically expected. Finally, in this experiment, the repetition of the same FP did not necessarily cause the greatest benefits for RT. Especially higher FP ranges showed the steepest increase in RT if the preceeding trial's FP was slightly longer.

Peaked Distribution Experiments. Again, we analyzed the data on the level of aggregated FP ranges as displayed in Table 1 using ANOVAs with FP range$_{trial}$ and FP range$_{trial-1}$ as within subject factors. Note that due to the non-uniform distribution of FPs (most of the trials had a FP of 500ms) the different FP ranges consist of different numbers of data points. Figure 3A, B show that RTs decrease with increasing FP range$_{trial}$, $F(3, 27) = 31.0,\ p < .001$ and depended on the preceding trials FP (FP range$_{trial-1}$), $F(3, 27) = 3.8,\ p < .05$. There is no interaction between FP range$_{trial}$ and FP range$_{trial-1}$ $F(9, 81) = 0.6,\ p = 0.67$. Interestingly, RTs seem to be generally much faster than in the uniform distribution experiment. Especially, the decrease from the shortest FP range to the next one is much more pronounced in the peaked distribution experiment than for the uniform distribution experiment, but RT decrease further

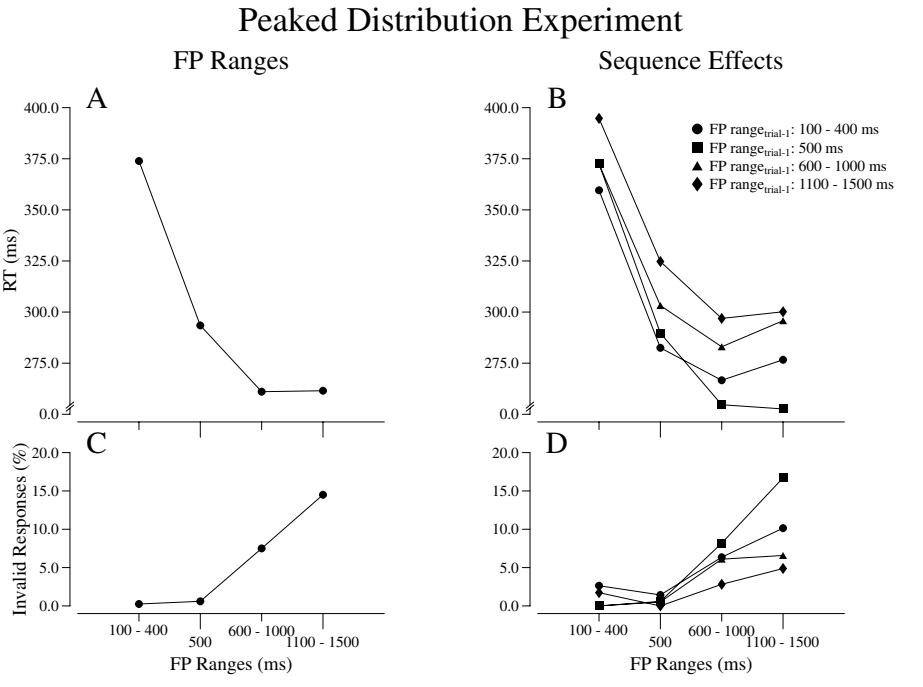

Fig. 3. The charts show the results of the peaked distribution experiment: the impact of the FP range on RT (A) and response validity (C) and the impact of specific sequences of FPs on RT (B) and response validity (D)

for the higher FP ranges. This may be due the high behavioral activation for the frequent FP of 500ms, which seems to be maintained for higher FP. The same trend can also be found in the response validity data, $F(3, 27) = 8.3,\ p < .01$. However, there was neither a significant main effect for FP range$_{trial-1}$ nor a significant interaction for response validity data, $F(3, 27) = 0.9,\ p = .48$, $F(9, 81) = 1.5,\ p = .23$. The analysis of sequence effects revealed an RT advantage if the preceding trial contained one of the shorter FP ranges. Participants responded especially fast in trials that followed a trial with the frequent FP of 500ms. We assume that the comparatively slow reactions following trials that did not contain the frequent FP 500ms may be attributed to RT costs caused by expectancy violations after trials with uncommon FPs.

4 Simulation

The following section mathematically formulates Machado's / Los and colleagues' model [1,24]. The model has a serial structure with interconnected timing nodes. Every node has two connections, one to the subsequent node and one to an output node. The links to the output node, which determines RT, are weighted, the weights are adjustable through learning processes. After the occurrence of a

WS activation is propagated through the nodes in the structure. Depending on the time elapsed since the WS occurred the single nodes of the serial structure contain a different amount of activation. Hence their contribution to the output differs over time, with a characteristic activation peak for every node. The basic structure of the model is shown in Figure 4. In this section, we now describe the propagation of activity through the nodes, the learning rules, and the response rule.

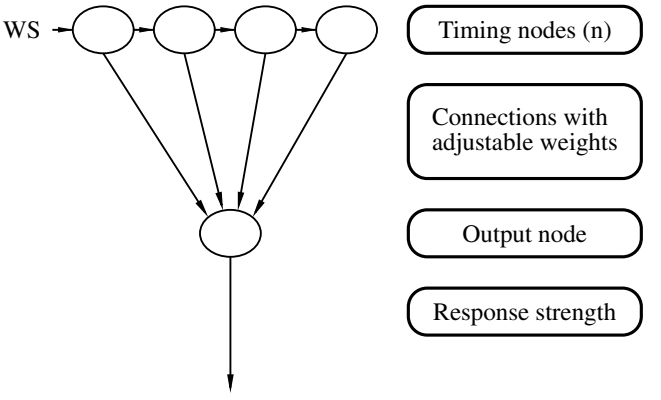

Fig. 4. The basic structure of the model adapted from Machado [24]

4.1 Formal Outline of the Model

Node Activations. When the WS appears, all activity is contained in the first node $X_0(t)$ and the remaining nodes with an n greater than 0 are not activated at all: $X_n(t) = 0 \quad n = 1, \ldots, N$, (Fig. 5A). This activation is then propagated through the system as time passes. The current activation of each of the remaining nodes depends on the activation that a given node receives from its predecessor and the activation it passes on to its successor. Figure 6A illustrates this process of a constant flow of activation by Machado's cascade analogy. The flow of activation is modeled by the following differential equations:

$$\frac{\delta}{\delta t} X_0(t) = -\lambda X_0(t) \tag{1}$$

$$\frac{\delta}{\delta t} X_n(t) = \lambda X_{n-1}(t) - \lambda X_n(t) \quad \text{for } n = 1, \ldots, N \tag{2}$$

where λ describes the time range and the speed of the activity propagation. The solution of (1) and (2) leads to the poisson density function

$$X_n(t) = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \quad . \tag{3}$$

The activation of one state over time can be described as a poisson process. With the exception of the first node, the activity in each node $X_n(t)$ rises continuously,

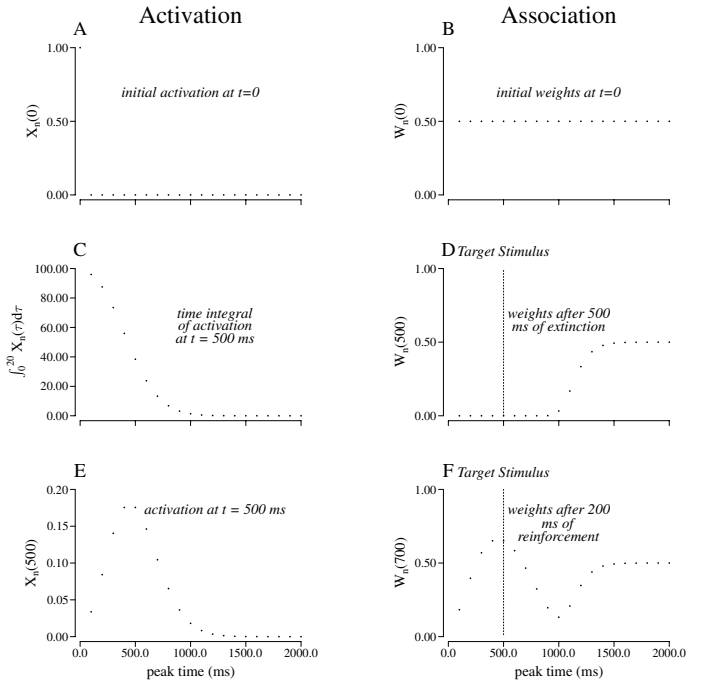

Fig. 5. The charts show different aspects of the model, occurring in a single trial with an FP of 500 ms. A: activation at $t = 0$; B: initial weight distribution; C: cumulative activation of nodes at the onset of the target stimulus ($t = 500ms$); D: Decreased weights (extinction) for nodes associated to time points before the onset of the target stimulus; E: activation of nodes at the onset of the target stimulus; F: Increased weights (reinforcement) after the end of the reinforcement period ($\lambda = 0.01,\ n = 20,\ \alpha = 2,\ \beta = 0.03,\ K = 1,\ d = 200ms$).

peaks at $t = \frac{n}{\lambda}$, and decreases afterward. The sum of activation is constant in the system, because the area of the poisson density function is $\frac{1}{\lambda}$, independent of the value of n . However, mean and variance of the activation curves are proportional to n yielding flatter and broader peaks for larger values of n. Consequently, the temporal resolution is high for short FPs and decreases for longer FPs. Figure 6B shows some activation curves for different values of n and $\lambda = 0.01$.

Extinction and Reinforcement. The weights of the links between the timing nodes and the output node are adjustable. The following section introduces the mathematical formulation of the learning rules.

The weight $W_n(t)$ of the connection between node n and the output node is subject to extinction and reinforcement during each trial. Initially, before the onset of the first trial of the experiment, no specific FP distribution can be expected and all weights are set to $W_n(0) = 0.5,\ n = 1, \ldots, N$ (Fig. 5B). All later trials start with the weight distribution that resulted from the previous trial.

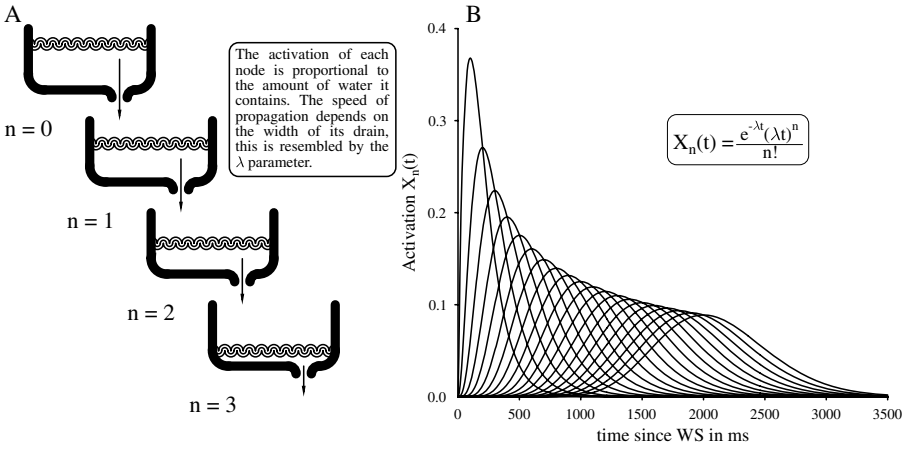

Fig. 6. A: The cascade analogy illustrates the propagation of activation through the series of nodes. B: The chart shows the poisson density distribution for different values of n and $\lambda = 0.01$.

Extinction takes place from the onset of the WS until the onset of the target stimulus. The decrease of the weights depends on the activation a node received during a trial, the activation of the node at the time the target stimulus occurs, and the initial weight of the connection. The adaption takes place dynamically, the changes are asymptotic, hence weights equal to zero are not possible, as well as weights equal to the specified upper bound. The following differential equation shows the dynamic extinction:

$$\frac{\delta}{\delta t} W_n(t) = -\alpha X_n(t) W_n(t) \quad \text{with } \alpha > 0 \text{ and } 0 \leq t \leq FP \tag{4}$$

with the following closed solution:

$$W_n(t) = W_n(0) e^{-\alpha \int_0^t X_n(\tau)\delta\tau} \tag{5}$$

where α is a learning rate parameter for the extinction process. The actual weight change is proportional to the initial weight $W_n(0)$. Extinction has a stronger effect on strong connections and only mildly affects weak connections. The weight change also depends on the cumulative activation of the respective state, that is, the whole activation that was propagated through this node in the time between WS and target stimulus (Fig. 5C). As shown in Fig. 5D, this results in a depression of all weights of nodes that are associated to time points before the appearance of the target stimulus.

Reinforcement is restricted to a fixed interval of duration d following the onset of the target stimulus at $t = FP$ and can be described through the differential equation:

$$\frac{\delta}{\delta t} W_n(t) = \beta X_n(FP)[1 - W_n(t)] \quad \text{with } \beta > 0 \text{ and } FP \leq t \leq FP + d \tag{6}$$

with the closed solution:

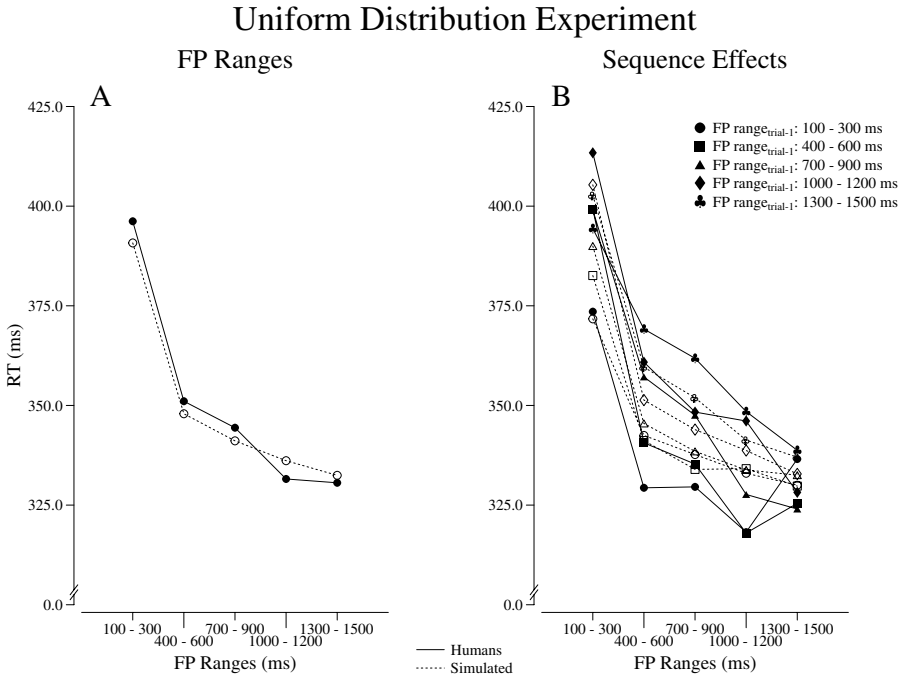

Fig. 7. The charts show simulated and empirical RTs depending on FP ranges (A) and depending on different FP ranges of the preceding and current trial (B) of the uniform distribution experiment

$$W_n(t) = K - (K - W_n(FP))e^{-\beta d X_n(FP)} \tag{7}$$

where β is a reinforcement learning rate parameter, t is the time that elapsed between WS and target stimulus, and K is the upper bound for every single weight. During reinforcement the weights at the beginning of the reinforcement period $W_n(t)$ are strengthened depending on the initial weight, the activation of the node at the time of reinforcement $X_n(t)$, and the reinforcement duration d. Figure 5E shows the activation of the nodes at the FP (i.e. at the appearance of the target stimulus) and Fig. 5F shows the resulting weights.

Response Rule. After the description of the time sensitive structure and the learning principles, we now turn to the response rule, which translates activations and weights into RTs. In our adaptation of the model[3], RT(t) is the RT that would result in a given trial if the target stimulus is displayed at time t:

[3] Note, that the response rule in [1] includes an additive term in the divisor, which was introduced to study tonic and phasic activation levels. As we do not deal with this topic here and to reduce the degrees of freedom of the model, we removed this term from the response rule.

Table 2. Uniform distribution experiment: R^2 of the prediction of the RT, averaged over sequences of FP ranges (sequential means, 25 predicted mean RTs per participant) and FP ranges (FP means, 5 predicted mean RTs per participant), for each particiant

participant	$R^2_{\text{sequential means}}$	$R^2_{\text{FP means}}$
9	0.90	0.98
3	0.87	0.98
4	0.87	0.98
6	0.82	0.96
7	0.69	0.91
5	0.68	0.81
2	0.63	0.85
8	0.58	0.99
10	0.58	0.93
1	0.57	0.99
M	0.72	0.94

$$RT(t) = RT_0 + \frac{A}{\sum_{n=1}^{N} X_n(t) W_n(t)} \tag{8}$$

where $X_n(t)$ is the activation of node n at time t, $W_n(t)$ the corresponding weight, RT_0 is an intercept, and A is a scaling coefficient. RT_0 represents the time taken by other processes contributing to the RT, like sensory stimulus processing or motor signal transmission. A is a necessary scaling factor, because the temporal regulation given by the sum in the divisor is bound between zero and K.

4.2 Simulation Method

To test if this adaptation of Los and colleagues' model [1] accounts for the behavioral data we estimated the model parameters and analyzed the overall fit to the empirical RTs. Following [1] and [24] we set the number of nodes to $N = 60$ and the reinforcement interval to $d = 200ms$. The remaining five parameters (RT_0, λ, A, α, and β) were fitted with the downhill simplex algorithm [30]. As each experimental participant received a different order of FPs and the model is sensitive to the order of FPs, we optimized the parameters of the model to predict each individual trial's RT as closely as possible (1-norm). We estimated individual parameter values for every participant. The resulting simulated RTs were aggregated similar to the empirical data to enable a direct comparison.

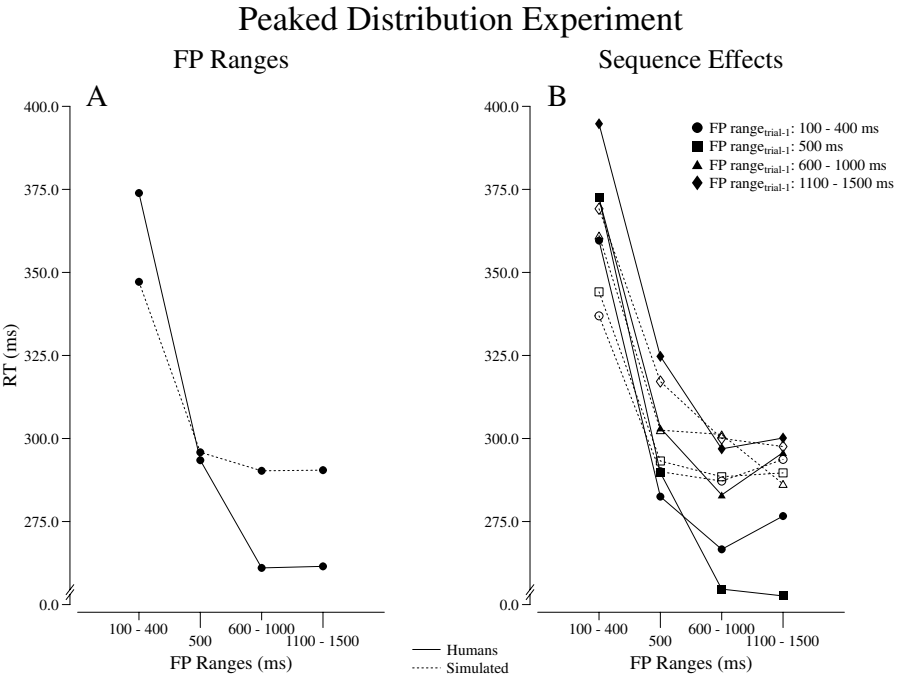

Fig. 8. The charts show simulated and empirical RTs depending on FP ranges (A) and depending on different FP ranges of the preceding and current trial (B) of the uniform distribution experiment

4.3 Simulation Results

Model Fit to Uniform Distribution Experiment. Figures 7A displays the results of our simulation of the uniform distribution experiment. The mean RTs for the different FP ranges simulated by the model correspond very well to the empirical data. Figure 7B shows the sequential effects of FP ranges for simulated and empirical RTs. Most of the qualitative features of the empirical data were reproduced by the model. However, the order of the impacts of previous trials on the short FP range 100-300ms could not be reproduced. Additionally, the impact of previous trial's FP seems to be somewhat reduced in the simulated data. Table 2 displays the amount of variance the simulated RTs can account for (indicated by R^2)[4] when considering RTs dependent on the FP and RTs dependent on the sequence of FPs (FP in trial n and trial n-1). Given the highly noisy individual data, the model acceptably reproduces the empirical RTs.

Model Fit to Peaked Distribution Experiment. Figure 8A displays the results of the simulation for the peaked distribution experiment. The empirical

[4] We used the coefficient of determination to estimate the goodness of fit of the model. $R^2 = \frac{SS_{\text{regression}}}{SS_{\text{total}}}$ is a measure for the amount of variance explained by the model.

Table 3. Peaked distribution experiment: R^2 of the prediction of the RT, averaged over sequences of FP ranges (sequential means, 25 predicted mean RTs per participant) and FP ranges (FP means, 5 predicted mean RTs per participant), for each participant

participant	$R^2_{\text{sequential means}}$	$R^2_{\text{FP means}}$
5	0.87	0.99
4	0.79	0.97
3	0.70	0.95
7	0.70	0.90
8	0.67	0.76
9	0.64	0.98
10	0.40	0.94
1	0.34	0.99
6	0.28	0.89
2	0.26	0.99
M	0.56	0.94

and simulated RTs corresponded for the different FP ranges. However, the between FP variability is smaller in the simulated data. This might be due to the high number of data points contributing to the 500ms FP. Figure 8B shows the empirical and simulated RTs as a function of FP range and FP range in the preceding trial. Again, the simulated RTs exhibit less variability than the empirical data. There are also some qualitative aspects of the empirical data that were not reproduced. Especially, the highly reduced RTs for trials following a trial with a 500ms FP could not be reproduced. We assume that this is caused by an effect which is systematically produced by the experimental design but not reflected by the model. The slower RTs for trials following a trial with a FP different from 500ms might be caused by cognitive processes, which result from the rather unexpected foreperiod in the previous trial. Probably, the model were suitable to reproduce the data if we put more weight on rare FP ranges for the fitting algorithm.

Table 3 displays the amount of variance the simulated RTs can account for (indicated by R^2) on different levels of aggregation. Similar to the results in the uniform distribution experiment, the accounted variance on the level of individual RTs is acceptable. However, the variance of the R^2 for sequential effects, which ranges between .26 to .87, is rather high.

5 Discussion

The purpose of the present study was to examine the adaptation of behavior in time to different distributions of many possible FPs. Our experimental results

show that humans are able to adjust their behavior to different predictability conditions. A comparison of both experiments reveals that RTs are much faster in the peaked distribution experiment than in the uniform distribution experiment. This implies that humans preactivate their behavioral system according to the predictability of a stimulus and that they are then able to quickly process that stimulus. Moreover, the described computational model accounts for human behavior in time under the conditions of our experiments.

5.1 Experimental Results

In detail, the conducted experiments replicated and extended typical findings. In general the RTs are the shorter the longer the FPs are. The typical sequential effects were also reproduced. RT increase with the length of the preceding FP relative to the current FP. These findings were in line with the common results reported for unpredictable FPs [7,2,8,9]. An interesting aspect of the sequential data is that participants did not respond fastest to direct repetitions of FPs. This was also true in the peaked distribution experiment where one FP was much more frequent than any other FP. Currently, we can only speculate how to interpret these findings. The shape of the different FP–RT functions might be the result of higher order sequential effects, also the FP distributions may be involved as well as the participants' ability to distinguish the FPs.

5.2 Computational Model

The computational model developed by Los and Agter [18] using the formal outlines of Machado [24] proved its ability to account for most of our experimental results. Please note that we fitted the model based on individual RTs of single trials and thus the fitting algorithm had to cope with very noisy data. We would expect even better fits if we ran several participants through identical sequences of FPs to average out RT variance that is caused by other than the preparatory mechanisms we want to study. Interestingly, the shifted minimum of the FP–RT functions was clearly reproduced. This was caused both by the model structure and the underlying learning mechanisms. The qualitative fit of the model was very good, even if the quantitative features of the empirical data could not be fully reproduced. Hence, in future research it might be beneficial to adjust the response rule to allow for a tighter replication of the data. In sum, both adjustments may improve fitting in future studies.

The supposed poisson process in connection with the applied learning rules seems to be able to account for a lot of qualitative aspects of the human ability of temporal anticipation. The quantitative fit may be improved by reformulating or extending the applied learning rules, as well as the response rule. For instance it seems to be quite simple to derive an expectancy of the length of the next FP after a single trial. The adjustment of the different weights cause the structure to be more or less "prepared" to react at different time points. The time point with the greatest product of activation and association strength could be conceived as a temporal expectancy or anticipation. The match or mismatch of

this "anticipation" with the subsequent FP could be used to predict erroneous behavior like premature responses or misses.

Please note, the presented model has five free parameters, which were used to predict five RT averages with high and 25 RT averages with acceptable accuracy. Due to the amount of free parameters, one may assume that the model can be fitted to a broad range of data. Indeed, the purpose of this paper is not to provide the most efficient model that accounts for the results but to present a model that is based on neurophysiological and psychological considerations. Given the neurologically derived architecture and the biologically plausible learning algorithms, the model may yield more explanational value than sparser, purely descriptive models.

5.3 Outlook

Another remarkable feature of the model is the possibility to adapt it to many experimental settings, which could differ from the FP paradigm. Machado proved the validity of the model in nearly all designs used to investigate the effects of temporal manipulations on the behavior of animals [24]. In all cases the structure of the model remained unchanged, only the response rule was adapted. Thus, the model reflects basic properties of the processing of temporal information in a wide range of species and behaviors [31].

References

1. Los, S., Knol, D., Boers, R.: The foreperiod effect revisited: conditioning as a basis for nonspecific preparation. Acta Psychologica 106, 121–145 (2001)
2. Bertelson, P., Tisseyre, F.: The time-course of preparation with regular and irregular foreperiods. Quarterly Journal of Experimental Psychology 20, 297–300 (1968)
3. Woodrow, H.: The measurement of attention. Psychological Monographs 5(76), 1–158 (1914)
4. Niemi, P., Näätänen, R.: Foreperiod and simple reaction time. Psychological Bulletin 89, 133–162 (1981)
5. Allan, L.G., Gibbon, J.: Human bisection at the geometric mean. Learning and Motivation 22, 39–58 (1991)
6. Wearden, J.H., Lejeune, H.: Scalar properties in human timing: conformity and violations. Quarterly Journal of Experimental Psychology 4, 569–587 (2008)
7. Elithorn, A., Lawrence, C.: Central inhibition - some refractory observations. Quarterly Journal of Experimental Psychology 11, 211–220 (1955)
8. Mattes, S., Ulrich, R.: Response force is sensitive to the temporal uncertainty of response stimuli. Perception and Psychophysics 59, 1089–1097 (1997)
9. Näätänen, R.: The diminishing time-uncertainty with the lapse of time after the warning-signal in reaction-time experiments with varying fore-periods. Acta Psycologica 34, 399–419 (1970)
10. Alegria, J.: Sequential effects of foreperiod duration: Some strategical factors in tasks involving time uncertainty. In: Alegria, J. (ed.) Attention and Performance, pp. 1–10. Academic Press, London (1975)

11. Karlin, L.: Reaction time as a function of foreperiod duration and variability. Journal of Experimental Psychology 16, 185–191 (1959)
12. Thomas, E.A.: Reaction-time studies: The anticipation and interaction of response. British Journal of Mathematical and Statistical Psychology 20, 1–29 (1967)
13. Drazin, B.: Effects of foreperiod, foreperiod variability, and probability of stimulus occurence on simple reaction time. Journal of Experimental Psychology 62, 43–50 (1961)
14. Bausenhart, K.M., Rolke, B., Hackley, S.A., Ulrich, R.: The locus of temporal preparation effects: Evidence from the psychological refractory period paradigm. Psychonomic Bulletin and Review 13, 536–542 (2006)
15. Fischer, R., Schubert, T., Liepelt, R.: Accessory stimuli modulate effects of nonconscious priming. Perception and Psychophysics 69(1), 9–22 (2007)
16. Hackley, S.A., Valle-Inclán, F.: Which stages of processing are speeded by a warning signal? Biological Psychology 64, 27–45 (2003)
17. Kiesel, A., Miller, J.: Impact of contingency manipulations on accessory stimulus effects. Perception and Psychophysics 69, 1117–1125 (2007)
18. Los, S.A., Agter, F.: Reweighting sequential effects across different distributions of foreperiods: Segregating elementary contributions to nonspecific preparation. Perception & Psychophysics 67(7), 1161–1170 (2005)
19. Miller, J., Franz, V., Ulrich, R.: Effects of auditory stimulus intensity on response force in simple, go/no-go, and choice rt tasks. Perception and Psychophysics 61, 107–119 (1999)
20. Müller-Gethmann, H., Ulrich, R., Rinkenauer, G.: Locus of the effect of temporal preparation: Evidence from the lateral readiness potential. Psychophysiology 40, 597–611 (2003)
21. Gibbon, J.: Scalar expectancy theory and weber's law in animal timing. Psychological Review 84, 279–325 (1977)
22. Killeen, P.R., Fetterman, J.G.: A behavioral theory of timing. Psychological Review 95(2), 274–295 (1988)
23. Treisman, M.: Temporal discrimination and the indifference interval: Implications for a model of the ïnternal clock: Psychological Monographs 77, 1–31 (1963)
24. Machado, A.: Learning the temporal dynamics of behavior. Psychological Review 104(2), 241–265 (1997)
25. Balkenius, C., Morén, J.: Dynamics of a classical conditioning model. Autonomous Robots 7, 41–56 (1999)
26. Grossberg, S., Schmajuk, N.: Neural dynamics of adaptive timing and temporal discrimination during associative learning. Neural Networks 2, 79–102 (1989)
27. Staddon, J.E.R., Higa, J.J.: Multiple time scales in simple habituation. Psychological Review 103(4), 720–733 (1996)
28. Los, S., Van den Heuvel, C.E.: Intentional and unintentional contributions of nonspecific preparation during reaction time foreperiods. Journal of Experimental Psychology: Human Perception and Performance 27, 370–386 (2001)
29. Schneider, W., Eschman, A., Zuccolotto, A.: E-prime user's guide. Psychology Software Tools Inc., Pittsburgh (2002)
30. Nelder, J.A., Mead, R.: A simplex method for function minimization. Computer Journal 7, 308–313 (1965)
31. Nobre, A.C., Correa, A., Coull, J.T.: The hazards of time. Current Opinion in Neurobiology 17, 465–470 (2007)