Neural repetition suppression modulates time perception: evidence from electrophysiology and pupillometry

Wouter Kruijne
Rijksuniversiteit Groningen

Christian N. L. Olivers
Vrije Universiteit Amsterdam

Hedderik van Rijn
Rijksuniversiteit Groningen

Abstract

Human time perception is malleable and subject to many biases. For example, it has repeatedly been shown that stimuli that are physically intense or that are unexpected seem to last longer. Two competing hypotheses have been proposed to account for such biases: one states that these temporal illusions are the result of increased levels of arousal which speeds up neural clock dynamics, whereas the alternative 'magnitude coding' account states that the magnitude of sensory responses causally modulates perceived durations. Common experimental paradigms used to study temporal biases can not dissociate between these accounts, as arousal and sensory magnitude covary and modulate each other. Here, we present two temporal discrimination experiments where two flashing stimuli demarcated the start and end of a to-be-timed interval. These stimuli could either be in the same or in a different location, which led to different sensory responses due to neural repetition suppression. Crucially, changes and repetitions were fully predictable, which allowed us explore effects of sensory response magnitude without impacting arousal or surprise. Intervals with changing markers were perceived as lasting longer than those with repeating markers. We measured EEG (Experiment 1) and pupil size (Experiment 2), and found that temporal perception was related to changes in event-related potentials (P2) and pupil constriction, both of which have been related to responses in sensory cortex. Conversely, correlates of surprise and arousal (P3 amplitude and pupil dilation) were unaffected by stimulus repetitions and changes. These results demonstrate, for the first time, that sensory magnitude affects time perception even under constant levels of arousal.

Introduction

In some of the most critical moments in life, making the right decision hinges on an accurate sense of time: at the start of a runner's sprint, or in the resolution of a cyclists race, the difference between victory and defeat hinges on choosing the right moment to break away; in our most emotional conversations, the slightest pauses between words can carry a world full of meaning; and during the final resolution of a poker game, the most subtle speedup and slowdown in on opponents' behavior can be taken as a tell about their odds of having the best hand. But while the clocks on our walls tick away at a steady rate, our 'internal clocks' that perceive such intervals are malleable and highly prone to biases, especially under conditions where we are emotional or have a heightened focus and sensitivity.

Identifying the neural mechanisms by which objective time is translated into a subjective experience of time lies at the heart of the study of time perception. To this end, psychophysical studies have attempted to identify the specific conditions where subjective time appears to speed up or to slow down. Converging evidence from such studies has given rise to the hypothesis that the neural 'clock dynamics' that give rise to a temporal percept are sped up under heightened states of arousal, which leads to overestimations of perceived durations. This explanation can account for a wide range of temporal illusions: for example, highly emotional stimuli are often perceived as lasting longer (Droit-Volet & Meck, 2007), time is dilated after a train of clicks or flashes (Penton-Voak, Edwards, Percival, & Wearden, 1996), and infrequent stimuli that violate participants' expectations are typically perceived as longer (Ulrich, Nitschke, & Rammsayer, 2006). The hypothesis is further supported by neurophysiological findings, which point to striatal dopamine release as a critical modulator of arousal-modulated speedup of clock dynamics (Paton & Buonomano, 2018; Soares, Atallah, & Paton, 2016).

However, certain experimental findings point to alternative sources of biases in time perception. For example, Johnston, Arnold, and Nishida (2006) showed that the duration of a laterally presented gabor stimulus was underestimated if it was preceded by a flickering stimulus at the same location. On trials where the flickering stimulus was presented in the opposite location of the screen, duration estimates were unaffected and comparable to trials without any flicker. This result points to a temporal bias constrained to the location of the flicker, and the authors suggested that the flicker resulted in retinotopically focused neural adaptation. This would have resulted in a subsequently weaker sensory response to the gabor, which might then underly underestimations of its duration. Not much later, Eagleman and Pariyadath (2009) proposed a similar explanation as a more general account of temporal biases. Reviewing a range of studies, they pointed out that overestimation of durations is generally found in experimental conditions where the to-be-timed stimulus evoked a larger neural response. For example, stimuli that are brighter, larger, or that are moving, appear to last longer when contrasted with stimuli that are dimmed, small, or stationary. The authors put forward the "sensory magnitude coding" hypothesis, which states that the magnitude of cortical responses plays a causal role in temporal perception. Neurophysiological support for this hypothesis was found by Mayo and Sommer (2013), who recorded individual neurons in macaque Frontal Eye Fields while the animals performed a temporal bisection task. Trials with incorrect 'longer'- and 'shorter'-

responses respectively showed enhanced and attenuated firing rates at the offset of the interval compared to correctly judged trials.

As there is considerable evidence for both arousal effects and sensory magnitude effects on time perception, it has proven difficult to disentangle both hypotheses. Fluctuations in arousal and differences in sensory responses strongly covary and can affect each other in several ways. Influential theories of perception posit that sensory responses are optimized to signal those stimuli that can be seen as arousing, either because they violate expectations (Clark, 2013; Friston, 2005; Horvitz, 2000; Rao & Ballard, 1999) or because they are associated with positive or negative reinforcement (Anderson, Laurent, & Yantis, 2011; Cho et al., 2017; Failing & Theeuwes, 2018; Keil & Ihssen, 2004). Conversely, heightened states of arousal can causally modulate the responses of sensory neurons (Buia & Tiesinga, 2006; Mather, Clewett, Sakaki, & Harley, 2016; Mather & Sutherland, 2011).

This complexity is particularly evident when considering the effects of stimulus repetition on perceived time, which is typically studied using an oddball paradigm: in a sequence of trials with frequent standard stimuli and infrequent oddball stimuli, the duration of the standards is underestimated whereas oddballs appear to last longer. Some have interpreted this as evidence for an arousal-mediated mechanism: compared to the standard, the more interesting oddball speeds up the clock dynamics, resulting in a longer perceived duration (New & Scholl, 2009; Ulrich et al., 2006). Others have argued that the oddball effect results from sensory magnitude coding (Kim & McAuley, 2013; Matthews, 2011b; Matthews & Meck, 2016; Pariyadath & Eagleman, 2008; Sadeghi, Pariyadath, Apte, Eagleman, & Cook, 2011), as predictive coding frameworks postulate that the repeating, predictable standard produces a weaker neural response than the oddball (Clark, 2013; Friston, 2005). Therefore, it remains unclear whether sensory magnitude affects perceived duration directly, or only when this is mediated by global levels of surprise or arousal (Ernst et al., 2017).

One shortcoming of the classical oddball paradigm is that stimulus repetitions are confounded with stimulus expectations (though see Kim & McAuley, 2013; Sadeghi et al., 2011). Previous neuroimaging studies have overcome this in designs that, for example, contrast fully expected stimulus changes with unanticipated stimulus repetitions. These studies demonstrate that 'repetition suppression' can be dissociated from 'expectation suppression' in both space, time and function (Kaliukhovich & Vogels, 2011; Larsson & Smith, 2012; Summerfield, Trittschuh, Monti, Mesulam, & Egner, 2008; Tang, Smout, Arabzadeh, & Mattingley, 2018; Todorovic & Lange, 2012). This points to a way to similarly dissociate the magnitude account from the arousal-based view: if stimulus repetitions and changes can compress and dilate perceived durations even when they are fully in line with an observers' expectations, this strongly suggests that repetition suppression causally affects temporal perception, without modulation by surprise or arousal.

Here, we present results from two experiments that demonstrate stimulus repetitions and changes can still bias temporal perception, even when they are fully anticipated. In Experiment 1, participants on each trial compared two intervals which were presented by means of flashing markers, one for the start and one for the end of the interval. These start- and end-markers could be in the same location or in opposite hemifields, and we unexpectedly found that a marker location change caused the interval to be perceived as longer than a location repetition. Even though the experiment was originally not designed to study repetition effects, the design permitted a clear dissociation of sensory magnitude accounts

and arousal-based accounts. We present analyses of behavior and event-related potentials (ERPs) that disentangle the two. In particular, we focus on three ERP components that are found at central electrodes and have previously been related to time perception, sensory processing, and expectations and arousal. First, we focus on the CNV, a slow-wave negative deflection that develops during an interval. In early work, the CNV was assumed to directly reflect internal clock computations (Macar & Vidal, 2003; Macar, Vidal, & Casini, 1999). More recently, a more nuanced interpretation has been put forward that the component reflects time-informed preparation for upcoming events and actions (Amit, Abeles, Carrasco, & Yuval-Greenberg, 2019; Boehm, van Maanen, Forstmann, & van Rijn, 2014; Kononowicz & Rijn, 2014; Van Rijn, Kononowicz, Meck, Ng, & Penney, 2011). The second component was the central P2, a positive peak around 200ms post-stimulus that is associated with pre-attentive sensory stimulus processing (Van der Molen et al., 2012). The P2 amplitude is modulated by physical stimulus properties, such as its brightness (Crowley & Colrain, 2004; Kaskey, Salzman, Klorman, & Pass, 1980), and previous studies have suggested a link between P2 amplitude at stimulus offset and perceived duration (Kononowicz & Rijn, 2014; Tarantino et al., 2010). Third, we assessed the P3 amplitude, an ERP complex that has been linked to a range of cognitive functions that include temporal decision making, (Lindbergh & Kieffaber, 2013), but also more broadly decision making and outcome evaluation (Jepma et al., 2016; Kelly & O'Connell, 2013; Nieuwenhuis, Aston-Jones, & Cohen, 2005). One canonical observation regarding P3 amplitude is that it relates to the degree of 'surprise' in response to a stimulus (Mars et al., 2008). Recently, P3 amplitude was found to predict biases in temporal perception in an oddball paradigm (Ernst et al., 2017).

Experiment 2 utilized a simpler, classical bisection paradigm in order to replicate the temporal bias found in Experiment 1. In this experiment, we recorded eye movements to strictly rule out their potential contribution to this bias, and we tested whether both horizontal and vertical changes yielded similar temporal dilation. Accompanying the behavioral results, we analyzed pupil size as another proxy for sensory magnitude and arousal. That is, previous work has shown that bottom-up stimulus strength can affect the pupil constriction response (Binda, Pereverzeva, & Murray, 2013; Mathôt, Dalmaijer, Grainger, & Stigchel, 2014; Naber, Frässle, Rutishauser, & Einhäuser, 2013), whereas dilation is often used as a signature of arousal or surprise (Akdoğan, Balcı, & Rijn, 2016; Braem, Coenen, Bombeke, van Bochove, & Notebaert, 2015; Satterthwaite et al., 2007).

To preview results from both experiments, the results from behavior, electroencephalograhpy (EEG) and pupillometry all could be readily accounted for by a magnitude coding account, and suggest that repetition suppression modulated perceived duration.

Experiment 1

In Experiment 1, participants compared two intervals spaced by a memory delay. This experiment was originally designed in order to investigate whether maintenance and retrieval of durations from working memory would yield lateralized signals in EEG if the interval was presented on either side of the display. These analyses are presented and discussed in a separate article (Kruijne, Olivers, & Van Rijn, in preparation), so as not to dilute the focus of the research question in the present article. Here, we will focus our analysis on one particular aspect of the task: stimulus locations either changed or repeated in a fully predictable manner, but nevertheless had a robust effect on perceived durations.

Method

Participants. Data were collected from 24 healthy participants who were compensated either by means of course credits or a monetary compensation of €10/hour. They were recruited through the research participants pool of the faculty of behavior and movement sciences at the Vrije Universiteit Amsterdam.Four participants were discarded on the basis of noisy EEG data, which was characterized by the preprocessing analyses described below: two participants had a too high number of noisy channels (8 and 10 EEG-electrodes out of 64) and two had a too high percentage of discarded data epochs (38% and 52%). None of the remaining participants were excluded on the basis of behavior. To initially assess behavior, we fit a psychometric curve for each participant by means of logistic regression that predicted the proportion of trials where the second interval was judged 'longer' than the first interval, as a function of the absolute difference in time between them. The slope coefficients of this regression were within 2SD from the group mean for each of the 20 participants. The final sample contained 20 participants (10 female, ages 19–26, Mean age 21.9).

Stimuli and procedure. Participants were seated in a sound-attenuated, dimly lit room at 75cm viewing distance from a monitor (22 inch, Samsung Syncmaster 2233) with 1680 × 1050 resolution and 120 Hz refresh rate. The experiment was programmed and presented using OpenSesame (Mathôt, Schreij, & Theeuwes, 2012), with the PsychoPy back-end (Peirce, 2007). Figure 1A depicts the stimulus sequence on an example trial with black and gray colors inverted for visibility. Trials started with a gray fixation cross (0.2°) presented on on a black background for 1000ms, followed by the onset of three, horizontally aligned gray placeholder circles (radius 1.97°, one in the center and two at 9.83°eccentricity) around a central gray fixation dot (0.2°). The placeholders and fixation dot stayed on screen until the end of the trial.

The first interval (interval 1) started 500ms after the onset of the placeholders, which was indicated by a red diamond (3.94°width and height) flashing for 125ms in the left or right circle. The duration of this interval was randomly sampled from a uniform distribution between 1250 and 2250ms. The end of the interval was again marked by a flashing red diamond either at the same or opposite location, dependent on block type ('same' or 'opposite' block). Then, a memory delay followed, which lasted 1250ms. After this delay, Interval 2 was presented by means of two green squares flashing in the center circle, one to mark the start and one to mark the end. The duration of this interval was determined relative to the duration of Interval 1, and was always either 10% or 20% shorter or longer. The offset of Interval 2 was followed by another 1250ms response delay, after which a response screen appeared where they were asked to indicate whether Interval 2 had been shorter or longer by pressing Z or M on a standard keyboard. On each trial, the meaning of these two keys could swap unpredictably, in order to prevent (lateralized) response preparation signals in the EEG. Participants were instructed to be accurate, and that there was no need for speeded responses.

EEG acquisition and data cleaning. EEG data were recorded at 512Hz using a BioSemi ActiveTwo system (biosemi.com, Amsterdam, The Netherlands) with 64 channels placed according to the standard 10–20 system. Six additional external electrodes were placed: two were placed on the earlobes to be used as offline reference; two were placed 1cm lateral to the external canthi, and two electrodes 2cm above and below the right eye. From

these latter four electrodes, horizontal and vertical electro-oculogram traces (H/VEOG) were constructed by subtracting data from the opposing pairs of electrodes.

Data were preprocessed and analyzed offline using MNE-python (Gramfort et al., 2013; Gramfort et al., 2014) and statistical functions in R. Supplemental Information with additional details regarding preprocessing are available at https://osf.io/7gpka/. Data were first high-pass filtered at 0.1Hz, removing slow drifts. In order to analyze data epochs with minimal artefacts, we used various algorithms to mark contaminated data for subsequent removal, which are described below. Many of these methods take epochs of equal length as their input, assumed to reflect on-task data of individual trials. To this end we created 'preprocessing epochs' with data from -2500ms to 2250ms around the onset of Interval 2. That way, they included a maximum amount of data on-task for each trial, based on the minimum duration of Interval 1, the duration of the memory delay, the minimum duration of Interval 2, and the response delay.

In order to detect channels with noisy data throughout the recording session, we followed the PREP-pipeline (Bigdely-Shamlo, Mullen, Kothe, Su, & Robbins, 2015) and made use of the RANSAC algorithm implemented in the autoreject package (Jas, Engemann, Bekhti, Raimondo, & Gramfort, 2017). This algorithm generates permutations of the dataset where it tries to predict channel data by interpolating the activity from a subset of channels (25%). If the correlation between the predicted and the observed data is too small ($r < 0.75$) in over 40% of epochs, that channel is marked as faulty. These channels were primarily located at peripheral sites and did not overlap with our electrodes of interest (see below). An overview of marked channels is given in the supplemental information. These channels were discarded from further preprocessing steps, and are interpolated from surrounding electrodes at the end of data cleaning. Note that this procedure had minimal impact on ERPs, as the analyzed amplitudes were computed from multiple averaged electrodes, and because discarded electrodes had virtually no overlap with our regions of interest.

Data segments contaminated by muscle artefacts were identified by means of a procedure adopted from the PREP-pipeline (Bigdely-Shamlo et al., 2015) and functions in FieldTrip (Oostenveld, Fries, Maris, & Schoffelen, 2011). Raw data were band-pass filtered at 110-140Hz and the hilbert envelope of the resulting signal was computed. The result was convolved with a 200ms boxcar averaging window, resulting in a per-channel time course estimate of high-frequency power characteristic for muscle activity. We determined a per-channel robust Z-score of such power based on the median and median absolute deviation computed from data within the preprocessing epochs. Data segments where the cross-channel average Z-score was larger than 5.0 were marked as contaminated and not considered in further analyses.

Artefacts caused by eye blinks were detected and filtered out by means of Independent Component Analysis (ICA), which was computed over the full dataset. Since the spatial ICA filters are sensitive to slow drifts, the raw data were first high-pass filtered at 1Hz, and data were subsampled to 102.4Hz ($=\frac{512}{5}$) to speed up computation. Independent components were computed using the infomax algorithm (the default in EEGLAB; Delorme & Makeig, 2004). Components were first visually identified as corresponding to blinks, an approach that was subsequently validated by correlating component activity with the VEOG signal during blinks. That is, we filtered the VEOG signal with a 1–10Hz bandpass filter and

extracted 1000ms epochs around local maxima. Components that were identified as blink-related, only those components, had a high correlation with VEOG amplitude ($r^2 > 0.5$) .

Horizontal saccades were identified by high-pass filtering the HEOG trace at 1Hz, and convolving the result with a step-wise zero-mean filter kernel of 325ms that went from -1 to 1 by means of a 25ms linear ramp (an approach adopted from ERPlab; Lopez-Calderon & Luck, 2014). In the resulting signal, we identified local maxima that exceeded the 99th percentile as moments of horizontal saccades. Data from the moment of such saccades until the end of a trial were marked as contaminated, and later epochs overlapping with annotated data segments were omitted from analyses.

As a final data cleaning step we applied the autoreject algorithm (Jas et al., 2017) in order to identify unreasonable fluctuations in amplitudes. Autoreject aims to improve on typical epoch rejection based on peak-to-peak amplitudes with a fixed threshold. It uses a bayesian optimization procedure to compute per-channel thresholds, as well as an additional integer value k. If a channel's peak-to-peak value in an epoch exceeds its threshold, autoreject will initially attempt to interpolate its data from neighboring channels. If the number of marked channels in an epoch is larger than k, the entire epoch is discarded. An overview of the fitted threshold values are given in the supplemental information, together with the number of epochs included after data cleaning.

ERP components: CNV, P2 and P3. Using the algorithms described in the previous section, 'cleaned' data epochs were created for ERP analyses. As stated in the introduction, we focused our analyses on data from electrodes at central sites (Cz, C1, C2, FCz, FC1, FC2, CPz, CP1, CP2) where we identified the CNV, the P2 and P3 (see Figure 2A and Figure 3A). In particular, data from these electrodes were spatially averaged and components were defined as the amplitudes in three different time windows in the resulting signal. First, we defined CNV amplitude on each trial as the average amplitude during the 250ms leading up to the end of the interval with respect to a pre-stimulus baseline (For Interval 1, this baseline was 100ms before the onset of the placeholders; for Interval 2 this baseline was 100ms before Interval 2 onset). For the P2, we averaged amplitudes from 200 to 250ms after the offset of each interval. In order to ensure that this measure captured the transient response with minimal interference from the ongoing slow-wave CNV amplitude, the P2 was computed with respect to a 100ms baseline surrounding the moment of interval offset (t=-50ms to +50ms, following Correa & Nobre, 2008; Kononowicz & Rijn, 2014). The final ERP component of immediate interest, The P3, was defined as the average amplitude between 350 and 400ms with respect to the same baseline as for the P2.

Statistical analyses. We analyzed behavioral, and electrophysiological data from regions- and time-windows of interest by means of generalized linear mixed models (GLMM) (Baayen, Davidson, & Bates, 2008). These analyses aimed to characterize the experimental variables that could best account for the behavioral change bias, capture variability in the EEG-data, and identify potential relations between the two. There are many approaches to statistical inference with GLMMs: for example, one could compare predefined models with prior plausibility on the basis of their fit values, or one might fit a single theory-driven model and determine a p-value for each of its predictors. As the change bias was an unanticipated finding, we chose a more data-driven model comparison approach. That is, we predefined a set of predictors of interest, and compared the statistical models that could be constructed

from these predictors. We base inferences on the BIC, which is typically a more conservative statistic than the AIC or χ^2. We report results regarding the 'best' model and its contrast with 'suboptimal' models with higher BIC values. In order to additionally report p-values we will report the outcomes of likelihood ratio tests where appropriate.

More precisely, we analyzed behavioral data by means of logistic regression on the response on each trial, and compared models constructed from the following predictors: the duration of Interval 1, the duration of Interval 2, the difference between them (Δt), the percentage change on a trial, and the trial type (same/change in marker location for Interval 1). The most complex model, with all predictors and their interactions was constructed, and as an initial model selection step was trimmed down on the basis of the AIC, as implemented by the `lmerTest::step()` function. Of this resulting model, we fit all less complex 'child' models that were nested variants of it, with the only constraint that interaction terms were only included alongside their respective main effects. All models included participant ID as a random intercept term. For each of the fixed effects in the 'best' model, we tested whether including a corresponding random slopes term should be included in the model, and did so if this improved a models' AIC score. Each of the ERP components (CNV, P2 and P3 for both Interval 1 and Interval 2) was analyzed using the same model selection procedure, albeit with linear mixed models, and with 'temporal percept' as an additional predictor (Interval 1 or 2 judged as longer, i.e. the response). This way, we could establish a direct link between ERP-components and behavior. For all analyses, we report the preferred model alongside comparison statistics for suboptimal simpler models where the predictor terms are individually removed We additionally will report comparison statistics with other models of theoretical interest. Our conclusions are based on the ΔBIC, where we use positive values to indicate that including a term is supported, and negative values for evidence against its inclusion. Wherever possible, we additionally report p-values based on the χ^2 statistic from likelihood ratio tests.

Note that most factorial analyses, such as traditional ANOVAs or t-tests, would be inadequate for our data, due to the multicollinearity among the predictors of interest. For example, a simple contrast among ERPs for trials where Interval 1 or Interval 2 was perceived as longer would be inappropriate: the latter of these classes would always include more trials where Interval 2 was objectively longer, relatively longer, or where the marker location changed. This similarly impedes most explorations of spatiotemporal differences using cluster-based permutation tests. Comparisons based on the BIC suffer less from this problem as they allow us to assess model complexity as a whole, rather than the estimated contrasts within the model. Nevertheless this issue of collinearity is to be heeded when interpreting our visualizations and precludes strong conclusions regarding the interpretation of model coefficients.

Results

Behavior. Behavior is depicted in Figure 1B, together with the predictions of the model with the best BIC value. This model had three additive fixed effects terms: First, the 'percentage change', expressed as a continuous linear term, was a strong predictor of behavior ($\Delta BIC = 15.0; \chi^2(1) = 23.8; p = 0.001$), which describes participants' ability to correctly perform the task. A random slope was included in the model to capture variability in this term across subjects. Second, participants were more likely to perceive Interval 2 as

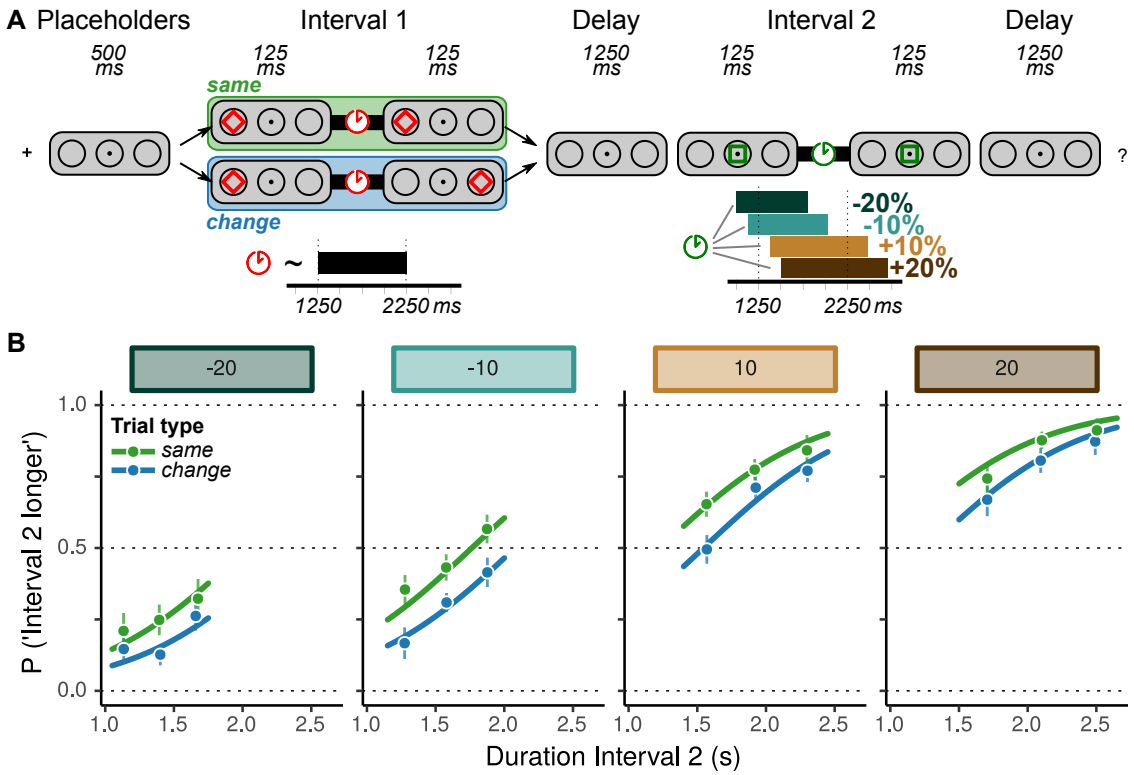

Figure 1. **A** Trial sequence. Participants compared Interval 1, which had a randomly sampled duration, to Interval 2, which was 10 or 20% shorter or longer. Intervals were indicated by flashing markers. Markers for Interval 1 appeared either in the same position (green shading) or changed in position (blue shading), which was varied across blocks. **B** Behavior. Participants' responses were driven by the relative duration of the intervals, but also by the absolute duration of Interval 2 and by the way Interval 1 was presented. Data points reflect average response rates in per-condition bins of Interval 2 duration, with 95% within subject Confidence interval (Cousineau, 2005; Morey, 2008). Curves reflect the fit of the best GLMM.

longer when it was also objectively long ("Duration Interval 2", expressed as a continuous linear predictor: $\Delta BIC = 296.0; \chi^2(1) = 298.8; p < 0.001$). Note that ideally, participants' responses would be driven by the relative duration alone, the objective duration is a useful heuristic for determining the correct response which participants might have exploited. Finally, we found that the trial type ("same" or "change" trial, referring to the markers of Interval 1), affected participants' responses: Interval 2 was less likely to be perceived as longer than Interval 1 on trials with a marker change ($\Delta BIC = 82.5; \chi^2(1) = 91.2; p < 0.001$). This effect will hereafter be referred to as the 'change bias'.

The predictions of this model are depicted in Figure 1B. In order to visualize the fit of binomial response data as a function of a continuous predictor ('Interval 2 duration'), we split data for each participant across three bins for each condition and computed the average response rate per bin. These averages were used to compute grand averages with

95% confidence intervals of within-subjects effects (Cousineau, 2005; Morey, 2008). Note that this was for visualization purposes only, and bins were not used in the analyses.

With respect to our research question it is of particular interest whether the change bias might be accounted for by a sense of 'surprise' on change trials. We do not expect this to be the case, as trial types were varied across rather than within blocks, and participants therefore should have been accustomed to the changing markers on each trial. We assessed this in more detail by exploring whether the bias effect was perhaps larger at the start of a block compared to the end. A factorial predictor term 'block phase' was defined through a median split on trial index, and was included in a more complex model as a predictor that interacted with 'trial type'. This model was not supported, however ($\Delta BIC = -17.5, \chi^2(2) = 0.06, p = 0.968$).

Conversely, one might hypothesize that the block-based manipulation of trial-type caused participants to be more aroused or more attentive in 'change' blocks, which in turn could have caused a speedup of clock dynamics under these conditions. If this were the case, one would expect the effect of trial type to be larger for trials with a relatively long Interval 1 duration. Such an interaction, however, had no statistical support ($\Delta BIC = -12.7; \chi^2(2) = 4.9; p = 0.085$). The χ^2 statistic here might marginally hint at a possible interaction, but we note that this was in fact under-additive, with a smaller bias for longer durations.

Taken together, the behavioral data support the interpretation that the change bias reflects a constant, additive bias that is driven by differences in low-level sensory processing on 'same' and 'change' trials, and does not reflect a speedup of clock computations or an effect driven by surprise or arousal. Next, we investigate whether ERP components lend further support to this view.

ERP Components during Interval 1. Figure 2A depicts the overall time course of the selected central electrodes, and highlights the CNV, the P2, and P3 as defined in the methods section. The models that best captured the data are listed in the table on the right. First, for the CNV, we found no evidence that any of the predictors modulated its amplitude: the second-best model included a term for which interval was perceived as longer (Interval 1 or Interval 2), but neither the BIC nor likelihood ratio tests gave an indication that this model was supported ($\Delta BIC = -7.5; \chi^2(1) = 0.29$).

For the P2, we found that it was modulated by the presentation type of the interval. That is, trials with a change in marker location elicited a larger P2 than trials with a repeated marker location ($\Delta BIC = 0.15; \chi^2(1) = 8.8; p = 0.003$). This effect is depicted in Figure 2B, accompanied by the time course of the ERP split for the two trial types in Figure 2C. To ensure that this modulation could not be partially attributed to potential baseline effects due to the CNV, we have also analyzed the P2 amplitude from band-pass filtered data (2 to 20Hz), which emphasizes the temporal properties of the P2 while attenuating the slower CNV component (following Kononowicz & Rijn, 2014) and filter out the slow-wave drifts of the CNV. Analyses on filtered data similarly supported the model with 'trial type' as an added predictor ($\Delta BIC = 1.05; \chi^2(1) = 9.7; p = 0.002$).

In a companion article (Kruijne, Olivers & Van Rijn, manuscript in preparation) we analyzed data from the same experiment and specifically aimed to uncover lateralized signals and their relation to time perception. There, we showed that cluster-based permutation analyses suggested that lateralized occipital ERPs were also modulated by trial type,

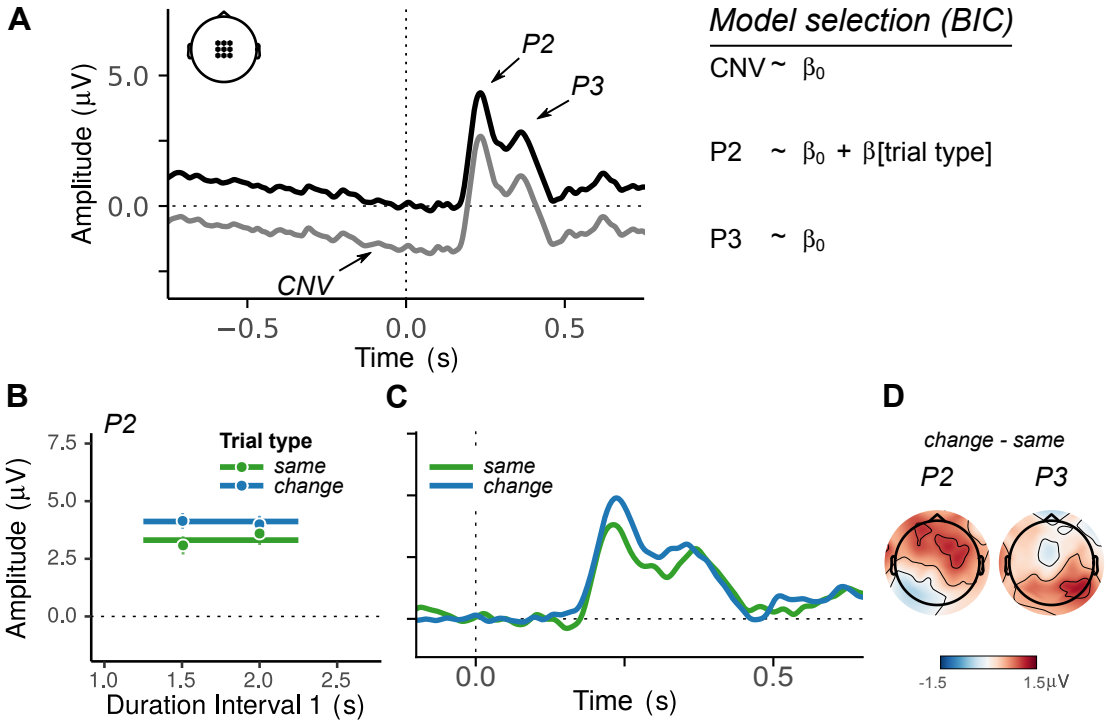

Figure 2. ERPs for Interval 1. **A** Grand-average waveforms at central sites. The gray curve has a baseline before the start of the interval to measure the CNV amplitude at the end of the interval. The black curve reflects the same potentials but with a baseline at the marker presentation, to measure the P2 and P3 amplitudes. **B** Model fit, **C** time course, and **D** spatial distribution of the P2 effect. Same- and change trials evoke P2 differences at central sites, but this difference has disappeared around the time of the P3. Around that time a contralateral positivity (plotted on the right) was present at occipital sites.

between 197ms and 455ms post-onset – which we interpreted as a difference in the N2pc between these conditions. When we reanalyzed the contralateral - ipsilateral signal focusing on the time associated with an N2pc, from 200-300ms, the best LMM was one where lateralized ERP amplitudes at occipital electrodes between 200 and 300ms were larger on 'change' trials ($\Delta BIC = 7.9; \chi^2(1) = 29.02; p < 0.001$). In Figure 2D, we plot the average scalp distribution of potentials at 200 to 250ms, and 300 to 350ms, contrasting 'change' and 'same' trials. These maps are derived from epochs where we mirrored data such that contralateral electrodes are plotted on the right. These maps illustrate how change trials are characterized by a positivity that peaks more frontally around the time of the P2, but that later in time manifests solely at contralateral, occipital electrodes.

The ERP time course and the scalp distributions depicted in Figure 2C and D suggest that the difference between different trial types is not modulating the central P3 component: around that time the difference is merely found at contralateral occipital electrodes. Indeed, analyses of the P3 found that it was not modulated by any of the predictors, with no support for an effect of trial type ($\Delta BIC = -8.3; \chi^2(1) = 0.29; p = 0.592$) nor for an effect of interval

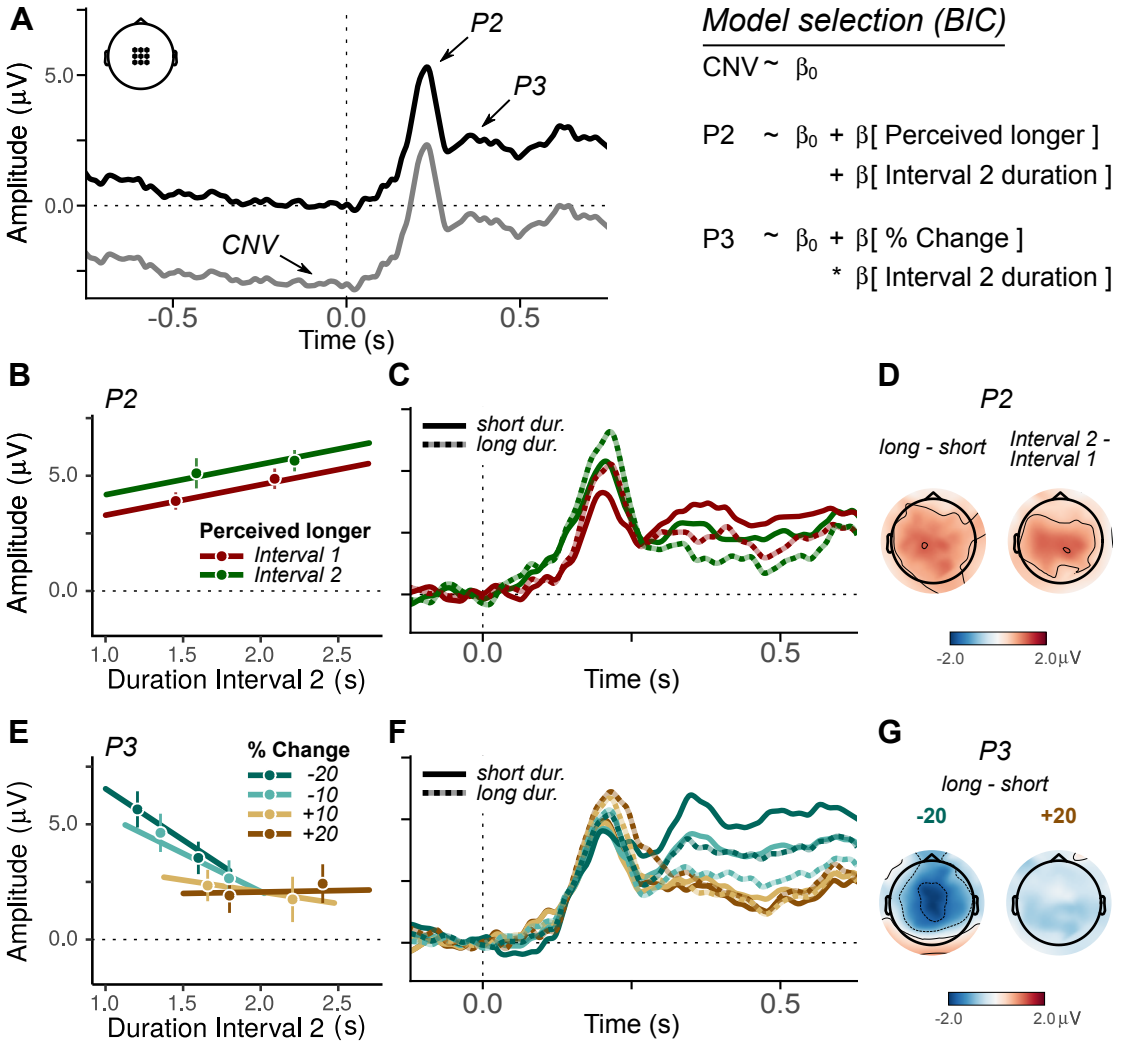

Figure 3. ERPs for Interval 2. **A** Grand-average waveforms at central sites, as in Figure 2, with on the right the best statistical models. P2 amplitudes were predicted by objective duration as well as by the temporal percept. The P3 amplitude was predicted by objective duration interacting with the relative (%) Change **B** Model fits, **C** time course, and **D** Spatial distribution of the P2 effects. Labels 'long' and 'short' in **C**&**D** correspond to bins in **B**. **E** Model fit, **F** time course, and **G** spatial distribution of the P3 effects. Again, 'short'/'long' correspond to the bins in **E**

duration ($\Delta BIC = -8.6; \chi^2(1) = 0.01; p = 0.907$).

ERP Components during Interval 2. In Figure 3A we illustrate the overall time course at central electrode sites around the offset of the second interval. While the CNV was more pronounced here compared to the offset of Interval 1, we still found no support that it was modulated by any of the predictors: the best model was the intercept-only model, followed by a model that included the comparison judgment as a predictor ($\Delta BIC = -7.7; \chi^2(1) = 0.93; p = 0.335$).

As for Interval 1, the P2 was again modulated by experimental variables. That is, the model that best described its amplitude included the predictors 'Interval 2 duration' ($\Delta BIC = 2.78; \chi^2(1) = 11.4; p < 0.001$) and 'perceived longer' ($\Delta BIC = 0.78; \chi^2(1) = 7.9; p = 0.005$), additive with evidence against an interaction ($\Delta BIC = -8.58; \chi^2(1) = 0.06; p = 0.798$). Again, we reran the analyses on band-passed filtered data to overcome potential carryover effects from the CNV, which resulted in the same statistical model with somewhat stronger evidence (both $\Delta BIC > 1.5; \chi^2(1) > 10.2; p = 0.0014$). The model fits, the time-course of the ERP, and the spatial distribution of these effects are depicted in Figure 3B-D.

The best model for the P3 was one with the predictors 'Percentage change' and 'Interval 2 duration', and their interaction (Percentage change: $\Delta BIC = 7.88; \chi^2(2) = 25.16; p < 0.001$, Interval 2 duration: $\Delta BIC = 8.1; \chi^2(2) = 25.39; p < 0.001$, interaction: $\Delta BIC = 7.2; \chi^2(1) = 15.8; p < 0.001$). As illustrated in Figure 3E-G, this interaction entailed that the contrast between 'short' and 'long' durations of Interval 2 yielded large P3 differences on trials with Interval 2 < Interval 1 (-20% or -10%), but did virtually not differ for trials where Interval 2 > Interval 1 (+10% or +20%).

P2 Amplitude predicts the temporal percept. The previous section shows that the P2 at the offset of Interval 2 was larger on trials where Interval 2 was perceived as longer. To further understand the link between this ERP component and perceived durations, we also assessed whether the inverse held. That is, we assessed whether the P2 was supported as another predictor for behavior, extending the behavioral model depicted in Figure 1B. Indeed, we found that the extended model including P2 amplitude as a continuous predictor improved the model ($\Delta BIC = 1.35; \chi^2(1) = 9.99; p = 0.002$). The best model still included the other three predictors (Removing either term: $\Delta BIC < -9.78$), and interactions between either of the predictors were not supported ($\Delta BIC < -6.87$).

The effect of P2 amplitude is depicted in Figure 4, where for each 'percentage change' condition the P2 amplitude is divided into three bins. The model prediction with the average P2 amplitude in each bin is plotted as in Figure 1B, but with three levels of transparency reflecting different P2 amplitude bins. The figure illustrates that for certain experimental conditions, participants are approximately 10% more likely to perceive an interval as longer in the highest P2 bin compared to the lowest bin. Note that this figure obscures the observation that generally, the P2 was also found to be larger for longer intervals (Figure 3B).

None of the other components under consideration, either for Interval 1 or for Interval 2, were statistically supported as a predictor for temporal judgments.

Discussion Experiment 1

Experiment 1 revealed that temporal perception was affected by the manner in which Interval 1 was presented: in blocks where the stimuli that marked the start- and end of the interval were in the same location, the time between these markers was perceived as shorter than when the markers were in opposite sides of the visual field. The characteristics of this 'change bias' fit with a magnitude coding account of temporal biases, and not with an arousal based speedup of clock dynamics. First, the change bias could be characterized as an additive effect that appeared to be invariable across the entire range of intervals: this fits with a sensory difference triggered by the change, and not with a speedup of clock dynamics.

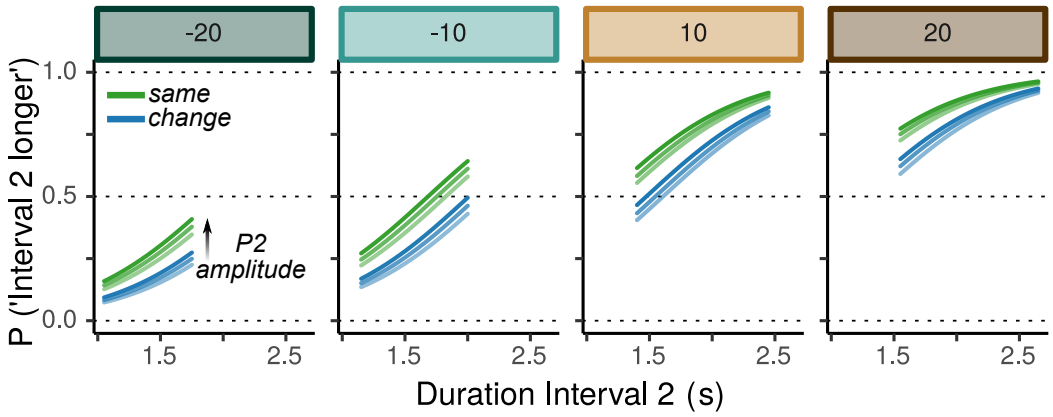

Figure 4. Model predictions for 'longer' responses, like the model fit in Figure 1B, but additionally including a predictor for the effect of P2 amplitude. In each condition, the predicted response rates are depicted for three different bins of P2 amplitude. Higher amplitudes are plotted with more opaque lines.

Second, the effect remained constant throughout blocks with the same presentation type, which renders an account based on 'surprise' highly unlikely: both trial types were fully predictable and only differed in terms of repetition of stimulus location. ERP analyses established a link between behavior and the P2 amplitude at the offset of both intervals. For Interval 1, trials with a marker change were associated with a larger P2. For Interval 2, the P2 amplitude predicted the upcoming comparison response. The CNV and the P3, respectively associated with temporal expectation and surprise, did not differ across trial types, nor did they relate directly to temporal judgments. Taken together, these findings indicate that sensory repetition suppression, reflected in P2 amplitude, affected time perception in the absence of heightened levels of arousal or surprise.

Although the results of Experiment 1 establish converging evidence that the change bias was caused by repetition suppression affecting magnitude coding of temporal durations, this study was originally not designed with the intention to investigate this bias, but rather to study neural markers of maintaining and retrieving durations in working memory. Our analyses could therefore be considered inherently exploratory and required subsequent confirmation. Experiment 2 therefore served to replicate and confirm the change bias phenomenon in a more canonical experimental setup, using a bisection paradigm with comparisons to a single fixed standard interval (Allan & Gibbon, 1991; Ulrich & Vorberg, 2009; Wearden, 1991). The bisection paradigm also discards the working memory component of Experiment 1, which allows us to test alternative explanations of the bias. For example, Experiment 1 leaves open the possibility that the change bias does not reflect a modulated temporal *percept*, but rather a modulated memory representation or facilitated retrieval of the perceived duration. Second, in Experiment 2 we tested whether the change bias could to a similar extent be found for horizontal and vertical changes in marker location. If the bias is found to be limited to horizontal changes, this might point to interhemispheric communication between localized clock computations as a potential source for the bias (e.g. Bueti, Bahrami, & Walsh, 2008; Christman, Garvey, Propper, & Phaneuf,

2003; Harrington, Haaland, & Knight, 1998; Paton & Buonomano, 2018; Vicario et al., 2008), rather than sensory magnitude coding. Third, given the effects that saccadic eye movements have on perceived durations (Burr, Ross, Binda, & Morrone, 2010), we relied on eye tracking for stringent control for eye movements, both for offline trial rejection as well as online feedback to participants. While we did not record EEG in Experiment 2, we explored the use of pupillary responses at interval offset as a potential measure to dissociate sensory magnitude effects from arousal. In particular, we assessed whether marker changes affected pupil sizes, either by causing larger pupil dilation, a signature of arousal or surprise (Akdoğan et al., 2016; Braem et al., 2015; Satterthwaite et al., 2007), or by evoking stronger constriction responses which are associated with stronger pre-attentive sensory processing (Binda et al., 2013; Mathôt et al., 2014; Naber et al., 2013).

Experiment 2

Method

Participants. Data were collected from 32 participants who were recruited through the first-year participant pool of the University of Groningen, and who did not participate in Experiment 1. The experimental procedure was approved by the Ethical committee of the Faculty of Behavioral and Social Sciences (PSY-1819-S-0019). All participants gave written informed consent before participating, and were treated in accordance with the Helsinki declaration. Data from five participants were omitted due to poor performance (as defined below), leaving 27 participants in the final sample (18 female, Mean age 20.6).

Stimuli and procedure. Overall trial design of Experiment 2 is depicted in figure Figure 5A. Participants sat in a dimly lit room in front of a 27́ LCD monitor (Iiyama prolite g2773hs-gb1) with a 60Hz refresh rate at 60cm distance from the screen, aided by a chin rest. The experiment started with an instruction phase, where participants familiarized themselves with the reference interval (1750ms) and practiced the task, after which the experimental phase followed.

On-screen stimulus dimensions were identical to Experiment 1, but participants sat 15cm closer to the screen. We report the new, slightly larger visual angles here. Each trial began with a small white fixation cross (0.25°) on a black background, that participants were asked to fixate and to then press the spacebar to initiate drift correction and subsequently start a trial. On experimental trials, drift correction was immediately followed by a cue informing participants on the trial type, by means of the word 'same' or 'opposite', presented for 400ms. In the instruction phase, a 400ms blank screen was shown instead of the cue. This was followed by a 500ms screen with only a fixation dot, after which five white placeholder circles appeared: one in the center, and four at 12.22° eccentricity on the left, right, top, and bottom of the display, each with a 2.45°radius. After another 500ms, the interval was presented as the SOA between the start- and end marker, each flashing for 125ms in one of the placeholders. On each trial, this interval was sampled from a uniform distribution between 1250 and 2250ms. In the experimental phase, markers were red diamonds (4.95°width and height) presented in the peripheral placeholders, either on the horizontal axes or the vertical axes as determined by the block type, and in the same or the opposite location as indicated by the cue. In the instruction phase, start- and end markers were green squares (3.47°) presented within the central placeholder. The placeholders re-

mained on screen for another 2000ms after the end of the interval, after which participants were probed for a response, to indicate whether the presented interval had been shorter ('F') or longer ('J') than the reference interval (1750ms). They were instructed to be as accurate as possible, without the need to respond fast.

Before the experimental blocks, participants completed an instruction- and practice phase: first, the reference interval was presented nine times to participants, with the instruction to to attend and memorize this duration. After this, they completed ten practice trials with randomly sampled intervals, where they received trial-wise feedback on their accuracy. The experimental phase consisted of eight blocks. Each block started with three 'refresher' presentations of the reference interval, after which participants were informed whether markers would appear on the horizontal or vertical axis in the upcoming block. As in Experiment 1, participants therefore could never predict the exact location of the onset marker, but could always infer the location of the offset marker before it was presented, based on the cue and the location of the onset marker.

Gaze and pupil data acquisition and analysis. Participants' right eye was tracked during the experiment using an EyeLink 1000 system (www.sr-research.com), which recorded gaze position in screen coordinates and pupil size (area in arbitrary units) at 1000Hz. Participants were instructed not to blink or to move their eyes during a trial, from the cue until the response. If the recorded gaze position was more than 2.45 °from the center and thus moved outside the central placeholder, participants saw a warning message in orange font at the end of the trial, and the trial was excluded from all further analyses (6.2% of trials).

The analysis of pupil size was conducted offline and followed (Mathôt, Fabius, Van Heusden, & Van der Stigchel, 2018). First, segments of pupil data recorded between 100 ms before and after a blink were replaced by linearly interpolated data. Two pupil epochs were subsequently created, one around interval onset (-1400 – 1250ms, so as to encompass the time from the start of trial up to the minimal interval duration) and one around interval offset (-500 – 2000ms, encompassing the end of the interval up to the moment of the response). Pupil sizes were baseline-corrected by subtracting the median pupil size in a 200ms window at the start of each trial, and samples were subsequently normalized by dividing pupil size samples by the average median baseline. This resulted in epochs of pupil size expressed as proportional change.

Statistical analyses. Overall performance of each participant was assessed by means of a logistic regression that predicted 'longer' responses as a function of interval duration, independent of condition. From the regression coefficients, a 50% Just-Noticeable Difference (JND) metric was computed. Participants were excluded if their JND was larger than 500ms which, following the traditional interpretation of the JND, implied that their response had been a guess on more than half of the trials.

Subsequently, behavior was analyzed by means of logistic general linear mixed effects regression, predicting responses based on the predictors 'interval duration' (numeric predictor), 'trial type' (same or change) and 'axis' (horizontal or vertical blocks), with a random intercept term and an uncorrelated random slope for interval duration. These two terms described participants individual differences in response biases and sensitivity to duration differences, respectively. Based on AIC scores, no other random effects were included. Models constructed from all combinations of predictors and their interactions

were again compared using the based on the BIC for statistical inference. For pupil sizes, effects of trial type were initially explored by averaging across trial type (same/change) and block type (horizontal/vertical), and exploring their average time courses on each trial by means of cluster-based permutation tests. The results, outlined below, suggested that pupil constriction at offset was modulated by trial type. In order to explore this in more detail we additionally computed a per-trial measure of constriction as the peak-to-peak difference in a window 100-700ms after interval offset, which we subjected to model comparisons with linear mixed effects models, similar to behavior.

Results

Data points in Figure 5B illustrates the average proportion of 'longer' responses in Experiment 2, computed separately for all levels of 'trial type' (same/change) and 'axis' (horizontal/vertical), for four bins of interval duration. The curves in this figure reflect the best statistical model based on the BIC. This model had two additive fixed effects: interval duration as a linear predictor ($\Delta BIC > 1000; \chi^2(2) = 3163.3; p < 0.001$), reflective of task performance, and trial type (same/change; $\Delta BIC = 163.2; \chi^2(1) = 172.23; p < 0.001$), which indicated we replicated the change bias effect from Experiment 1. Again, we found no evidence for an interaction between duration and trial type ($\Delta BIC = -6.0; \chi^2(1) = 2.98; p = 0.084$). As in Experiment 1, the potential trend hinted at by the χ^2-statistic actually pointed to an under-additive interaction, once again arguing against a clock-speed modulation in change trials. Figure 5B also illustrates that behavior was almost identical in horizontal or vertical blocks. Indeed, there was no support for the additive predictor 'axis' ($\Delta BIC = -8.9; \chi^2(1) = 0.053; p = 0.818$) nor for an interaction term that captured a modulation of the change bias by the axis ($\Delta BIC = -17.9; \chi^2(2) = 0.061; p = 0.970$). This illustrates that the change bias was not limited to horizontal displacements, rendering an explanation based on interhemispheric communication unlikely.

At the onset of the interval, we did not observe any differences in pupil size that could be related to the change bias. Exploring the data by means of cluster-based permutation analysis marked one significant cluster that was suggestive of a difference between same- and change trials, from -1150ms to -20ms with respect to interval onset ($p = 0.002$). In this window, change trials were found to have somewhat smaller pupil sizes. Given the timing of this cluster, this finding likely reflects differences in pupil constriction in response to the cue rather than any differences related to temporal perception. Note that change-trials were signaled by the word 'opposite', and same-trials by the word 'same' – the latter of which was smaller and therefore less bright than the former. The permutation test did not identify any significant differences during the interval itself. A LMM predicting the average pupil size in a window of interest at the end of the interval (750 to 1250ms) provided no support for a difference between trial types ($\Delta BIC = -19.5; \chi^2(1) = 0.12; p = 0.734$).

At the offset of the interval, the cluster-based permutation analysis indicated one significant cluster, for the main effect of trial type, between 429 and 1084ms after interval offset ($p = 0.044$). This was suggestive of more constriction on change trials, as is illustrated in Figure 5C. The traces in this figure also suggest another difference, in that constriction responses were stronger in horizontal blocks compared to vertical blocks. While this difference was indeed reflected in a cluster in the same time range (426 to 771ms), that cluster did not survive permutation correction ($p = 0.084$). We explored this constriction response

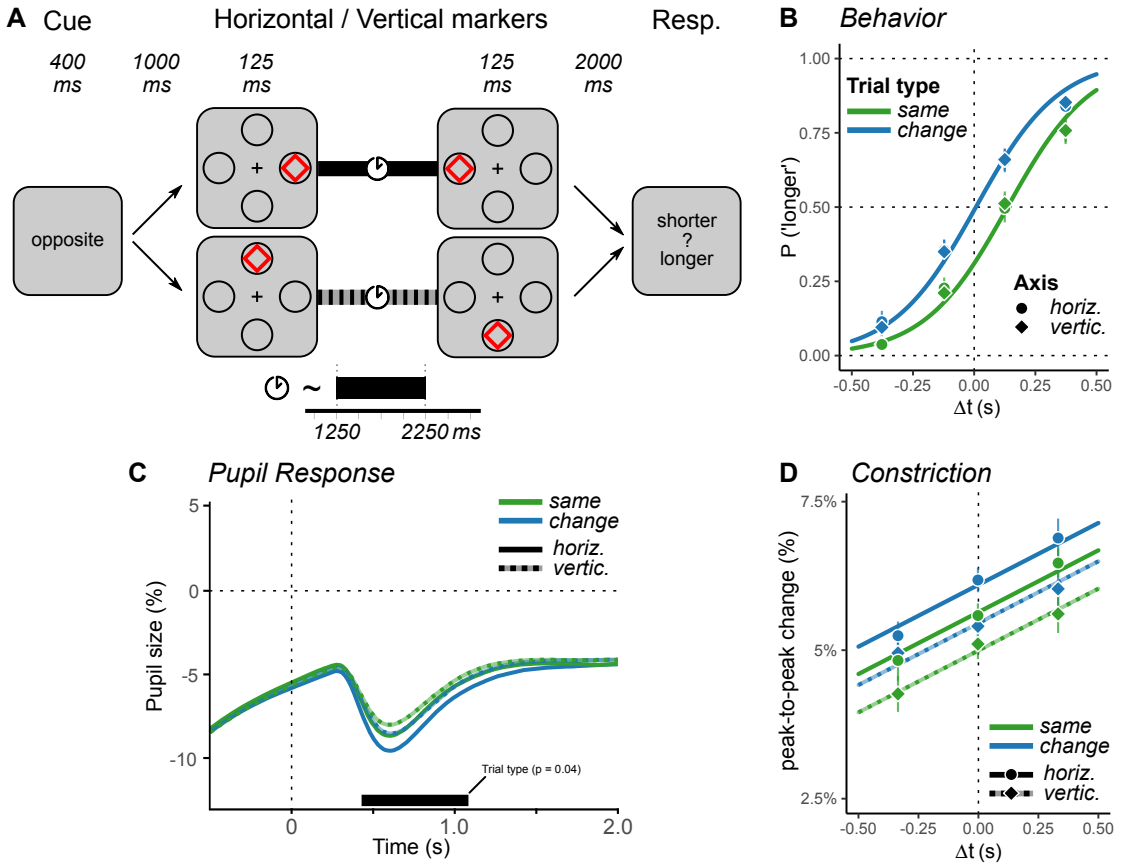

Figure 5. Design and results of Experiment 2. **A** General trial design. Two example 'change' trials are illustrated, for horizontal (upper path) and vertical (lower path) blocks. After the cue, the interval was presented through markers on the horizontal or vertical axis. Participants compared a randomly determined interval to the reference duration of 1750ms. **B** Behavior. The x-axis marks deviation (Δt) from the reference interval. As in Experiment 1, 'change' trials were perceived as longer. Performance was not affected by the marker presentation axis. Data points and error bars reflect average response in four bins of interval duration, as in Figure 1B. **C** Pupil size in response to the offset of the interval. A significant cluster indicates that the pupil constricts more on 'change' trials. **C** The peak-to-peak constriction response at interval offset. Constriction was larger with longer intervals, was larger in horizontal compared to vertical blocks, and larger on change- than on same-trials. Data bins and error bars as in **B**

in more detail with a LMM, but note that it would be misleading to use the pupil size as depicted in Figure 5C as the dependent variable for this LMM. This is because due to the sluggishness of the pupil response, carryover effects from the constriction response to the onset markers can still be observed. Therefore, we instead analyze the constriction response in isolation, by determining the peak-to-peak difference in pupil size in a window 100-700ms following the offset marker.

This peak-to-peak constriction response was best modeled as a function of three

additive predictors, as illustrated in Figure 5D. In line with the observations of the cluster test, one of these predictors was the 'trial type', with more constriction on change- compared to same-trials ($\Delta BIC = 24.01; \chi^2(1) = 45.75; p < 0.001$). Additionally, model comparisons indicated weaker constriction on 'vertical' compared to 'horizontal' trials ($\Delta BIC = 67.75; \chi^2(1) = 89.5; p < 0.001$), for which the permutation analysis had also suggested a trend. The LMM also allowed us to map the effect of interval duration as a continuous predictor. We found more constriction for longer intervals ($\Delta BIC = 284.26; \chi^2(1) = 304.96; p < 0.001$). There was some support for an interaction between 'axis' and 'duration', though this was not warranted on the basis of the BIC ($\Delta BIC = -13.68; \chi^2(1) = 5.56; p = 0.018$). This interaction (not depicted in Figure 5D) suggested that the duration effect on constriction was somewhat weaker in 'vertical' trials compared to horizontal trials. To some extent, the effects of duration might reflect a ceiling-effect on peak-to-peak constriction: on shorter trials, the pupil might already be in a constricted state due to the flashing onset marker. However, this explanation is rendered unlikely given the observation that even for short durations, the effect of trial type is observed with no support for an under-additive interaction ($\Delta BIC = -17.71; \chi^2(1) = 1.53; p = 0.216$). Instead, the effect of duration on pupil constriction might be similar to that which was observed for the P2 in Experiment 1: the effects of constriction might reflect stronger repetition suppression on shorter trials, as well as on 'same' trials. Unlike for P2 amplitude, however, we found no support that pupil constriction was predicted by the temporal percept ($\Delta BIC = -21.26; \chi^2(1) = 0.01; p = 0.905$).

Discussion Experiment 2

Experiment 2 confirmed the behavioral findings of Experiment 1, that a fully predictable stimulus change causes temporal dilation compared to stimulus repetition. The results provided further support for the magnitude coding hypothesis over alternative explanations of the bias, including arousal-based accounts. Behaviorally, the bias was constant across the tested range of durations, which once again suggested that there was no clock speedup in anticipation of an upcoming stimulus change. By strictly controlling for eye movements, both online and offline, we ruled out their potential role in causing the temporal bias. In addition, we found that the bias was virtually identical for horizontal and vertical changes, indicating that this axis does not play a role for the mechanism underlying the bias.

Pupil size measurements further supported a magnitude coding account of the change bias, more so than an account involving arousal modulations. We found stronger pupil constriction responses to the offset marker on change-trials. Even though we are not aware of studies that directly test whether sensory repetition suppression affects pupil constriction, the response has been shown to depend on physical stimulus properties as well as pre-attentive bottom-up salience (Binda et al., 2013; Mathôt et al., 2014). Our findings certainly do not reflect an effect of increased pupil dilation, which in previous work has been been associated with motivational salience, arousal, and violation of expectations (Akdoğan et al., 2016; Naber et al., 2013; Satterthwaite et al., 2007).

General Discussion

In two temporal discrimination tasks, we have demonstrated a bias in temporal perception caused by fully predictable changes in sensory input, and we have illustrated the behavioral and physiological properties of this bias. We found that a change in the location of a visual start- or end-marker leads to a longer perceived duration of the interval compared to a location repetition. In Experiment 1 we found that this bias in behavior was not modulated by the absolute or relative duration of the intervals, and did not attenuate as participants became more accustomed to changes or repetitions throughout a block: rather, the behavioral bias manifested as an additive effect, consistent across conditions throughout the experiment. ERP analyses indicated that stimulus changes and repetitions were paired with amplified and attenuated central P2 amplitudes, respectively. At the offset of the reference interval, always presented at a central location, the same P2 component was modulated by the interval duration, and was predictive of the upcoming discrimination judgement. These findings strongly relate the P2 amplitude to temporal perception. The P3 measured at the same sites was affected by the objective duration of the reference interval, but was not modulated by the bias nor was it predictive of the upcoming response. In Experiment 2, we replicated the behavioral bias effect in a bisection task, with similar magnitude for horizontal and vertical marker changes. Pupil size analyses showed that the bias was paired with larger constriction in response to change, similar to the P2 modulation in Experiment 1. Taken together, our findings offer direct evidence of magnitude coding effects on time perception. Stimulus repetitions and changes affect the relative magnitude of the sensory response, reflected in the P2 amplitude and pupil constriction, and subsequently affect perceived time. The bias did not seem to be mediated by surprise or arousal, as stimuli were fully predictable and the bias could not be related to P3 amplitude or pupil dilation.

Several differences set our experimental design apart from previous work on temporal biases, in particular from those using oddball paradigms. Most crucially, the sensory changes and repetitions in our task were fully predictable by design, either through blocked presentation or cueing. This allowed us to isolate the effects of repetition independent of surprise. A similar approach has been used previously to dissociate repetition suppression from expectation suppression effects (Larsson & Smith, 2012; Summerfield et al., 2008; Todorovic & Lange, 2012). Another difference is that in our design, stimulus repetitions and changes were physically identical up to the end of the interval, whereas in most oddball designs the deviant stimulus differs from the onset onwards throughout its duration. This means that in oddball designs, sensory magnitude effects could have ongoing effects on time perception throughout the interval, which would give rise to 'clock speedup'. This would explain why the change bias presented here was characterized by a shift of the psychometric curve, whereas in oddball designs it is often reported that the slope of the psychometric curve is additionally affected (Ernst et al., 2017; Sadeghi et al., 2011; Tse, Intriligator, Rivest, & Cavanagh, 2004). This interpretation aligns with previous MEG results (Noguchi & Kakigi, 2006) from a study where two intervals were compared that could either be the same or different. Here, it was found that changing stimuli were perceived as longer, and were characterized by a larger visual onset response followed by faster changes in climbing neural activity which the authors localized to the supplementary motor area and compared

to the CNV-component in EEG.

Regarding the CNV, it may seem remarkable that even though it was clearly present in our task, we did not find that it was related to the temporal bias, nor that it was related to discrimination responses. This extends earlier work that suggests that the CNV does not reflect the timing process itself, but instead marks processes that are contingent on time, such as preparing for upcoming stimuli or actions (Boehm et al., 2014; Kononowicz & Rijn, 2014; Van Rijn et al., 2011). In both experiments, several aspects of our design might have prevented such preparation. Intervals were uniformly distributed, responses were far removed from the offset of the interval, and the mapping of the percept to either of the keys was not known to the participant during estimation. Together, these factors might have prevented preparation-effects on the CNV that previous studies interpreted as a marker of the internal clock (Hiltraut, Rolf, & Gerhard, 2003; Macar & Vidal, 2003; Trillenberg, Verleger, Wascher, Wauschkuhn, & Wessel, 2000).

The P3 on the other hand, did seem to reflect aspects of the temporal comparison. At the end of Interval 1, the P3 was constant in amplitude across experimental factors, in line with our assumption that stimulus changes and repetitions were both fully anticipated. However, when the temporal comparison was to be made at the end of Interval 2, P3 amplitude was smaller for longer durations (cf. Lindbergh & Kieffaber, 2013), interacting with duration relative to Interval 1. This aligns with the classical interpretation of the P3 as index of surprise (Mars et al., 2008), and in addition point to the relation between the P3 and decision making processes (Twomey, Murphy, Kelly, & O'Connell, 2015). Taken together, these results suggest that the P2 and the P3 might index different aspects of temporal decision making in order to arrive at a comparison judgment.

Thus far we have phrased our findings within the context of the magnitude coding account: stimulus repetitions and changes cause attenuation or enhancement of sensory responses due to neural repetition suppression, which causally affects the perception of time. A subtly different interpretation of our findings would be that neural repetition suppression itself reliably depends on the time between repetitions, and that the resulting sensory magnitude is thus used by the brain as a heuristic to tell time. Not many studies have explored the effects of inter-stimulus intervals on repetition suppression, but some findings in rodent auditory cortex (Budd et al., 2013) and human visual cortex (Noguchi, Inui, & Kakigi, 2004) suggest that sensory response magnitude indeed monotonically increases as a function of this duration. This implies that the amount of repetition suppression to two subsequent stimuli could in theory be used to infer how much time has elapsed between them. Although our results might fit this interpretation, future research is necessary to fully understand the temporal dynamics of repetition suppression effects and whether they indeed contribute to our perception of time.

By dissociating stimulus repetition from stimulus expectation, we were able to identify a temporal bias that we could attribute to neural repetition suppression. We emphasize, however, that the inverse relation might not always hold, as not every form of repetition will necessarily affect temporal percepts. The precise effects of stimulus repetition are known to vary across stimuli (Amado & Kovács, 2016) and the cortical area under consideration (Larsson & Smith, 2012), to the extent that certain configurations demonstrate repetition enhancement rather than suppression (de Gardelle, Waszczuk, Egner, & Summerfield, 2013; Segaert, Weber, de Lange, Petersson, & Hagoort, 2013). It is to be expected that not every

sensory area plays a crucial part in temporal cognition and in computing perceived durations. In fact, paradigms like the one presented here might provide critical insights into which sensory systems do, without being affected by global modulations through arousal or surprise. Previous research suggests that 'change biases' similar to the location change effects presented here arise with changes in size (Matthews, 2011a) or changes in the direction of two arrowheads (Noguchi & Kakigi, 2006). Visual processing of stimulus location, size, and direction is classically attributed to the dorsal 'where' pathway (Haxby et al., 1991). These temporal change biases therefore converge with recent findings from neuroimaging that the the dorsal pathway, and parietal cortex in particular, plays a critical role in perceiving the timing of visual events (Battelli, Walsh, Pascual-Leone, & Cavanagh, 2008; Harvey, Dumoulin, Fracasso, & Paul, 2020; Harvey, Fracasso, Petridou, & Dumoulin, 2015; Walsh, 2003).

In conclusion, we believe that our findings present compelling evidence that sensory repetition effects affect human time perception, even when repetitions and changes are fully in line with expectations. Results from behavior, electrophysiology and pupillometry all supported for this interpretation, demonstrating that neural magnitude can modulate time perception in the absence of arousal-based effects. In neuroimaging, repetition effects sensory neurons have a strong tradition of being used as a tool to map out neural tuning in cortical areas, uncovering the pathways that give rise to a percept (Grill-Spector, Henson, & Martin, 2006). In the same vein, temporal biases like the change bias can be used to determine how low-level sensory responses give rise to a temporal percept.

References

Akdoğan, B., Balcı, F., & Rijn, H. v. (2016). Temporal expectation indexed by pupillary response. *Timing & Time Perception*, *4*(4), 354–370. doi:10.1163/22134468-00002075

Allan, L. G. & Gibbon, J. (1991). Human bisection at the geometric mean. *Learning and Motivation*, *22*(1), 39–58. doi:10.1016/0023-9690(91)90016-2

Amado, C. & Kovács, G. (2016). Does surprise enhancement or repetition suppression explain visual mismatch negativity? *European Journal of Neuroscience*, *43*(12), 1590–1600. doi:10.1111/ejn.13263

Amit, R., Abeles, D., Carrasco, M., & Yuval-Greenberg, S. (2019). Oculomotor inhibition reflects temporal expectations. *NeuroImage*, *184*, 279–292. doi:10.1016/j.neuroimage.2018.09.026

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Learned value magnifies salience-based attentional capture. *PLoS ONE*, *6*(11), e27926. doi:10.1371/journal.pone.0027926

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*. Special Issue: Emerging Data Analysis, *59*(4), 390–412. doi:10.1016/j.jml.2007.12.005

Battelli, L., Walsh, V., Pascual-Leone, A., & Cavanagh, P. (2008). The 'when' parietal pathway explored by lesion studies. *Current Opinion in Neurobiology*. Cognitive neuroscience, *18*(2), 120–126. doi:10.1016/j.conb.2008.08.004

Bigdely-Shamlo, N., Mullen, T., Kothe, C., Su, K.-M., & Robbins, K. A. (2015). The prep pipeline: Standardized preprocessing for large-scale eeg analysis. *Frontiers in Neuroinformatics*, *9*. doi:10.3389/fninf.2015.00016

Binda, P., Pereverzeva, M., & Murray, S. O. (2013). Attention to bright surfaces enhances the pupillary light reflex. *Journal of Neuroscience*, *33*(5), 2199–2204. doi:10.1523/JNEUROSCI.3440-12.2013

Boehm, U., van Maanen, L., Forstmann, B., & van Rijn, H. (2014). Trial-by-trial fluctuations in cnv amplitude reflect anticipatory adjustment of response caution. *NeuroImage*, *96*, 95–105. doi:10.1016/j.neuroimage.2014.03.063

Braem, S., Coenen, E., Bombeke, K., van Bochove, M. E., & Notebaert, W. (2015). Open your eyes for prediction errors. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(2), 374–380. doi:10.3758/s13415-014-0333-4

Budd, T. W., Nakamura, T., Fulham, W. R., Todd, J., Schall, U., Hunter, M., ... Michie, P. T. (2013). Repetition suppression of the rat auditory evoked potential at brief stimulus intervals. *Brain Research*, *1498*, 59–68. doi:10.1016/j.brainres.2012.12.033

Bueti, D., Bahrami, B., & Walsh, V. (2008). Sensory and association cortex in time perception. *Journal of Cognitive Neuroscience*, *20*(6), 1054–1062. doi:10.1162/jocn.2008.20060

Buia, C. & Tiesinga, P. (2006). Attentional modulation of firing rate and synchrony in a model cortical network. *Journal of Computational Neuroscience*, *20*(3), 247–264. doi:10.1007/s10827-006-6358-0

Burr, D. C., Ross, J., Binda, P., & Morrone, M. C. (2010). Saccades compress space, time and number. *Trends in Cognitive Sciences*, *14*(12), 528–533. doi:10.1016/j.tics.2010.09.005

Cho, J. R., Treweek, J. B., Robinson, J. E., Xiao, C., Bremner, L. R., Greenbaum, A., & Gradinaru, V. (2017). Dorsal raphe dopamine neurons modulate arousal and promote wakefulness by salient stimuli. *Neuron*, *94*(6), 1205–1219.e8. doi:10.1016/j.neuron.2017.05.020

Christman, S. D., Garvey, K. J., Propper, R. E., & Phaneuf, K. A. (2003). Bilateral eye movements enhance the retrieval of episodic memories. *Neuropsychology*, *17*(2), 221–229. doi:10.1037/0894-4105.17.2.221

Clark, A. (2013). Whatever next? predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(03), 181–204. doi:10.1017/S0140525X12000477

Correa, A. & Nobre, A. C. (2008). Neural modulation by regularity and passage of time. *Journal of Neurophysiology*, *100*(3), 1649–1655.

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. doi:10.20982/tqmp.01.1.p042

Crowley, K. E. & Colrain, I. M. (2004). A review of the evidence for p2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology*, *115*(4), 732–744. doi:10.1016/j.clinph.2003.11.021

de Gardelle, V., Waszczuk, M., Egner, T., & Summerfield, C. (2013). Concurrent repetition enhancement and suppression responses in extrastriate visual cortex. *Cerebral Cortex*, *23*(9), 2235–2244. doi:10.1093/cercor/bhs211

Delorme, A. & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. doi:10.1016/j.jneumeth.2003.10.009

Droit-Volet, S. & Meck, W. H. (2007). How emotions colour our perception of time. *Trends in Cognitive Sciences*, *11*(12), 504–513. doi:10.1016/j.tics.2007.09.008

Eagleman, D. M. & Pariyadath, V. (2009). Is subjective duration a signature of coding efficiency? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1525), 1841–1851. doi:10.1098/rstb.2009.0026

Ernst, B., Reichard, S. M., Riepl, R. F., Steinhauser, R., Zimmermann, S. F., & Steinhauser, M. (2017). The p3 and the subjective experience of time. *Neuropsychologia*, *103*, 12–19. doi:10.1016/j.neuropsychologia.2017.06.033

Failing, M. & Theeuwes, J. (2018). Selection history: How reward modulates selectivity of visual attention. *Psychonomic Bulletin & Review*, *25*(2), 514–538. doi:10.3758/s13423-017-1380-y

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1456), 815–836. doi:10.1098/rstb.2005.1622

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... Hämäläinen, M. (2013). Meg and eeg data analysis with mne-python. *Frontiers in Neuroscience*, *7*. doi:10.3389/fnins.2013.00267

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., ... Hämäläinen, M. S. (2014). Mne software for processing meg and eeg data. *NeuroImage*, *86*, 446–460. doi:10.1016/j.neuroimage.2013.10.027

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*(1), 14–23. doi:10.1016/j.tics.2005.11.006

Harrington, D. L., Haaland, K. Y., & Knight, R. T. (1998). Cortical networks underlying mechanisms of time perception. *Journal of Neuroscience*, *18*(3), 1085–1095. doi:10.1523/JNEUROSCI.18-03-01085.1998

Harvey, B. M., Dumoulin, S. O., Fracasso, A., & Paul, J. M. (2020). A network of topographic maps in human association cortex hierarchically transforms visual timing-selective responses. *Current Biology*. doi:10.1016/j.cub.2020.01.090

Harvey, B. M., Fracasso, A., Petridou, N., & Dumoulin, S. O. (2015). Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. *Proceedings of the National Academy of Sciences*, *112*(44), 13525–13530. doi:10.1073/pnas.1515414112

Haxby, J. V., Grady, C. L., Horwitz, B., Ungerleider, L. G., Mishkin, M., Carson, R. E., ... Rapoport, S. I. (1991). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. *Proceedings of the National Academy of Sciences*, *88*(5), 1621–1625. doi:10.1073/pnas.88.5.1621

Hiltraut, M.-G., Rolf, U., & Gerhard, R. (2003). Locus of the effect of temporal preparation: Evidence from the lateralized readiness potential. *Psychophysiology*, *40*(4), 597–611. doi:10.1111/1469-8986.00061

Horvitz, J. (2000). Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience*, *96*(4), 651–656. doi:10.1016/S0306-4522(00)00019-1

Jas, M., Engemann, D. A., Bekhti, Y., Raimondo, F., & Gramfort, A. (2017). Autoreject: Automated artifact rejection for meg and eeg data. *NeuroImage*, *159*, 417–429. doi:10.1016/j.neuroimage.2017.06.030

Jepma, M., Murphy, P. R., Nassar, M. R., Rangel-Gomez, M., Meeter, M., & Nieuwenhuis, S. (2016). Catecholaminergic regulation of learning rate in a dynamic environment. *PLoS Computational Biology*, *12*(10). doi:10.1371/journal.pcbi.1005171

Johnston, A., Arnold, D. H., & Nishida, S. (2006). Spatially localized distortions of event time. *Current Biology*, *16*(5), 472–479. doi:10.1016/j.cub.2006.01.032

Kaliukhovich, D. A. & Vogels, R. (2011). Stimulus repetition probability does not affect repetition suppression in macaque inferior temporal cortex. *Cerebral Cortex*, *21*(7), 1547–1558. doi:10.1093/cercor/bhq207

Kaskey, G. B., Salzman, L. F., Klorman, R., & Pass, H. L. (1980). Relationships between stimulus intensity and amplitude of visual and auditory event related potentials. *Biological Psychology*, *10*(2), 115–125. doi:10.1016/0301-0511(80)90032-0

Keil, A. & Ihssen, N. (2004). Identification facilitation for emotionally arousing verbs during the attentional blink. *Emotion*, *4*(1), 23–35. doi:10.1037/1528-3542.4.1.23

Kelly, S. P. & O'Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *Journal of Neuroscience*, *33*(50), 19434–19441. doi:10.1523/JNEUROSCI.3355-13.2013

Kim, E. & McAuley, J. D. (2013). Effects of pitch distance and likelihood on the perceived duration of deviant auditory events. *Attention, Perception, & Psychophysics*, *75*(7), 1547–1558. doi:10.3758/s13414-013-0490-5

Kononowicz, T. W. & Rijn, H. v. (2014). Decoupling interval timing and climbing neural activity: A dissociation between cnv and n1p2 amplitudes. *Journal of Neuroscience*, *34*(8), 2931–2939. doi:10.1523/JNEUROSCI.2523-13.2014

Larsson, J. & Smith, A. T. (2012). Fmri repetition suppression: Neuronal adaptation or stimulus expectation? *Cerebral Cortex*, *22*(3), 567–576. doi:10.1093/cercor/bhr119

Lindbergh, C. A. & Kieffaber, P. D. (2013). The neural correlates of temporal judgments in the duration bisection task. *Neuropsychologia*. Special Issue: How Does the Brain Process Time? *51*(2), 191–196. doi:10.1016/j.neuropsychologia.2012.09.001

Lopez-Calderon, J. & Luck, S. J. (2014). Erplab: An open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*. doi:10.3389/fnhum.2014.00213

Macar, F. & Vidal, F. (2003). The cnv peak: An index of decision making and temporal memory. *Psychophysiology*, *40*(6), 950–954. doi:10.1111/1469-8986.00113

Macar, F., Vidal, F., & Casini, L. (1999). The supplementary motor area in motor and sensory timing: Evidence from slow brain potential changes. *Experimental Brain Research*, *125*(3), 271–280. doi:10.1007/s002210050683

Mars, R. B., Debener, S., Gladwin, T. E., Harrison, L. M., Haggard, P., Rothwell, J. C., & Bestmann, S. (2008). Trial-by-trial fluctuations in the event-related electroencephalogram reflect dynamic changes in the degree of surprise. *Journal of Neuroscience*, *28*(47), 12539–12545. doi:10.1523/JNEUROSCI.2925-08.2008

Mather, M., Clewett, D., Sakaki, M., & Harley, C. W. (2016). Norepinephrine ignites local hotspots of neuronal excitation: How arousal amplifies selectivity in perception and memory. *Behavioral and Brain Sciences*, *39*. doi:10.1017/S0140525X15000667

Mather, M. & Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspectives on Psychological Science*, *6*(2), 114–133. doi:10.1177/1745691611400234

Mathôt, S., Dalmaijer, E., Grainger, J., & Stigchel, S. V. d. (2014). The pupillary light response reflects exogenous attention and inhibition of return. *Journal of Vision*, *14*(14), 7–7. doi:10.1167/14.14.7

Mathôt, S., Fabius, J., Van Heusden, E., & Van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior Research Methods*, *50*(1), 94–106. doi:10.3758/s13428-017-1007-2

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. doi:10.3758/s13428-011-0168-7

Matthews, W. J. (2011a). Can we use verbal estimation to dissect the internal clock? differentiating the effects of pacemaker rate, switch latencies, and judgment processes. *Behavioural Processes*, *86*(1), 68–74. doi:10.1016/j.beproc.2010.09.006

Matthews, W. J. (2011b). Stimulus repetition and the perception of time: The effects of prior exposure on temporal discrimination, judgment, and production. *PLOS ONE*, *6*(5), e19815. doi:10.1371/journal.pone.0019815

Matthews, W. J. & Meck, W. H. (2016). Temporal cognition: Connecting subjective time to perception, attention, and memory. *Psychological Bulletin*, *142*(8), 865–907. doi:10.1037/bul0000045

Mayo, J. P. & Sommer, M. A. (2013). Neuronal correlates of visual time perception at brief timescales. *Proceedings of the National Academy of Sciences*, *110*(4), 1506–1511. doi:10.1073/pnas.1217177110

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64. doi:10.20982/tqmp.04.2.p061

Naber, M., Frässle, S., Rutishauser, U., & Einhäuser, W. (2013). Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. *Journal of Vision*, *13*(2), 11–11. doi:10.1167/13.2.11

New, J. J. & Scholl, B. J. (2009). Subjective time dilation: Spatially local, object-based, or a global visual experience? *Journal of Vision*, *9*(2), 4–4. doi:10.1167/9.2.4

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the p3, and the locus coeruleus–norepinephrine system. *Psychological Bulletin*, *131*(4), 510–532. doi:10.1037/0033-2909.131.4.510

Noguchi, Y., Inui, K., & Kakigi, R. (2004). Temporal dynamics of neural adaptation effect in the human visual ventral stream. *Journal of Neuroscience*, *24*(28), 6283–6290. doi:10.1523/JNEUROSCI.0655-04.2004

Noguchi, Y. & Kakigi, R. (2006). Time representations can be made from nontemporal information in the brain: An meg study. *Cerebral Cortex*, *16*(12), 1797–1808. doi:10.1093/cercor/bhj117

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 1–9. doi:10.1155/2011/156869

Pariyadath, V. & Eagleman, D. M. (2008). Brief subjective durations contract with repetition. *Journal of Vision*, *8*(16), 11–11. doi:10.1167/8.16.11

Paton, J. J. & Buonomano, D. V. (2018). The neural basis of timing: Distributed mechanisms for diverse functions. *Neuron*, *98*(4), 687–705. doi:10.1016/j.neuron.2018.03.045

Peirce, J. W. (2007). Psychopy-psychophysics software in python. *Journal of Neuroscience Methods*, *162*(1-2), 8–13. doi:10.1016/j.jneumeth.2006.11.017

Penton-Voak, I. S., Edwards, H., Percival, A., & Wearden, J. H. (1996). Speeding up an internal clock in humans? effects of click trains on subjective duration. *Journal of Experimental Psychology: Animal Behavior Processes*, *22*(3), 307.

Rao, R. P. N. & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. doi:10.1038/4580

Sadeghi, N. G., Pariyadath, V., Apte, S., Eagleman, D. M., & Cook, E. P. (2011). Neural correlates of subsecond time distortion in the middle temporal area of visual cortex. *Journal of Cognitive Neuroscience*, *23*(12), 3829–3840. doi:10.1162/jocn_a_00071

Satterthwaite, T. D., Green, L., Myerson, J., Parker, J., Ramaratnam, M., & Buckner, R. L. (2007). Dissociable but inter-related systems of cognitive control and reward during decision making: Evidence from pupillometry and event-related fmri. *NeuroImage*, *37*(3), 1017–1031. doi:10.1016/j.neuroimage.2007.04.066

Segaert, K., Weber, K., de Lange, F. P., Petersson, K. M., & Hagoort, P. (2013). The suppression of repetition enhancement: A review of fmri studies. *Neuropsychologia*, *51*(1), 59–66. doi:10.1016/j.neuropsychologia.2012.11.006

Soares, S., Atallah, B. V., & Paton, J. J. (2016). Midbrain dopamine neurons control judgment of time. *Science*, *354*(6317), 1273–1277. doi:10.1126/science.aah5234

Summerfield, C., Trittschuh, E. H., Monti, J. M., Mesulam, M.-M., & Egner, T. (2008). Neural repetition suppression reflects fulfilled perceptual expectations. *Nature Neuroscience*, *11*(9), 1004–1006. doi:10.1038/nn.2163

Tang, M. F., Smout, C. A., Arabzadeh, E., & Mattingley, J. B. (2018). Prediction error and repetition suppression have distinct effects on neural representations of visual information. *eLife*, *7*, e33123. doi:10.7554/eLife.33123

Tarantino, V., Ehlis, A.-C., Baehne, C., Boreatti-Huemmer, A., Jacob, C., Bisiacchi, P., & Fallgatter, A. J. (2010). The time course of temporal discrimination: An erp study. *Clinical Neurophysiology*, *121*(1), 43–52. doi:10.1016/j.clinph.2009.09.014

Todorovic, A. & Lange, F. P. d. (2012). Repetition suppression and expectation suppression are dissociable in time in early auditory evoked fields. *The Journal of Neuroscience*, *32*(39), 13389–13395. doi:10.1523/JNEUROSCI.2227-12.2012

Trillenberg, P., Verleger, R., Wascher, E., Wauschkuhn, B., & Wessel, K. (2000). Cnv and temporal uncertainty with 'ageing' and 'non-ageing' s1-s2 intervals. *Clinical Neurophysiology*, *111*(7), 1216–1226. doi:10.1016/S1388-2457(00)00274-1

Tse, P. U., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics*, *66*(7), 1171–1189. doi:10.3758/BF03196844

Twomey, D. M., Murphy, P. R., Kelly, S. P., & O'Connell, R. G. (2015). The classic p300 encodes a build-to-threshold decision variable. *European Journal of Neuroscience*, *42*(1), 1636–1643. doi:10.1111/ejn.12936

Ulrich, R., Nitschke, J., & Rammsayer, T. (2006). Perceived duration of expected and unexpected stimuli. *Psychological Research*, *70*(2), 77–87. doi:10.1007/s00426-004-0195-4

Ulrich, R. & Vorberg, D. (2009). Estimating the difference limen in 2afc tasks: Pitfalls and improved estimators. *Attention, Perception, & Psychophysics*, *71*(6), 1219–1227. doi:10.3758/APP.71.6.1219

Van der Molen, M. J. W., Van der Molen, M. W., Ridderinkhof, K. R., Hamel, B. C. J., Curfs, L. M. G., & Ramakers, G. J. A. (2012). Auditory and visual cortical activity during selective attention in fragile x syndrome: A cascade of processing deficiencies. *Clinical Neurophysiology*, *123*(4), 720–729. doi:10.1016/j.clinph.2011.08.023

Van Rijn, H., Kononowicz, T. W., Meck, W. H., Ng, K. K., & Penney, T. B. (2011). Contingent negative variation and its relation to time estimation: A theoretical evaluation. *Frontiers in Integrative Neuroscience*, *5*. doi:10.3389/fnint.2011.00091

Vicario, C. M., Pecoraro, P., Turriziani, P., Koch, G., Caltagirone, C., & Oliveri, M. (2008). Relativistic compression and expansion of experiential time in the left and right space. *PLOS ONE*, *3*(3), e1716. doi:10.1371/journal.pone.0001716

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. doi:10.1016/j.tics.2003.09.002

Wearden, J. H. (1991). Do humans possess an internal clock with scalar timing properties? *Learning and Motivation*, *22*(1), 59–83. doi:10.1016/0023-9690(91)90017-3