Temporal Scaling of Human Scalp-Recorded Potentials

Cameron D. Hassall, Jack Harley, Nils Kolling, and Laurence T. Hunt*

Oxford Centre for Human Brain Activity, Wellcome Centre for Integrative Neuroimaging, Department of Psychiatry, University of Oxford, Oxford OX3 7JX, United Kingdom

*laurence.hunt@psych.ox.ac.uk

Abstract

Much of human behaviour is governed by common processes that unfold over varying timescales. Standard event-related potential analysis assumes *fixed-latency* responses relative to experimental events. However, recent single unit recordings in animals have revealed neural activity *scales* to span different durations during behaviours demanding flexible timing. Here, we introduce a novel approach to unmix fixed-time and scaled-time components in human magneto/electroencephalography (M/EEG) data. We use this to reveal consistent temporal scaling of human scalp-recorded potentials across four independent EEG datasets, including interval perception, production, prediction and value-based decision making. The topography of these potentials is consistent with a source in medial frontal cortex, as studied in animal recordings. Between-trial variation in the temporally scaled response predicts between-trial variation in subject reaction times, demonstrating the relevance of this temporally scaled signal for temporal variation in behaviour. Our results provide a general approach for studying flexibly timed behaviour in the human brain.

Action and perception in the real world require flexible timing. We can walk quickly or slowly, recognize the same piece of music played at different tempos, and form temporal expectations over long and short intervals. Reaction time variation in many cognitive tasks is modelled in terms of internal evidence accumulation[1], whereby the same dynamical process unfolds at different speeds on different trials.

Flexible timing is critical in our lives, yet its neural correlates have proven difficult to study. One source of difficulty is disagreement over how the brain represents time. For example, the classic pacemaker-accumulator model[2] relies on a dedicated timing mechanism. Other models represent time intrinsically through oscillatory alignment[3] or network population dynamics[4]. However, it has recently been shown that brain activity at the level of individual neurons can be best explained by a *temporal scaling* model[5,6].

When monkeys are cued to produce either a short or a long interval, medial frontal cortex (MFC) unit activity can be explained by a single response that is stretched or compressed according to the length of the produced interval – a temporally scaled response[5]. This suggests that flexible motor timing is achieved by adjusting the speed of a common neural process. Temporal scaling is also implicit in the neural correlates of evidence integration during sensory and value-based decision making[7]. Indeed, recent approaches to studying time-warped responses in neural populations have revealed time-warping as a common property across many different population recordings and behavioural tasks[8].

Successfully characterising scaled-time components in humans could open the door to studying flexible timing in more complex, hierarchical tasks such as music production or language perception, as well as in patient populations in which timing is impaired[9]. Yet it is currently unclear how temporal scaling of neural responses may manifest at the scalp (if at all) using non-invasive recording in humans. Although magnetoencephalography and electroencephalography (M/EEG) have played a prominent role in understanding the neural

basis of timing[10], the method typically used to analyse such data has been the event-related potential (ERP), which averages event-locked responses across multiple repetitions. This implicitly assumes that neural activity occurs at *fixed-time* latencies with respect to experimental events. Sometimes these event-related potentials have been found to ramp at different speeds for different temporal intervals[10], suggestive of temporal scaling - but crucially, they appear mixed at the scalp with fixed-time components due to the *superposition problem*[11].

We therefore developed an approach to *unmix* scaled-time and fixed-time components in the EEG (Fig. 1a). Our proposed method builds on recently developed regression-based approaches[12,13] that have proven useful in unmixing fixed-time components that overlap with one another, such as stimulus-related activity and response-related activity. We estimate the fixed-time ERPs using a general linear model (GLM) in which the design matrix is filled with time-lagged dummy variables (1s around the events of interest, 0s otherwise). Importantly, these 'stick functions' can overlap in time to capture overlap in the underlying neural responses (Fig. 1b). In situations without any overlap, the GLM would exactly return the conventional ERP. To reveal scaled-time responses, we allowed the duration of the stick function to vary depending upon the interval between stimulus and response, meaning that the same neural response could span different durations on different trials. We used a 'box-shaped' kernel to perform this interpolation, although we found that other interpolation methods (see Supplementary Fig. 1) produced similar results. The returned scaled-time potential is no longer a function of real-world ('wall clock') time, but instead a function of the *percentage of time elapsed* between stimulus and response.

As a proof of concept, we simulated data for at a single EEG sensor for an interval timing task, consisting of two *fixed-time components* (locked to cues and responses), and one

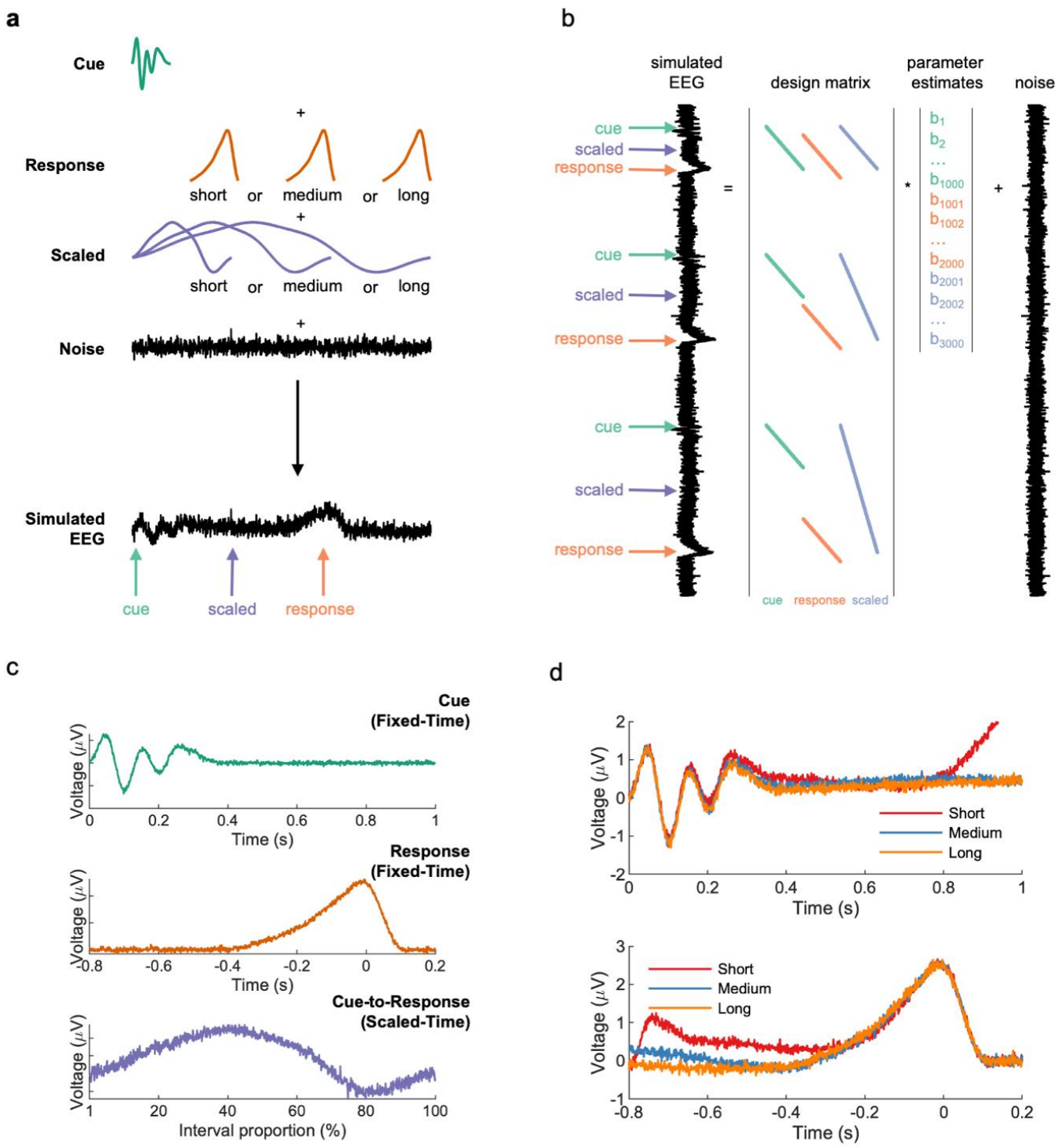

Figure 1. **Regression based unmixing of simulated data successfully recovers scaled-time and fixed-time components.** (a) EEG data were simulated by summing fixed-time components (cue and response), a scaled-time component with differing durations for different trials (short, medium, or long), and noise. (b) The simulated responses were unmixed via a GLM with stick basis functions: cue-locked, response-locked, and a single scaled-time basis spanning from cue to response. (c) The GLM successfully recovered all three components, including the scaled-time component. (d) A conventional ERP analysis (cue-locked and response-locked averages) of the same data obscured the scaled-time component.

scaled-time component spanning between cues and responses (Fig. 1a). Our proposed method was successful in recovering all three components (Fig. 1c), whereas a conventional ERP

approach obscured the scaled-time component (Fig. 1d). Crucially, in real EEG data we also repeat this approach across all sensors, potentially revealing different scalp topographies (and hence different neural sources) for fixed-time versus scaled-time components.

We used this approach to analyze EEG recorded across four independent datasets, comprising three interval timing tasks and one decision-making task. In the first task, participants produced a target interval (short, medium, or long) following a cue. Feedback was provided, and participants were able to closely match the target intervals. In the second, participants evaluated a computer-produced interval (Supplementary Fig. 2). The closer the produced interval was to the target interval, the more likely participants were to judge the response as 'on time' (Supplementary Fig. 3). In the third (previously analyzed[14,15]) task, participants made temporal predictions about upcoming events based on rhythmic predictions. In the fourth task (also previously analyzed[16,17]) participant chose between pairs of snack items – a process in which reaction time variability can be modelled as a process of internal evidence accumulation across time[18].

In all four tasks, we observed a scaled-time component that was distinct from the preceding and following fixed-time components (Fig. 2), which resembled conventional ERPs (Supplementary Fig. 4). Typically, ERP components are defined by their polarity, scalp distribution, and latency[19]. The observed scaled-time components shared a common polarity (negative) and scalp distribution (central), which is notably consistent with the medial frontal recording site where temporally scaled single-unit responses were previously identified[5]. Although our scaled-time components were estimated by time-warping a common signal so that they could span a variable delay period, their 'latency' was nevertheless consistent, in that the scaled-time signal grew and appeared to peak later in the timed interval. This is again reminiscent of the time course of scaled time components across the neural population in

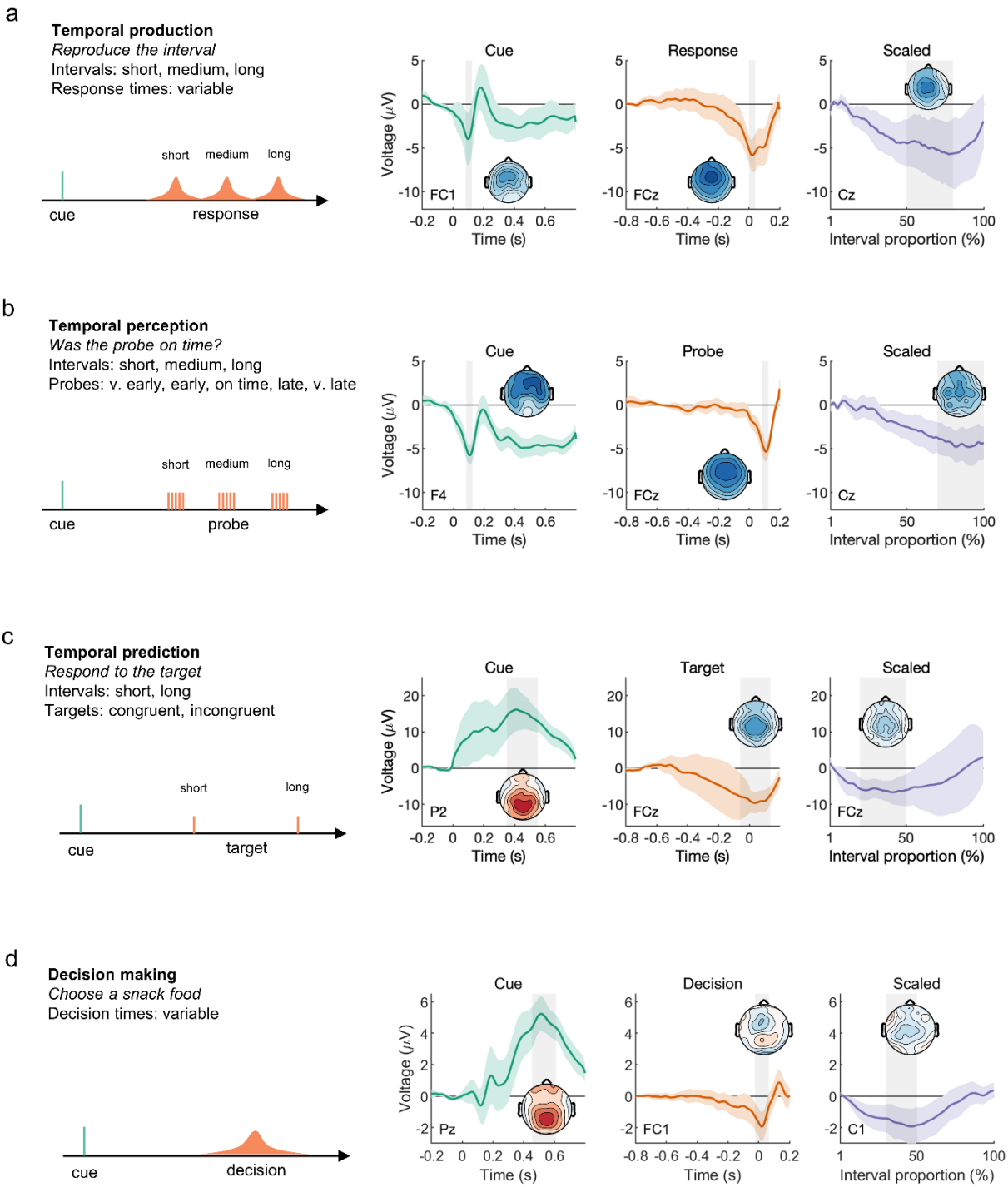

Figure 2. **Scaled-time components were consistently observed across three time-estimation paradigms and one decision-making paradigm, with distinct scalp topographies from fixed-time components.** Data were analysed from: (a) a temporal production task; (b) a temporal judgement task; (c) a temporal prediction task[11,12]; (d) a decision-making task[13,14]. All had distinct fixed-time components relative to task-relevant events (left/middle columns), and a common negative scaled-time component over central electrodes, reflecting interval time (right column). Error bars represent 95% confidence intervals.

medial frontal recordings[5]. Single-sample *t*-tests of the mean voltage in the shaded regions in Fig. 2 revealed a significant scaled-time signal in the production task ($t(9) = -3.19$, $p = .01$, Cohen's $d = -1.01$), the perception task ($t(9) = -4.79$, $p = .001$, Cohen's $d = -1.52$), the

prediction task ($t(18) = -4.03$, $p < .001$, Cohen's $d = -0.92$), and the decision-making task $t(17) = -3.40$, $p = .003$, Cohen's $d = -0.80$ (Supplementary Table 1). In many cases, scaled-time components were reliably observed at the single-subject level (Supplementary Figs. 4-7).

We then examined how the scaled-time component related to behavioural variability: does the latency of the scaled-time component predict participants' response time? We focussed on the temporal production and decision-making tasks, in which the scaled-time component spanned the duration from cue to response. To measure component latency, we applied an approach developed in [20,21], using principal component analysis (PCA) to model delay activity over central electrodes in the temporal production task, after first regressing out fixed-time components from the data. PCA was applied separately to each of three produced intervals. This consistently revealed a first principal component that matched the shape of the scaled-time component and a second principal component that matched its *temporal derivative*. This analysis confirms the presence of the scaled-time component in our data, as it is the first principal component of the residuals after removing fixed-time components. Crucially, adding or subtracting the second principal component captures variation in the *latency* of the scaled-time component (Fig. 3a). Across response time quantiles, we found that PC2 scores (Supplementary Table 2) were significantly related to response times (Fig. 3b; $F(2,18) = 9.05$, $p = .002$, $\eta_p^2 = 0.59$, $\eta_g^2 = 0.51$). This implies that the earlier in time that the scaled-time component peaked, the faster the subject would respond on that trial. This result was replicated in the decision-making task, $F(2,34) = 176.55$, $p < .001$, $\eta_p^2 = 0.91$, $\eta_g^2 = 0.80$ (Fig. 3c).

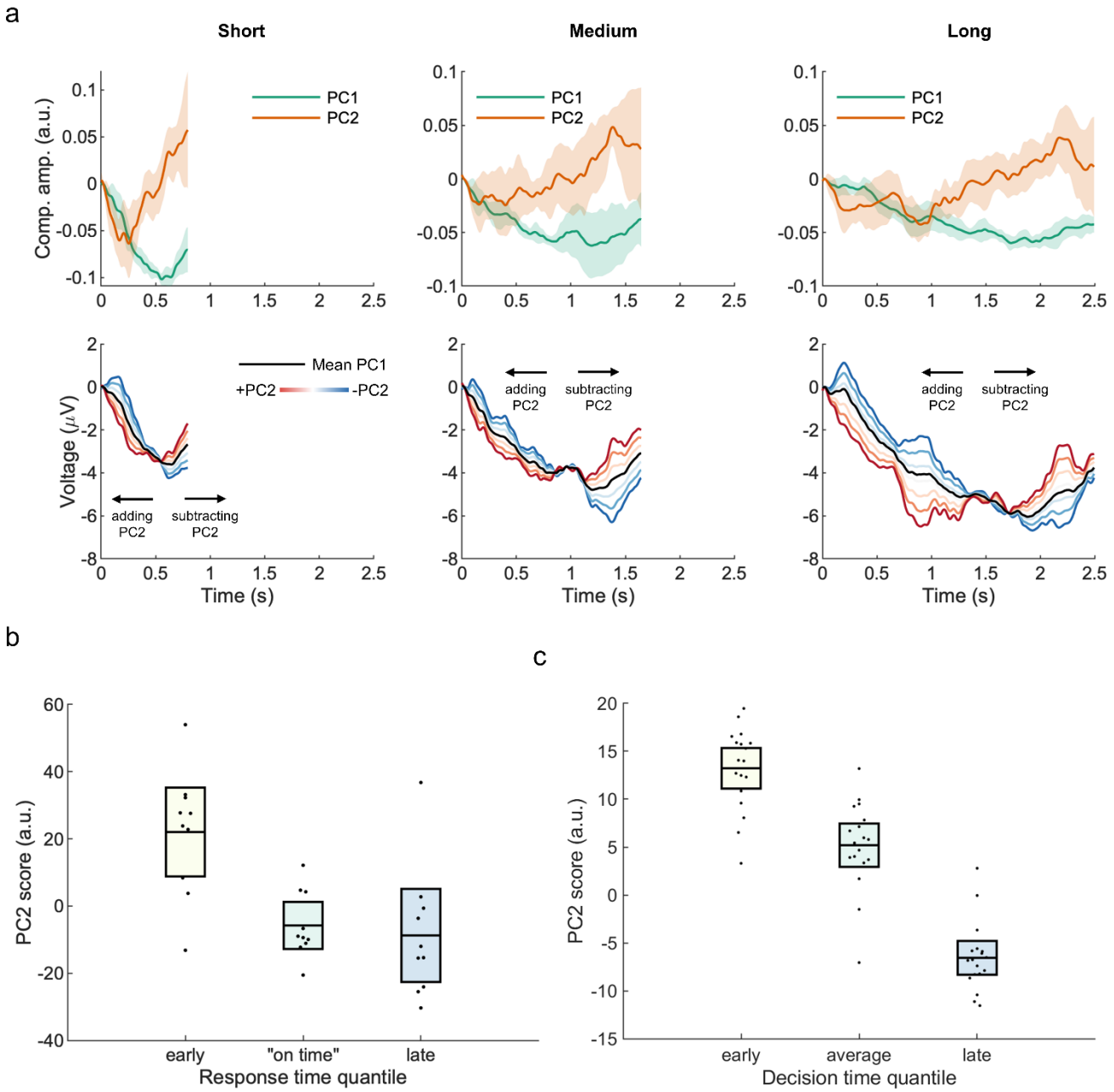

Figure 3. **Variation in scaled-time components predict behavioural variation in time estimation.** Cue-locked EEG over central electrodes (FC1, FCz, FC2, Cz, CP1, CPz, CP2) was grouped by response time (early, on time, or late), averaged, and stacked for each target interval (short, medium, or long). A separate PCA was run for each target interval and participant. (a) The first two principal components for each target interval represent the amplitude (PC1) and first derivative (PC2) of the time-scaled component (top panel). Adding or subtracting different amounts of PC2 to PC1 shifted the peak earlier or later in time (bottom panel). (b) PC2 scores depended on response time, implying the scaled-time component peaked earlier for fast responses and later for slow responses. (c) The effect replicated in a decision-making task. Error bars represent 95% confidence intervals.

Discussion

Our results provide a general method for recovering temporally scaled signals in human M/EEG, where scaled-time components are mixed at the scalp with conventional

fixed-time ERPs. We focused here on tasks that have been widely used in the timing literature, namely interval production, perception, and prediction, as well as an example of a cognitive task that exhibits variable reaction times across trials (value-based decision making). Distinct scaled-time components and scalp topographies were revealed in all four tasks. These results suggest that flexible cognition relies on temporally scaled neural activity, as seen in recent animal work[5,6].

The existence of temporally scaled signals at the scalp may not be surprising to those familiar with the study of time perception. Because of its excellent temporal resolution, EEG has long been used to study delay activity in interval timing tasks. These studies have revealed a ramping signal over frontal-central sites called the contingent negative variation[22] (CNV), which we observed in our conventional ERP analysis (Supplementary Fig. 4). Notably, CNV slope varies as a function of interval duration, and has been interpreted as an accumulation signal in pacemaker-accumulation models of timing[10]. Our work differs from these previous studies in one important respect. In a conventional ERP analysis, delay activity is assumed to occur over fixed latencies. The CNV is thus computed by averaging over many cue-locked EEG epochs of the same duration. In contrast, we have considered the possibility that scalped-recorded potentials reflect a mixture of both fixed-time and scaled-time components. By modelling fixed-time and scaled-time components separately, we revealed scaled activity that was common across all timed intervals. This, in turn, is more consistent with a recent class of models of timing that propose time estimation reflects the variable speed over which an underlying dynamical system unfolds[6].

We also observed temporally scaled activity in a decision-making task, a somewhat surprising result given that the task did not have an explicit timing component (participants made simple binary decisions[16]). Nevertheless, time is the medium within which decisions are made[23]. Computationally, the timing of binary decisions can be captured in a drift

diffusion model as the accumulation of evidence in favour of each alternative[1]. This accumulation is thought to be indexed by an ERP component called the central parietal positivity[24] (CPP). There is evidence that the slope of the CPP – which can be either stimulus-locked or response-locked – captures the rate of evidence accumulation[25]. For faster/easier decisions the CPP climbs more rapidly compared to slower/harder decisions[24,25]. We argue here that these effects can be explained by stretching/compressing a common scaled-time component while holding stimulus- and response-related activity constant. Furthermore, variation in the scaled-time component is relevant to decision making according to our results: it predicts *when* a decision will be made.

Although we have focused here on interval timing and decision-making tasks, we anticipate other temporally-scaled EEG and MEG signals will be discovered for cognitive processes known to unfold over varying timescales. For example, the neural basis of flexible sequential behaviours (such as speech) is still unknown, but may involve a form of temporal scaling[6]. Flexible timing is important across a vast array of decision-making tasks, where evidence accumulation can proceed quickly or slowly depending on the strength of the evidence[7]. Flexible timing can help facilitate a range of adaptive behaviours via temporal attention[26], while disordered timing characterizes several clinical disorders[27], underscoring the importance of characterising temporal scaling of neural responses in human participants.

Acknowledgements

This research was funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) Postdoctoral Fellowship to C.D.H., a BBSRC AFL Fellowship (BB/R010803/1) to N.K., a Sir Henry Dale Fellowship from the Royal Society and Wellcome (208789/Z/17/Z) to L.T.H., a NARSAD Young Investigator Award from the Brain and Behavior Research Foundation to L.T.H., and supported by the NIHR Oxford Health Biomedical Research Centre. The Wellcome Centre for Integrative Neuroimaging was supported by core funding from Wellcome Trust (203139/Z/16/Z). For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Author Contributions

N.K. and L.T.H. conceived the experiments and methodology. C.D.H. and L.T.H. designed the experiments and developed the methodology. C.D.H. and J.H. performed the experiments (except for the prediction task, where data was downloaded from a previous publication[14,15]). C.D.H. and L.T.H. analyzed the data. C.D.H. and L.T.H. wrote the manuscript with input from the other authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper.

Correspondence and requests for materials should be addressed to L.T.H.

Methods

Simulations

We simulated cue-related and response-related EEG in a temporal production task using MATLAB 2020a (Mathworks, Natick, USA). Cue and response were separated by either a short, medium, or long interval. During the delay period, we simulated a scaled response that stretched or compressed to fill the interval. All three responses (cue, response, scaled) were summed together at appropriate lags (short, medium, or long), with noise – see Fig. 1a. In total, we simulated 50 trials of each condition (short, medium, long).

To unmix fixed-time and scaled-time components, we used a regression-based approach[12,13,28] in which the continuous EEG at one sensor Y is modelled as a linear combination of the underlying event-related responses β, which are unknown initially. The model can be written in equation form as:

$$Y = X\beta + \varepsilon$$

where X is the design matrix and ε is the residual EEG not accounted for by the model. X contains as many rows as EEG data points, and as many columns as predictors (that is, the number of points in the estimated event-related responses). In our case, X was populated by 'stick functions' – non-zero values around the time of the modelled events, and zeros otherwise. We included in X two fixed-time components, the cue and the response, as stick functions of set EEG duration (with variables set to 1). In other words, the height of the fixed-time stick function was constant across events of the same type and equal to its width. To model a temporally-scaled response, we used the MATLAB *imresize* function (Image Processing Toolbox, R2020b) with 'box' interpolation to stretch/compress a stick function so that it spanned the duration between cue and response (other interpolation methods were tried – see Supplementary Fig. 1 – but this choice had little effect on the results). Thus, the duration of the scaled stick function varied from trial to trial (Fig. 1b). The goal here was to

estimate a single scaled-time response to account for EEG activity across multiple varying delay periods. For the fixed-time responses, each column of X represents a latency in ms before/after an experimental event; by contrast, for the scaled-time responses, each column of X represents the *percentage* of time that has elapsed between *two* events (stimulus and response). Simulation code is available at https://git.fmrib.ox.ac.uk/chassall/temporal-scaling.

Production and Perception Tasks

Participants

Participants completed both the production and perception tasks within the same recording session. We tested ten university-aged participants, 5 male, 2 left-handed, M_{age} = 23.40, 95% CI [21.29, 25.51]. Participants had normal or corrected-to-normal vision and no known neurological impairments. Participants provided informed consent approved by the Medical Sciences Interdivisional Research Ethics Committee at the University of Oxford. Following the experiment, participants were compensated £20 (£10 per hour of participation) plus a mean performance bonus of £3.23, 95% CI [2.92, 3.55].

Apparatus and Procedure

Participants were seated approximately 64 cm from a 27-inch LCD display (144 Hz, 1 ms response rate, 1920 by 1080 pixels, Acer XB270H, New Taipei City, Taiwan). Visual stimuli were presented using the Psychophysics Toolbox Extension[29,30] for MATLAB 2014b (Mathworks, Natick, USA). Participants were given written and verbal instructions to minimize head and eye movements. The goal of the production task was to produce a target interval and the goal of the perception task was to judge whether or not a computer-produced interval was correct.

The experiment was blocked with ten trials per block. There were 18 production blocks and 18 perception blocks, completed in random order. Prior to each block, participants listened to five isochronic tones indicating the target interval. Beeps were 400 Hz sine waves

of duration 50 ms and an onset/offset ramping to a point 1/8 of the length of the wave (to avoid abrupt transitions). The target interval was either short (0.8 s), medium (1.65 s), or long (2.5 s).

In production trials, participants listened to a beep then waited the target time before responding. Feedback appeared after a 400-600 ms delay (uniform distribution) and remained on the display for 1000 ms. Feedback was a ‘quarter-to’ clockface to indicate ‘too early’, a ‘quarter-after’ clockface to indicate ‘too late’, or a checkmark to indicate an on-time response. Feedback itself was determined by where the participant’s response fell relative to a window around the target duration. The response window was initialized to +/- 100 ms around each target, then changed following each feedback via a staircase procedure: increased on each side by 10 ms following a correct response and decreased by 10 ms following an incorrect response (either too early or too late).

In perception trials, participants heard two beeps, then were asked to judge the correctness of the interval, that is, whether or not the test interval matched the target interval. Test intervals (very early, early, on time, late, very late) were set such that each subsequent interval was 25% longer than the previous (see Supplementary Table 3). Participants were then given feedback on their judgement – a checkmark for a correct judgement, or an ‘x’ for an incorrect judgement.

For each task, participants gained 2 points for each correct response and lost 1 point for each incorrect response. At the end of the experiment points were converted to a monetary bonus at a rate of £0.01 per point.

Data Collection

In the perception task we recorded participant response time from cue, trial outcome (early, late, on time), and staircase- response window. In the production task, we recorded trial ‘on time’ judgements (yes/no), and trial outcome (correct/incorrect).

We recorded 36 channels of EEG, referenced to AFz. Data were recorded at 1000 Hz using a Synamps amplifier and CURRY 8 software (Compumetrics Neuroscan, Charlotte, USA). The electrodes were sintered Ag/AgCl (EasyCap, Herrsching, Germany). 31 of the electrodes were laid out according to the 10-20 system. Additional electrodes were placed on the left and right mastoids, on the outer canthi of the left and right eyes, and below the right eye. The reference electrode was placed at location AFz, and the ground electrode at Fpz.

Prediction Task

In this previously published[14,15] experiment, 19 participants responded to the onset of a visual target following a visual warning cue. The delay between cue and target was either short (700 ms) or long (1300 ms) and, in some conditions, congruent with a preceding stimulus stream. Only these predictable trials were included in the current analysis (i.e., the 'valid' trials in the 'Rhythmic' and 'Repeated-Interval' conditions – see [14,15] for more detail).

Decision-Making Task

In this experiment, also previously published[16,17], 18 participants were presented with two snack foods and asked to pick one. This was not an interval timing task and on average participants took 763 ms, 95% CI [713, 813], to respond.

Data Analysis

Behavioural data

For the perception task, we computed mean window size and mean produced interval for each participant. For the production task, we computed mean likelihood of responding yes to the 'on time' prompt, for each condition (short, medium, long) and interval (very early, early, on time, late, very late).

EEG Preprocessing

For all three timing tasks, EEG was preprocessed in MATLAB 2020b (Mathworks, Natick, USA) using EEGLAB[31]. We first down-sampled the EEG to 200 Hz, then applied a

0.1-20 Hz bandpass filter and 50 Hz notch filter. The EEG was then re-referenced to the average of the left and right mastoids (and AFz recovered in the production/perception tasks). Ocular artifacts were removed using independent component analysis (ICA). The ICA was trained on 3-second epochs of data following the appearance of the fixation cross at the beginning of each trial. Ocular components were identified using the *iclabel* function and then removed from the continuous data.

EEG for the decision-making task was already preprocessed prior to our analysis. This was a simultaneous EEG-fMRI recording, and preprocessing included the removal of MR-related artifacts via filtering and principal component analysis, as well as a 0.5-40 Hz bandpass filter. In line with our other analyses, we re-referenced the EEG to the average of TP7 and TP8 (located close to the mastoids) and applied an additional 20 Hz low-pass filter.

ERPs

To construct conventional event-related potentials (ERPs), we first created epochs of EEG around cues (all tasks), responses (perception task), probes (production task), targets (prediction task), and decisions (decision-making task). Cue-locked ERPs extended from 200 ms pre-cue to either 800, 1650, or 2500 ms post-cue (the short, medium, and long targets) in the perception/production tasks, 700 or 1300 ms in the prediction task (the short and long targets), and 600 ms in the decision-making task. Epochs were baseline-corrected using a 200 ms pre-cue window. We also constructed epochs from 800, 1650, or 2500 ms prior to the response/probe in the production/perception tasks, 700 or 1300 ms prior to the target in the prediction task, and 600 prior to the decision in the decision-making task to 200 ms after the response/probe/target/decision. A baseline was defined around the event of interest (mean EEG from -20 to 20 ms) and removed in all cases except for the decision-making task, in line with the original analysis[16]. We then removed any trials in which the sample-to-sample voltage differed by more than 50 μV or the voltage change across the entire epoch exceeded

150 μV. We then created conditional cue and response/probe/target/decision averages for each participant and task: production/perception (short, medium, and long), prediction (short and long), and decision-making (early and late, via a median split[16]). Finally, participant averages in the timing tasks were combined into grand-average waveforms at electrode FCz, a location where timing-related activity has been previously observed[10] and Pz in the decision-making task, in line with the previously published analysis[16] (Supplementary Fig. 4).

rERPs

To unmix fixed-time and scaled-time components in our EEG data, we estimated regression-ERPs (rERPs) following the same GLM procedure we used with our simulated data, but now applied to each sensor. We used a design matrix consisting of a regular stick functions for cue and response/probe/target and a stretched/compressed stick function spanning the interval from cue to response/probe/target/decision. In particular, we estimated cue-locked responses that spanned from 200 ms pre-cue to 800 ms post-cue. The response/probe/target/decision response interval spanned from -800 to 200 ms. Each fixed-time response thus spanned 1000 ms, or 200 EEG sample points. The scaled-time component, as described earlier, was modelled as a single underlying component (set width in X) that spanned over multiple EEG durations (varying number of rows in X). Thus, our method required choosing how many scaled-time sample points to estimate (the width in X). For the production/perception tasks, we chose to estimate 330 scaled-time points, equivalent to the duration of the 'medium' interval. For the prediction task, we chose to estimate 200 scaled-time points, equivalent to the mean of the short and long conditions (700 ms, 1300 ms). For the decision-making task, we estimated 153 scaled-time points (roughly equivalent to the mean decision time). Unlike the conventional ERP approach, this analysis was conducted on the continuous EEG. To identify artifacts in the continuous EEG, we used the *basicrap* function from the ERPLAB[32] toolbox with a 150 μV threshold (2000 ms window, 1000 ms

step size). A sample was flagged if it was ‘bad’ for any channel. Flagged samples were excluded from the GLM (samples removed from the EEG and rows removed from the design matrix). Additionally, we removed samples/rows associated with unusually fast or slow responses in the production task (less than 0.2 s or more than 5 s). On average, we removed 2.17 % of samples in the production task (95% CI [1.49, 2.86]), 3.75 % of samples in the perception task (95% CI [2.39, 5.10]), 1.03% of samples in the prediction task (95% CI [0.95, 1.10]), and 5.56% of samples in the decision-making task (95% CI [4.99, 6.12]).

To impose a smoothness constraint on our estimates, we used a first-derivative form of Tikhonov regularization. Tikhonov regularization reframes the GLM solution as the minimization of:

$$\|X\beta - \mathrm{Y}\|^2 + \lambda\|\mathrm{L}\beta\|^2$$

where L is the regularization operator and λ is the regularization parameter. In other words, we aimed to minimize a penalty term in addition to the usual residual. This has the solution

$$(X^T X + \lambda\mathrm{L})^{-1} X^T Y$$

In our case, L approximated the first derivative as a scaled finite difference[33]:

$$L = \frac{1}{2}\begin{bmatrix} 1 & -1 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -1 \end{bmatrix}$$

We then chose regularization parameters for each participant using 10-fold cross validation. Our goal here was to minimize the mean-squared error of the residual EEG at electrode FCz, our electrode of interest. The following λs were tested on each fold: 0.001, 0.01, 0, 1, 10, 100, 1000, 10000, 100000. An optimal λ was chosen for each participant corresponding to the parameter with the lowest mean mean-squared error across all folds. See Supplementary Table 4 for a summary.

Inferential Statistics

We quantified the amplitude of the scaled-time component in two ways. First, we computed a 95% confidence interval at each 'timepoint' in the scaled-time signal. Next, in line with conventional ERP analyses[19], we computed the mean signal around the apparent peak. We then confirmed the existence of the signal using a single-sample two-sided repeated-measures *t*-test. The assumption of normality was tested for each task using the Shapiro-Wilk test, revealing that the mean signal in the perception and decision-making tasks violated the assumption of normality. However, as repeated-measures *t*-tests are robust to violations of normality, no corrections were made. For all *t*-tests we also computed Cohen's *d* by dividing the mean of the mean scaled components by the standard deviation of the mean scaled components. These analyses were done at the scalp location where the mean of the mean signals was greatest, i.e. electrode Cz in the production/perception tasks, FCz in the prediction task, and C1 in the decision-making task.

PCA

To explore the link between the scaled-time component and behaviour, we regressed out the fixed-time components from the EEG in the temporal production task – that is, we reconstructed the preprocessed data using only the scaled-time regressors plus residuals. Only mid-frontal electrodes were considered: FC1, FCz, FC2, Cz, CP1, CPz, and CP2. We then constructed epochs starting at the cue and ending at the target interval (800 ms, 1650 ms, or 2500 ms). Epochs within each condition (short, medium, long) were further grouped into three equal-sized response-time bins (early, on time, late) and averaged for each electrode and participant. We then conducted a PCA for each condition (short, medium, long) and participant. See Supplementary Table 5 for amount of variance explained by PC1 and PC2. To visualize the effect of PC2, we computed the mean PC2 across all participants. We then added more or less of the mean PC2 to the mean PC1 projection and applied a 25-point moving-mean window for visualization purposes (Fig. 3a). In order to choose a reasonable

range of PC2 scores, we examined the average minimum and maximum PC2 score for each participant and condition (short, medium, long). The PC2 score ranges were -21 to 15 (short), -41 to 38 (medium), and -40 to 55 (long). To assess the relationship between PC2 score and behaviour, we binned PC2 scores according to our response time bins (early, on time, late) and collapsed across conditions (short, medium, long). This gave us as single mean PC2 score for each participant and response time bin (early, on time, late), which we analyzed using a two-sided repeated-measures ANOVA (Fig. 3b) after verifying the assumption of normality using the Shapiro-Wilk test. Two different effect sizes, η_p^2 and η_g^2, were computed, according to:

$$\eta_p^2 = \frac{SS_Q}{SS_Q + SS_{sQ}} \qquad \eta_g^2 = \frac{SS_Q}{SS_Q + SS_S + SS_{sQ}}$$

where SS_Q is the sum of squares of the quantile effect (early, on time, late), SS_{sQ} is the error sum of squares of the quantile effect, and SS_S is the sum of squares between subjects[34].

We then replicated the PCA procedure for the decision-making task using an epoch extending 1000 ms from the cue at a central electrode cluster (FC3, FC1, FC2, FC4, C3, C1, Cz, C2, C4, CP3, CP1, CP2, CP4, P3, P1, Pz, P2, and P4). Note that the assumption of normality was violated for ‘late’ responses in the decision-making task – again, no statistical correction was made.

Data Availability

Raw and preprocessed EEG for the production and perception tasks will be made publicly available at the time of publication. Raw data for the prediction task is available at https://doi.org/10.5061/dryad.5vb8h. Raw data for the decision-making task is available at https://openneuro.org/datasets/ds002734/versions/1.0.2.

Code Availability

Simulation and analysis scripts are available athttps://git.fmrib.ox.ac.uk/chassall/temporal-scaling.

References

1. Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. The Physics of Optimal Decision Making: A Formal Analysis of Models of Performance in Two-Alternative Forced-Choice Tasks. 66.
2. Treisman, M. Temporal discrimination and the indifference interval: Implications for a model of the 'internal clock'. *Psychological Monographs: General and Applied* **77**, 1–31 (1963).
3. Matell, M. S. & Meck, W. H. Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cognitive Brain Research* **21**, 139–170 (2004).
4. Buonomano, D. V. & Maass, W. State-dependent computations: spatiotemporal processing in cortical networks. *Nat. Rev. Neurosci.* **10**, 113–125 (2009).
5. Wang, J., Narain, D., Hosseini, E. A. & Jazayeri, M. Flexible timing by temporal scaling of cortical responses. *Nat Neurosci* **21**, 102–110 (2018).
6. Remington, E. D., Egger, S. W., Narain, D., Wang, J. & Jazayeri, M. A Dynamical Systems Perspective on Flexible Motor Timing. *Trends in Cognitive Sciences* **22**, 938–952 (2018).
7. O'Connell, R. G., Shadlen, M. N., Wong-Lin, K. & Kelly, S. P. Bridging Neural and Computational Viewpoints on Perceptual Decision-Making. *Trends in Neurosciences* **41**, 838–852 (2018).
8. Williams, A. H. *et al.* Discovering Precise Temporal Patterns in Large-Scale Neural Recordings through Robust and Interpretable Time Warping. *Neuron* **105**, 246-259.e8 (2020).
9. Meck, W. H. Neuropsychology of timing and time perception. *Brain and Cognition* **58**, 1–8 (2005).

10. Macar, F. & Vidal, F. Event-Related Potentials as Indices of Time Processing: A Review. *Journal of Psychophysiology* **18**, 89–104 (2004).

11. Luck, S. J. A Closer Look at ERPs and ERP Components. in *An introduction to the event-related potential technique* 35–70 (The MIT Press, 2014).

12. Smith, N. J. & Kutas, M. Regression-based estimation of ERP waveforms: I. The rERP framework. *Psychophysiology* **52**, 157–168 (2015).

13. Ehinger, B. V. & Dimigen, O. Unfold: an integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis. *PeerJ* **7**, e7838 (2019).

14. Breska, A. & Deouell, L. Y. Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLOS Biology* **15**, e2001665 (2017).

15. Breska, A. & Deouell, L. Y. Data from: Neural mechanisms of rhythm-based temporal prediction: delta phase-locking reflects temporal predictability but not rhythmic entrainment. (2017) doi:doi:10.5061/dryad.5vb8h.

16. Pisauro, M. A., Fouragnan, E., Retzler, C. & Philiastides, M. G. Neural correlates of evidence accumulation during value-based decisions revealed via simultaneous EEG-fMRI. *Nat Commun* **8**, 1–9 (2017).

17. Pisauro, M. A., Fouragnan, E., Retzler, C. & Philiastides, M. G. Evidence Accumulation in Value-Based decisions. (2020) doi:10.18112/openneuro.ds002734.v1.0.2.

18. Krajbich, I., Armel, C. & Rangel, A. Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience* **13**, 10 (2011).

19. Luck, S. J. *An introduction to the event-related potential technique*. (The MIT Press, 2014).

20. Möcks, J. The Influence of Latency Jitter in Principal Component Analysis of Event-Related Potentials. *Psychophysiology* **23**, 480–484 (1986).

21. Hunt, L. T., Behrens, T. E., Hosokawa, T., Wallis, J. D. & Kennerley, S. W. Capturing the temporal evolution of choice across prefrontal cortex. *eLife* **4**, e11945 (2015).

22. Walter, W. G., Cooper, R., Aldridge, V. J., McCallum, W. C. & Winter, A. L. Contingent Negative Variation : An Electric Sign of Sensori-Motor Association and Expectancy in the Human Brain. *Nature* **203**, 380–384 (1964).

23. Ariely, D. & Zakay, D. A timely account of the role of duration in decision making. *Acta Psychologica* **108**, 187–207 (2001).

24. O'Connell, R. G., Dockree, P. M. & Kelly, S. P. A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. *Nature Neuroscience* **15**, 1729–1735 (2012).

25. Kelly, S. P. & O'Connell, R. G. Internal and External Influences on the Rate of Sensory Evidence Accumulation in the Human Brain. *Journal of Neuroscience* **33**, 19434–19441 (2013).

26. Nobre, A. C. & van Ede, F. Anticipated moments: temporal structure in attention. *Nature Reviews Neuroscience* **19**, 34–48 (2018).

27. Allman, M. J. & Meck, W. H. Pathophysiological distortions in time perception and timed performance. *Brain* **135**, 656–677 (2012).

28. Smith, N. J. & Kutas, M. Regression-based estimation of ERP waveforms: II. Nonlinear effects, overlap correction, and practical considerations. *Psychophysiology* **52**, 169–181 (2015).

29. Brainard, D. H. The psychophysics toolbox. *Spatial vision* **10**, 433–436 (1997).

30. Pelli, D. G. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision* **10**, 437–442 (1997).

31. Delorme, A. & Makeig, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* **134**, 9–21 (2004).

32. Lopez-Calderon, J. & Luck, S. J. ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front. Hum. Neurosci.* **8**, (2014).

33. Reichel, L. & Ye, Q. Simple Square Smoothing Regularization Operators. *Electronic Transactions on Numerical Analysis* **33**, 63–83 (2008).

34. Olejnik, S. & Algina, J. Generalized eta and omega squared statistics: measures of effect size for some common research designs. *Psychol Methods* **8**, 434–447 (2003).