Neural signatures of evidence accumulation in temporal decisions

Highlights

- Neural responses to stimuli offsets depend on their duration and participants’ decision
- These responses reflect subjective decisions even for identical durations
- A computational model captures behavioral performance and predicts physiology
- This finding generalizes to vision and touch as well as to different time ranges

Authors

Nir Ofir, Ayelet N. Landau

Correspondence

nir.ofir@mail.huji.ac.il (N.O.), ayelet.landau@mail.huji.ac.il (A.N.L.)

In brief

Ofir and Landau show that the EEG response to stimuli offsets in a temporal categorization task reflect their duration and participants’ decisions. The responses match the distance to bound in a drift-diffusion model and generalize across sensory modalities and time ranges. The work provides a new way to study the cognitive processes underlying temporal decisions.
Neural signatures of evidence accumulation in temporal decisions

Nir Ofir1,2,3,* and Ayelet N. Landau1,2,4,*

1Department of Psychology, Hebrew University of Jerusalem, Mt. Scopus, Jerusalem 9190501, Israel
2Department of Cognitive and Brain Sciences, Hebrew University of Jerusalem, Mt. Scopus, Jerusalem 9190501, Israel
3Edmond and Lily Safra Center for Brain Sciences, Hebrew University of Jerusalem, Edmond J. Safra Campus, Jerusalem 9190401, Israel
4Lead contact
*Correspondence: nir.ofir@mail.huji.ac.il (N.O.), ayelet.landau@mail.huji.ac.il (A.N.L.)

https://doi.org/10.1016/j.cub.2022.08.006

SUMMARY

Cognitive models of interval timing can be formulated as an accumulation-to-bound process.1–5 However, the physiological manifestation of such processes has not yet been identified. We used electroencephalography (EEG) to measure the neural responses of participants while they performed a temporal bisection task in which they were requested to categorize the duration of visual stimuli as short or long.6 We found that the stimulus-offset and response-locked activity depends on both stimulus duration and the participants’ decision. To relate this activity to the underlying cognitive processes, we used a drift-diffusion model.7 The model includes a noisy accumulator starting with the stimulus onset and a decision threshold. According to the model, a stimulus duration will be categorized as “long” if the accumulator reaches the threshold during stimulus presentation. Otherwise, it will be categorized as “short.” We found that at the offset of stimulus presentation, an EEG response marks the distance of the accumulator from the threshold. Therefore, this model offers an accurate description of our behavioral data as well as the EEG response using the same two model parameters. We then replicated this finding in an identical experiment conducted in the tactile domain. We also extended this finding to two different temporal ranges (sub- and supra-second). Taken together, the work provides a new way to study the cognitive processes underlying temporal decisions, using a combination of behavior, EEG, and modeling.

RESULTS

Humans and non-human animals routinely make decisions about elapsed time. Such decisions are encountered daily, when waiting for a meal to cook or when preparing to move in anticipation of a changing traffic signal. These decisions can typically be phrased as “has it been long enough since a certain event?” Theoretical studies of interval timing often describe the representation of duration as an accumulation of internally produced ticks (pacemaker-accumulator models),1–5 which can be formulated as a bounded decision process.8,9 The threshold in such models marks when a target duration has elapsed. However, clear physiological manifestations of the abstract accumulation process or the threshold have not yet been identified.10,11

Offset event-related potential (ERP) amplitude depends on stimulus duration and participants’ decision

We examined the electroencephalography (EEG) response of participants while they categorized stimuli based on their duration (Figure 1). The task started with a short familiarization stage in which participants viewed only the shortest (0.5 s) and longest (2.1 s) durations as reference stimuli. After familiarization, participants viewed stimuli that were presented at a variable duration between the two reference durations, and they were required to classify each stimulus as being more similar to the short or the long reference (STAR Methods).6 Their EEG was measured and analyzed for the stimulus-offset response.

About 300 ms after stimulus offset, there is a large positive deflection that depends on stimulus duration: shorter durations elicit a larger potential (Figure 2A). The potential amplitude diminishes with longer durations and saturates at a certain stimulus duration (Figures 2A and 2B). Importantly, the stimulus duration at which the response saturates is near the duration wherein group performance is closest to 50% (i.e., the bisection point; compare Figures 2B and 2D). Finally, these responses predict the subjective perceived time: a given stimulus with a duration closest to an individual’s bisection point elicited a larger potential if it was categorized as “short” rather than “long” (Figure 2C). Three main factors could contribute to the offset response: stimulus duration, participants’ decision, and response preparation, with the first two being the factors of interest. As these three factors are not independent, we used a hierarchical mass univariate general linear model (GLM) approach to test the dependence of the offset-evoked EEG on all three.13 At the first level, the data of each participant were modeled using stimulus duration (continuous predictor, twelve durations between 0.5 and 2.1 s), participant response to that stimulus (categorical predictor, “long” or “short”), and their interaction while controlling for the response time (continuous predictor; STAR Methods). Beta coefficients were estimated for each electrode
and time point. At the second level, beta estimates for each predictor were entered into a non-parametric cluster-based analysis to evaluate the significance of these effects at the group level over time and scalp location (STAR Methods; Figure S2)\textsuperscript{14,15}.

This analysis revealed a clear effect of stimulus duration, participants’ decision, and their interaction on the offset EEG responses. First, the offset response was related to stimulus duration most clearly for “short” trials. This was supported by a significant interaction between duration and decision (p < 10\textsuperscript{-4}, cluster extending from 271 to 500 ms following offset). Within stimuli categorized as “short,” shorter stimuli elicited larger offset responses (Figure 2A; post hoc test for duration effect, p < 10\textsuperscript{-4}, cluster extending from 273 to 500 ms following offset). In contrast, within stimuli categorized as “long,” we did not find a significant effect of duration. Last, for stimuli presented for the mean duration (1.3 s, close to the mean bisection point of 1.25 s), stimuli categorized as “short” were associated with a larger offset response compared with those categorized as “long.” This was supported by the significant effect of the participants’ decision (Figure 2C; p < 10\textsuperscript{-4}, cluster extending from 283 to 500 ms following offset). This effect strongly supports the interpretation of the EEG response as reflecting subjective rather than physical duration: although the same physical duration is presented, “short” decisions elicited a significantly larger offset response in the EEG compared with “long” decisions. The effects of duration, decision, and their interaction were most evident in a group of frontocentral electrodes 300–500 ms after stimulus offset (for a full report of the spatiotemporal extent of significant clusters, see Figure S2). Within this cluster, we used a subset of electrodes to relate the physiology to the underlying cognitive processes in the task (STAR Methods).

**Offset ERPs reflect the distance from the decision threshold**

In summary, the EEG offset responses are modulated both by the duration of the stimulus and the decision of the participant: stimuli categorized as “long” elicit virtually no response, regardless of the stimulus duration, while for stimuli categorized as “short,” shorter durations elicit larger responses. This pattern suggests that the EEG reflects the cognitive process that underlies temporal bisection performance. Timing behaviors are often modeled by pacemaker-accumulator models, which can be formulated as different versions of accumulation-to-bound processes. We use a simplification of ToPDDM\textsuperscript{7} (time-adaptive, opponent Poisson drift-diffusion model), the only model to explicitly consider the decisions being made before the stimulus ends.\textsuperscript{16} The basis of the model is a drift-diffusion process that reflects an internal dynamic representation of duration (i.e., evidence accumulation). This process starts at the stimulus onset and continues until either the decision threshold is reached or the stimulus offsets. If the threshold is reached during stimulus presentation, evidence accumulation terminates,\textsuperscript{17} and the stimulus is categorized as “long”; otherwise, the stimulus is categorized as “short” (Figures 2G and 2H). The EEG offset response reflects the distance from the decision threshold (Figures 2E and 2F): for “long” trials there is no response, regardless of stimulus duration, because the threshold has already been reached, and for “short” trials, the shorter the stimulus, the larger the distance to the threshold at the offset (for a full description of the model and its relation to the original ToPDDM, see STAR Methods).

To assess the qualitative fit of the model to the EEG data, we estimated the model parameters (the ratio of the threshold to the drift rate and the squared ratio of the threshold to the diffusion rate\textsuperscript{8}; STAR Methods) using the average behavioral data (Figure 2D). We then used the same parameters to predict the average EEG response with a simulation (STAR Methods). The model parameters fit to participants’ behavioral performance capture the offset response pattern accurately (R\textsuperscript{2} = 0.98; Figure 2B; for parameter estimates, see Table S1). We emphasize that the purpose of introducing the cognitive model is to explain the EEG pattern established by the GLM analysis rather than to estimate the cognitive parameters using the EEG\textsuperscript{18} (STAR Methods).

To investigate the scope of the reported pattern, we examined whether the findings generalize to another sensory modality. Finding the same pattern in another modality would serve as a conceptual replication as well as an important extension indicating that this temporal-evidence accumulation process is modality independent.

**The accumulated evidence index is modality independent**

In experiment 2, individuals performed the exact same task in the tactile modality. Participants had to judge the duration of a vibrotactile stimulus that was presented to the fingers of their left hand. Instructions, experimental design, and EEG data handling were identical to those in experiment 1 (STAR Methods).

As in experiment 1, the statistical analysis, using the same GLM, reveals the clear effect of stimulus duration, participants’ decision,
Figure 2. EEG tracks accumulated evidence in temporal bisection

(A–D) Neural responses to stimuli offsets depend on stimulus duration and participants’ decision.
(A) EEG time-locked to stimulus offset as a function of stimulus duration (color coded; data for EEG traces were lowpass filtered for visualization only). Shaded areas mark the SEM across participants. The gray horizontal line marks the epoch of the most significant cluster of the effect of duration for trials categorized as “short.”
(B) Average potential across the frontocentral electrodes in a time window of 300–500 ms post-offset, per stimulus duration (same color code as A). Circles show the observed average potentials, error bars depict the SEM across participants, and the solid curve shows the responses predicted by the cognitive model fit to the behavior ($R^2 = 0.98$; see offset ERPs reflect the distance from the decision threshold section in the results). The interaction of the stimulus duration and participant response is evident in the saturating (i.e., decreasing) slope. The inset shows the scalp distribution of the cognitive model fits in the same time window. The frontocentral electrodes depicted in (A)–(C) are marked in white.
(C) Offset responses to the stimuli with durations closest to individual participants’ bisection points, separated by participants’ decision. Shaded areas mark the SEM across participants. The gray horizontal line marks the time window of the most significant cluster of the effect of decision for the mean duration.
(D) The behavioral results along with the model fit. Circles show the average percent of “long” responses across participants for each stimulus duration, and error bars depict the SEM. The solid curve shows the fit of the model to the data.
(E–H) Cognitive model of temporal decisions. The model includes a noisy accumulator (orange line, E and G) and a decision threshold (black horizontal line, E and G).
(E) The distance of the accumulator from the threshold for different stimulus durations. Blue shading shows the distribution of accumulator values at the stimulus offset for different durations. Red vertical lines correspond to the mean distance of the accumulator from the threshold.
(F) The distance from the decision threshold is the predicted offset response curve of the model.
(G) Accumulator dynamics during stimulus presentation. The thick blue line shows the mean value of the accumulator over time, and the blue shading shows the 95% highest density interval for accumulator values. The red curve (top) depicts the distribution of threshold-crossing times (for which the cumulative distribution appears in H).
(H) The psychometric curve produced by the model corresponds to the probability of the accumulator reaching the threshold by stimulus offset.

See also Figure S1.
and their interaction on the EEG. Offset responses were related to duration most clearly for “short” trials, as indicated by a significant interaction between duration and decision (p < 10^{-3}, cluster extending from 389 to 500 ms following offset). For “short” trials, shorter stimuli elicited larger offset responses (Figure 3A; duration effect, p < 10^{-4}, cluster extending from 334 to 500 ms following offset). Last, for stimuli presented for the mean duration (1.3 s, compared with the average bisection point of 1.07 s), stimuli categorized as “short” were associated with a larger offset response compared with those categorized as “long.” This was supported by the significant effect of the participants’ decision (Figure 3C; p < 10^{-3}, cluster extending from 281 to 461 ms following offset). Because in this experiment the mean bisection point was farther from the mean duration, we complemented the GLM analysis with a direct comparison of the mean amplitude in the frontocentral electrodes 300–500 ms following stimulus offset between near-bisection trials categorized as “short” versus “long.” This was explored by the significant effect of the participants’ decision (Figure 3C; p < 10^{-3}, cluster extending from 281 to 461 ms following offset). Because in this experiment the mean bisection point was farther from the mean duration, we complemented the GLM analysis with a direct comparison of the mean amplitude in the frontocentral electrodes 300–500 ms following stimulus offset between near-bisection trials categorized as “short” versus “long” (i.e., tailored per participant as depicted in Figure 3C). This analysis corroborated the result of the GLM, with “short” categorized trials associated with a larger amplitude (paired t test, t_{28} = 4.23, p < 10^{-3}). The effects found using the GLM were evident in a time range similar to that found in experiment 1, with a similar topographical distribution (Figure S2). Hence, we focused on the same epoch and electrodes when evaluating the fit of the cognitive model to these data. The full report of the spatiotemporal extent of significant clusters is described in the supplemental material (Figure S2).

We fitted the cognitive model parameters to the average behavioral data and used these to predict the average EEG responses. Here again, the EEG pattern was reproduced by the model (R^2 = 0.96; Figure 3B; see parameter estimates in Table S1). In summary, by replicating our previous findings in a tactile experiment, we show that the EEG response to stimulus offset reflects a temporal-evidence accumulation process that is not specific to the visual system and is therefore, potentially, modality independent.

Finally, to convincingly state that the processes delineated in the EEG offset response index accumulated evidence to a decision threshold rather than the specific, physical durations used in this task, we next set out to demonstrate that the same results are found when a task consists of a different stimulus set with different durations. In other words, we asked whether the offset response scales to reflect accumulation-to-threshold under different temporal contexts. The behavioral performance (summarized by the psychometric curve) of participants will change to accommodate the different temporal contexts in this task. Under the model we focus on in this work, this is explained by an adaptation of the accumulation rate.9

The accumulation index tracks relative rather than absolute duration

In experiment 3, individuals performed the same visual-temporal bisection task as in experiment 1, but they performed the task in two separate blocks. The blocks included stimuli in two time ranges: either 0.2–0.8 or 1–2 s. Block order was counterbalanced across participants, and each block was preceded by a familiarization stage. EEG data were acquired and handled with the exact same procedures as in the previous experiments. In this experiment, we assigned different hands to “long” and “short” responses, counterbalanced across participants, to control for motor preparation confounds in the EEG (STAR Methods).
Figure 4. EEG tracks accumulated evidence in short-duration (0.2–0.8 s) and long-duration (1–2 s) contexts (A and E) EEG time-locked to stimulus offset as a function of stimulus duration (colors, lines, shades, and significance as in Figure 2) for (A) short-duration and (E) long-duration contexts. (B and F) Average potential across the frontocentral electrodes in a time window of 300–500 ms post-offset, per stimulus duration, along with model predictions (short-duration block, $R^2 = 0.98$; long-duration block, $R^2 = 0.97$). The inset shows the scalp distribution of the cognitive model fits in the same time window. The frontocentral electrodes depicted in (A)–(C) and (E)–(G) are marked with white-filled circles. (C and G) Offset responses to the stimuli with durations closest to individual participant’s bisection points, separated by the participant’s decision. (D and H) The behavioral results along with the cognitive model fit. See also Figure S1.
Our model predicts that the offset response should track the relative distance of accumulated evidence to a decision threshold rather than absolute duration. Indeed, comparing offset responses shows that they are maximal to the shortest duration for both blocks (0.2 or 1 s; Figures 4A and 4E). As in the previous experiments, the GLM analysis revealed a clear effect of stimulus duration, participants’ decision, and their interaction on the offset EEG responses. A stronger relation between stimulus duration and offset responses is seen in “short” trials, as indicated by a significant interaction between duration and decision (short-duration block, p < 10^-3, cluster extending from 307 to 500 ms following offset; long-duration block, p < 0.03, cluster extending from 244 to 299 ms following offset). For “short” trials, shorter stimuli elicited larger offset responses (Figures 4A and 4E; duration effect, p < 10^-3 in both blocks, cluster extending from 357 to 500 ms and from 342 to 500 ms following offset for the short-duration and long-duration blocks, respectively). Last, for stimuli presented for the mean duration (short-duration block, 0.5 s compared with mean bisection point of 0.47 s; long-duration block, 1.5 s compared with mean bisection point of 1.49 s), stimuli categorized as “short” were associated with a larger offset response compared with those categorized as “long.” This was supported by the significant effect of the participants’ decision (Figures 4C and 4G; p < 10^-3 in both blocks, cluster extending from 328 to 496 ms and from 260 to 500 ms following offset for the short-duration and long-duration blocks, respectively). Note that offset responses for stimuli shorter than 500 ms (3 of the 7 durations in experiment 3a) contain onset responses to the same stimuli, most strongly seen in the epoch 0–300 ms after stimulus offset. However, the latter part of the offset response on which we test the cognitive model follows the same pattern seen in other experiments with longer intervals. The effects were evident in the same time range as in experiment 1, with a similar topographical distribution. The full report of the spatiotemporal extent of significant clusters is described in the supplemental material (Figure S2).

We fitted the model parameters to the average behavioral data and used these to predict the EEG responses in the same electrodes and epoch as in experiments 1 and 2. The model predictions matched the observed EEG offset responses (Figures 4B and 4F; short-duration block, R^2 = 0.98; long-duration block, R^2 = 0.97; for parameter estimates, see Table S1). In conclusion, if the offset response is interpreted as tracking temporal evidence, then experiment 3 supports another prediction: namely, that the offset responses will scale according to the set of durations used in the task.

DISCUSSION

We used a temporal bisection task in three different experiments. Participants had to categorize stimuli presented at varied durations as either “long” or “short.” Experiment 1 presented visual stimulation. Experiment 2 presented tactile stimulation using identical timing parameters as in experiment 1. Experiment 3 compared two different temporal ranges, sub- and supra-second durations, in the visual modality. We found that the EEG responses to stimulus offset reflect the distance of accumulated evidence toward a decision threshold in different sensory modalities and for different temporal ranges: stimuli categorized as “long” elicit an equal response regardless of duration, while for “short” stimuli, shorter durations elicit larger responses. Accordingly, stimuli of identical duration elicit larger responses if they are categorized as “short.” We also find evidence that participants with short bisection points exhibit EEG curves that reach a plateau at earlier durations (Figure S4; STAR Methods).

Late ERP components (e.g., P300) have previously been implicated in research on surprisal and prediction.19 Our findings can also be interpreted as a surprisal response to the stimulus offset, which is inversely related to the hazard rate.20 Future experiments could directly assess the separate contributions of the hazard rate and temporal evidence to the offset response by manipulating the distribution of the intervals.

The current report adds to the growing literature about EEG signatures that track interval timing,21 including contingent negative variation (CNV),22 lateralized readiness potential (LRP),23 and beta oscillations.24,25 Recently, offset responses have gained interest as signatures that are strongly modulated in interval timing experiments.26,27–31 Two responses have been highlighted, an early offset response, which we briefly describe in Figure S2, and a late offset response, which is the focus of this work. By focusing on evidence accumulation, which previous work relates to late ERPs,32 our findings contribute a direct link between EEG, behavior, and specific underlying cognitive processes and computations.

At the neural network level, previous work explored population clocks33 and low-dimensional trajectories34 as potential representations of elapsed time. Both can be conceptualized in terms of a process evolving toward a set point, similar to a bounded accumulation process. Thus, our findings can be viewed as a unidimensional representation of the complex activity patterns seen at the levels of neural networks and provide a putative link between human cognitive neuroscience and animal physiology. This link supports a common framework for the study of temporal cognition across species and at different levels of description in the neural system.

**STAR METHODS**

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Experiments 1 and 3 - Visual stimuli
  - Experiment 2 - Tactile stimuli
  - Design and procedure
  - Summary of experimental design and procedures for all experiments
- **EEG acquisition**
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - EEG preprocessing
  - General linear model
  - Cognitive process model
Fitting the model to behavior and simulating the EEG response
Estimating the correlation between the EEG and psychometric curves at the level of single participants
Motor preparation control analysis
Response time control analysis

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.cub.2022.08.006.

ACKNOWLEDGMENTS
The authors would like to thank Omri Manor, Gal Moscona, Noa Itzhaki, Segev Feinstein, Yael Gordon, and Gal Samuel for assistance in data acquisition. We thank Lorenzo Pisoni for the development of the tactile stimulation system. We thank Dr. Flor Kusnir, Maya Inbar, and Dr. Shany Grossman for providing meaningful input on the work and manuscript. We thank Prof. Leon Deouell (HCNIL) for making his laboratory available for part of the data acquired for this report. The Brain Attention and Time Lab (PI: A.N.L.) is supported by the James McDonnell Scholar Award in Understanding Human Cognition, ISF grant 958/16. This project has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 852387).

AUTHOR CONTRIBUTIONS
Conceptualization, A.N.L. and N.O.; methodology, A.N.L. and N.O.; software programming, N.O.; formal analysis, N.O.; investigation, N.O.; data curation, N.O.; writing – original draft, A.N.L. and N.O.; writing – review & editing, A.N.L. and N.O.; visualization, A.N.L. and N.O.; supervision, A.N.L.; funding acquisition, A.N.L.

DECLARATION OF INTERESTS
The authors declare no competing interests.

INCLUSION AND DIVERSITY
We worked to ensure gender balance in the recruitment of human subjects. While citing references scientifically relevant for this work, we also actively worked to promote gender balance in our reference list.

REFERENCES
1. Treisman, M. (1963). Temporal discrimination and the indifference interval. Implications for a model of the “internal clock”. Psychol. Monogr. 77, 1–31.
2. Gibbon, J. (1977). Scalar expectancy theory and Weber’s law in animal timing. Psychol. Rev. 84, 279–325.
3. Ivry, R.B., and Schneider, L.J. (2008). Dedicated and intrinsic models of time perception. Trends Cogn. Sci. 12, 273–280.
4. Bueti, D. (2011). The sensory representation of time. Front. Integr. Neurosci. 5, 34.
5. Simen, P., Rivest, F., Ludvig, E.A., Balci, F., and Killeen, P. (2013). Timescale invariance in the pacemaker-accumulator family of timing models. Timing Time Percept. 1, 159–188.
6. Gronlund, S. (2010). Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. Atten. Percept. Psychophys. 72, 561–582.
7. Balci, F., and Simen, P. (2014). Decision processes in temporal discrimination. Acta Psychol. 149, 157–168.
8. Simen, P., Balci, F., de Souza, L., Cohen, J.D., and Holmes, P. (2011). A model of interval timing by neural integration. J. Neurosci. 31, 9238–9253.
9. Balci, F., and Simen, P. (2016). A decision model of timing. Curr. Opin. Behav. Sci. 8, 94–101.
10. Merchant, H., Harrington, D.L., and Meck, W.H. (2013). Neural basis of the perception and estimation of time. Annu. Rev. Neurosci. 36, 313–336.
11. Paton, J.J., and Buonomano, D.V. (2018). The neural basis of timing: distributed mechanisms for diverse functions. Neuron 98, 687–705.
12. de Cheveigne, A., and Neißen, I. (2019). Filters: when, why, and how (not) to use them. Neuron 102, 280–293.
13. Pernet, C.R., Chauveau, N., Gaspar, C., and Rousselet, G.A. (2011). LIMO EEG: a toolbox for hierarchical linear modeling of electroencephalographic data. Comput. Intell. Neurosci. 2011, 814109.
14. Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. J. Neurosci. Methods 164, 177–190.
15. Pernet, C.R., Latinius, M., Nichols, T.E., and Rousselet, G.A. (2015). Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: a simulation study. J. Neurosci. Methods 250, 85–93.
16. Machado, A., Malheiro, M.T., and Erhagen, W. (2009). Learning to time: a perspective. J. Exp. Anal. Behav. 92, 423–458.
17. Klapproth, F., and Wearden, J.H. (2011). Why do temporal generalization gradients change when people make decisions as quickly as possible? Q. J. Exp. Psychol. 64, 1646–1664.
18. Wilson, R.C., and Niv, Y. (2015). Is model fitting necessary for model-based fMRI? PLoS Comput. Biol. 11, e1004237.
19. Verleger, R. (2020). Effects of relevance and response frequency on P3b amplitudes: review of findings and comparison of hypotheses about the process reflected by P3b. Psychophysiology 57, e13542.
20. Nobre, A.C., Correa, A., and Coull, J. (2007). The hazards of time. Curr. Opin. Neurobiol. 17, 465–470.
21. Kononowicz, T.W., van Rijn, H., and Meck, W.H. (2018). Timing and time perception: a critical review of neural timing signatures before, during, and after the to-be-timed interval. In Stevens’ Handbook of Experimental Psychology and Cognitive Neuroscience (Wiley), pp. 1–38.
22. Pflüger, M., Flotzar, R., and Pouthas, V. (2005). Relationship between CNV and timing of an upcoming event. Neurosci. Lett. 382, 106–111.
23. Voelber, G., and Thomaschke, R. (2017). Time-based expectations entail preparatory motor activity. Cortex 92, 261–270.
24. Kononowicz, T.W., and van Rijn, H. (2015). Single trial beta oscillations index time estimation. Neuropsychologia 75, 381–389.
25. Wiener, M., Parikh, A., Krakow, A., and Coslett, H.B. (2018). An intrinsic role of beta oscillations in memory for time estimation. Sci. Rep. 8, 7992.
26. Bueno, F.D., and Cravo, A.M. (2021). Post-interval EEG activity is related to task-goals in temporal discrimination. PLoS One 16, e0257378.
27. Kononowicz, T.W., and van Rijn, H. (2014). Decoupling interval timing and climbing neural activity: a dissociation between CNV and N1P2 amplitudes. J. Neurosci. 34, 2931–2939.
28. Wiener, M., and Thompson, J.C. (2015). Repetition enhancement and memory effects for duration. Neuroimage 113, 268–278.
29. Bannier, D., Wearden, J., Le Dantec, C.C., and Rebaï, M. (2018). Differences in the temporal processing between identification and categorization of durations: a behavioral and ERP study. Behav. Brain Res. 356, 197–203.
30. Kruizjé, W., Olivers, C.N.L., and van Rijn, H. (2021). Neural repetition suppression modulates time perception: evidence from electrophysiology and pupillometry. J. Cogn. Neurosci. 33, 1230–1252.
31. Silvestrini, M., Claessens, P.M.E., and Cravo, A.M. (2022). Temporal learning modulates post-interval ERPs in a categorization task with hidden reference durations. Preprint at bioRxiv. https://doi.org/10.1101/2022.01.25.477745.
32. O’Connell, R.G., and Kelly, S.P. (2021). Neurophysiology of human perceptual decision-making. Annu. Rev. Neurosci. 44, 495–516.
33. Zhou, S., Masmanidis, S.C., and Buonomano, D.V. (2020). Neural sequences as an optimal dynamical regime for the readout of time. Neuron 108, 651–658.e5.
34. Wang, J., Narain, D., Hosseini, E.A., and Jazayeri, M. (2018). Flexible timing by temporal scaling of cortical responses. Nat. Neurosci. 21, 102–110.
35. Kleiner, M., David, B., and Pelli, D. (2007). What’s new in Psychtoolbox-3? Percept. ECVP Abstr Suppl. 36, 437–442.
36. Mathôt, S., Schreij, D., and Theeuwes, J. (2012). OpenSesame: an open-source, graphical experiment builder for the social sciences. Behav. Res. Methods 44, 314–324.
37. Prins, N., and Kingdom, F.A.A. (2018). Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. Front. Psychol. 9, 1250.
38. Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J.M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 156869.
39. Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134, 9–21.
40. de Cheveigné, A., and Arzounian, D. (2018). Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data. Neuroimage 172, 903–912.
41. O’Connell, R.G., Dockree, P.M., and Kelly, S.P. (2012). A supramodal accumulation-to-bound signal that determines perceptual decisions in humans. Nat. Neurosci. 15, 1729–1735.
42. Higham, D.J. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. SIAM Rev. 43, 525–546.
43. Schurger, A., Sitt, J.D., and Dehaene, S. (2012). An accumulator model for spontaneous neural activity prior to self-initiated movement. Proc. Natl. Acad. Sci. USA 109, E2904–E2913.
KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Deposited data      |        |            |
| Raw behavioral data | This paper | https://www.doi.org/10.17605/OSF.IO/MRA6D |
| Processed EEG data  | This paper | https://www.doi.org/10.17605/OSF.IO/MRA6D |
| GLM results         | This paper | https://www.doi.org/10.17605/OSF.IO/MRA6D |

Software and algorithms

| SOFTWARE or ALGORITHM | SOURCE                        | IDENTIFIER |
|-----------------------|-------------------------------|------------|
| MATLAB 2018b          | The MathWorks                 | https://www.mathworks.com/; RRID: SCR_001622 |
| Experiment code       | This paper                    | https://www.doi.org/10.17605/OSF.IO/MRA6D |
| Analysis code         | This paper                    | https://www.doi.org/10.17605/OSF.IO/MRA6D |
| Psychotoolbox          | Kleiner et al.35              | http://psychtoolbox.org/; RRID: SCR_002881 |
| OpenSesame             | Mathôt et al.36               | https://osdoc.cogsci.nl/; RRID: SCR_002849 |
| Palamedes toolbox     | Prins and Kingdom37           | https://www.palamedestoolbox.org/; RRID: SCR_006521 |
| Fieldtrip toolbox     | Oostenveld et al.38           | https://www.fieldtriptoolbox.org/; RRID: SCR_004849 |
| LIMO EEG               | Pernet et al.13               | https://limo-eeg-toolbox.github.io/limo_meeg/; RRID: SCR_009592 |
| EEGLAB                 | Delorme and Makeig39          | https://sccn.ucsd.edu/eeglab/index.php; RRID: SCR_007292 |

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Ayelet N. Landau (ayelet.landau@mail.huji.ac.il).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- De-identified human data (behavior, processed EEG and GLM results) have been deposited at OSF. They are publicly available as of the date of publication. DOIs are listed in the key resources table.
- All original code for running the experiments as well as the analysis reported in this paper has been deposited at OSF and is publicly available as of the date of publication. DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

A total of 116 individuals participated in a total of three experiments. Experiment 1 n=45 (31 women, average age = 22.9 (SD 2.7)), experiment 2 n=40 (23 women, average age = 25 (SD 4.2)) and experiment 3 n=31 (13 women, average age = 24 (SD 3.1)). Participants were recruited from the university community and were compensated for their time with either money (10 euro per hour) or class credit. All procedures were approved by the institutional review board of ethical conduct. 2 subjects were removed from experiment 2 after collection due to technical issues in the EEG recording: One had faulty triggers and the other had high levels of noise in M1, which is used as one of the reference electrodes.

METHOD DETAILS

Three experiments are included in this report. The basic structure of the experiments was identical and therefore, for most, methods and task instructions are shared unless otherwise specified.

Experiments 1 and 3 - Visual stimuli

Visual stimuli consisted of a square-wave grating presented in a circular window on a BenQ XL2420Z monitor running on 144 Hz (experiment 1) or 100 Hz (experiment 3) using Psychtoolbox in Matlab.35 The grating had a spatial frequency of 3 cycles per visual field.40
degree, a diameter of 8° visual degree and was positioned at the center of the screen. During the experiment, stimuli were presented for different durations (see design and procedure) at two different levels of contrast, 100% and 50%. To ameliorate adaptation to the stimuli, gratings were presented randomly with a tilt of 45° or 135°, and a phase of 0°, 90° or 180°.

**Experiment 2 - Tactile stimuli**

Stimulation was produced with a vibrotactile coin stimulator connected to an open-source hardware, Arduino (Uno Rev3), programmed with C++ on a compatible IDE. The experiment was built and run on OpenSesame v3.1.26 The vibration produced by the Arduino had a frequency of approximately 120 Hz. Headphones were used to administer white noise throughout the experiment, in order to prevent participants from hearing the vibration. Vibration was presented to the left hand, to either the index finger, the middle finger or both fingers leaving the right hand free to deliver responses with a keypress. In addition, two different levels of vibration-intensity were presented.

**Design and procedure**

In experiment 1 there were twelve levels of stimulus duration in the test phase (500, 700, 900, 1050, 1150, 1250, 1350, 1450, 1550, 1700, 1900 and 2100 milliseconds). To make sure that no reliable associations can be formed between stimulus duration and its phase and orientation, all combinations were presented during both the familiarization and test phases, while only one contrast level was used to prevent participants from forming separate internal references for the different contrast levels. The training contrast level was counterbalanced across participants. The familiarization phase included 12 trials per duration. The test phase included a total of 504 trials, with 48 trials for durations 1050-1550 msec each, and 36 trials for each of the other durations. During the test phase, a break was given to the participants every 126 trials (every ~10 minutes). A red fixation point was displayed at the center of the screen (atop the gratings) during the entire experiment, except breaks. For a trial to start, participants had to fixate within a 1.5° radius of the fixation dot for a continuous second.

In experiment 2 we used the same twelve levels of stimulus duration in the test phase (500, 700, 900, 1050, 1150, 1250, 1350, 1450, 1550, 1700, 1900 and 2100 milliseconds). During the familiarization phase, the short reference was always presented with the weak intensity, while the long reference was always presented with the strong intensity. The familiarization phase included 12 trials per duration. The test phase included a total of 540 trials, with 40 trials for durations 1050-1550 msec each, and 30 trials for each of the other durations. During the test phase, a break was given to the participants every 108 trials (every ~10 minutes). A white fixation cross was displayed at the center of the screen throughout the entire experiment, except during breaks.

Experiment 3 was divided into two parts, where each part had the same structure, but a different set of durations. Both sets included 7 levels. The short durations included [200, 300, 400, 500, 600, 700, 800] milliseconds. The long durations included [1000, 1170, 1330, 1500, 1670, 1830, 2000] milliseconds. To ensure that no reliable associations could be formed between stimulus duration and its phase and orientation, all combinations were presented during both the familiarization and test phases, while only one contrast level was used to prevent participants from forming separate internal references for the different contrast levels. The training contrast level was counterbalanced across participants. Each participant completed two blocks, one for each duration set, and the order of the blocks was counterbalanced across participants. The familiarization phase included 12 trials per duration. The test phase included a total of 420 trials, with 40 trials for each duration. During the test phase, a break was given to the participants every 105 trials (every ~10 minutes). A red fixation point was displayed at the center of the screen (atop the gratings) throughout the entire experiment, except during breaks. For a trial to start, participants had to fixate within a 1.5° radius of the fixation dot for a continuous second.

In all experiments, participants could only respond once the stimulus was over.

**Summary of experimental design and procedures for all experiments**

| Experiment | Modality | Stimulus | Duration range (ms) | Sample size |
|------------|----------|----------|---------------------|-------------|
| 1          | Visual   | Grating  | 500 - 2100          | 45          |
| 2          | Tactile  | Vibration| 500 - 2100          | 31          |
| 3a         | Visual   | Grating  | 200 - 800           | 40          |
| 3b         | Visual   | Grating  | 1000 - 2000         |             |

**EEG acquisition**

We recorded the EEG of participants while they performed the temporal bisection task described above. For all participants but 10 in experiment 2 (tactile) we used a g.GAMMACap (gTec, Austria) and a g.Hiamp amplifier (gTec, Austria). For the other 10 we used an Active-Two acquisition system (Biosemi, the Netherlands). The gTec participants had a cap with 62 active electrodes distributed over the scalp, with the addition of two active earlobe electrodes. The Biosemi participants had a cap with 64 active electrodes distributed over the scalp. In both cases, electrodes were positioned according to the extended 10–20 system. We removed electrodes Iz, P9 and P10 from the Biosemi group prior to analysis because it did not have a counterpart in the gTec cap, and removed F9 and F10 electrodes from the gTec participants, as they often included strong muscle activity and were far from regions of interest. In addition,
for all participants we recorded the horizontal electrooculogram (EOG) using passive (gTec) or active (Biosemi) electrodes placed at the outer canthi of both eyes and the vertical EOG using electrodes placed above and below the left eye. The EEG was continuously sampled at 512 Hz. During both visual experiments (1 and 3), we monitored the eye position using an infrared EyeLink camera (SR Research, Canada), sampling at 1000 Hz. The EyeLink signal and the EEG signal were time-aligned and stored for offline-analysis using a Simulink model (MathWorks, MA).

### QUANTIFICATION AND STATISTICAL ANALYSIS

**EEG preprocessing**

The EEG was referenced offline to the average of the earlobes for the gTec participants, or to the average of the mastoids for the Biosemi group, as these references are close to each other, and were used previously. All offline preprocessing and analyses were done using a combination of FieldTrip, EEGLAB, LIMO and custom Matlab code. Bad electrodes were removed by visual inspection. On average, 0.70% of electrodes per participant were removed in experiment 1 (max 5%), 1.95% in experiment 2 (max 10%) and 0.75% in experiment 3 (max 6.67%). 6 participants had bad electrodes within the electrodes of interest, 2 per experiment. Slow drifts in the EEG were removed in segments between breaks using the baseline removal function msbackadj(), part of the Matlab Bioinformatics toolbox. This function uses a moving window to estimate the baseline value within each window. The baseline values are then interpolated with a spline, and the resulting spline is removed from the original data. The window size was 4 seconds, with a step of 0.75 second between two consecutive windows, and the median of each window was taken as the baseline value. This method is very close to recent robust detrending approaches. Although both methods aim to avoid artifacts arising from high-pass filtering of the data, the spline baseline removal has an additional advantage that its required parameters can be defined based on the experiment’s design. In contrast, the robust detrending method necessitates choosing artifact threshold and polynomial order, which are by nature a more subjective choice. Trials were defined from 500 msec before stimulus onset to 500 msec after the participant responded. Artifactual trials were removed using visual inspection guided by the summary statistics in the ft_rejectvisual() function. In addition, trials in which participants responded later than 5 seconds after stimulus-offset were removed from the analysis. On average, 3.26% of trials per participant were removed in experiment 1 (max 9.44%), 5.53% in experiment 2 (max 19.63%) and 2.32% in experiment 3 (max 5.83%). The larger number of rejected trials in experiment 2 was a result of instructing the participants that they could withhold responses when they wanted a break. 4 participants had more than 10% of their reaction times longer than 5 seconds. After these trials were rejected, we ran ICA on data, from which we visually selected components corresponding to eye movements and blinks and removed them. Finally, we interpolated rejected electrodes using spline interpolation as implemented in ft_channelrepair().

Within these trials, we focused on two windows of interest: The EEG response to the stimulus-offset, and the EEG response leading up to the keypress, indicating the participants’ temporal decision. As participants responded at variable intervals following the stimulus-offset, these 2 windows contain partially overlapping data. The EEG was baseline to a 100-msec window around stimulus-offset, 50 msec prior to 50 msec after offset prior to the extraction of both windows. This baseline was chosen to minimize the impact of slow varying fluctuations, such as the CNV, in the signal of interest. The full intervals, from stimulus onset to offset, are shown in Figure S3. In this report we focus on the stimulus-offset response, but all analyses were also performed on the keypress-locked potentials, which provided equivalent results (Figure S1).

This data was used for the GLM analysis (See general linear model). To evaluate how well the cognitive model fit the EEG, we averaged the potential for electrodes F1, Fz, F2, FC1, FCz, FC2, C1, Cz and C2, and during 300-500 msec after stimulus-offset. The data were first averaged within participants, and the results were then averaged across participants. The model was fit to these averages (See cognitive process model). This window of interest was supported by the statistical model in experiment 1 and is consistent with previous experiments. Although small variations in the location and specific epoch of significant clusters were seen from one experiment to the next, we decided to select a fixed subset of electrodes and a reasonably wide time window. A more tailored choice of parameter combination would likely favorably affect statistical and fitting procedures, but we found the data were sufficiently robust to apply this conservative choice without risking cherry picking the data. We carried out a post-hoc analysis to comprehensively explore the spatiotemporal extent of the effect. In this analysis we fitted the model to potentials from single electrodes in non-overlapping time windows of 10 msec, spanning 0-500 msec after stimulus-offset (Figure S2).

Additionally, as decision signals in the EEG are often analyzed with respect to participants’ response (keypress in our design), we repeated the cognitive model fits for the response-locked signals as well. For this, we selected electrodes CP1, CPz and CP241 and a time window of the last 100 msec prior to a keypress.

We applied an FIR windowed sync zero-phase lowpass filter for visualization purposes only, with 30 Hz as the cut-off frequency. The data visualizing the effect of decision (“short” vs. “long” trials for the same stimulus duration, Figures 2C, 3C, 4C, and 4G) was generated as follows. First, the bisection point of each subject was found by fitting their behavioral data with the cognitive model described below. We then took the stimulus that was closest to the bisection point for each participant and extracted the offset response for those stimuli. Finally, we averaged the offset responses within participants and then across participants for trials that were categorized as “short” and “long”, separately.
General linear model

We analyzed both windows of the EEG (500 ms following stimulus-offset and 500 ms leading to participants’ keypress) using the hierarchical mass univariate general linear modeling approach implemented in LIMO.13 At the first level, we modeled the EEG amplitude at each timepoint and electrode combination per participant using the following formula:

\[ EEG = \beta_0 + \beta_1 \cdot decision_{short} + \beta_2 \cdot decision_{long} + \beta_3 \cdot duration \cdot decision_{short} + \beta_4 \cdot duration \cdot decision_{long} + \beta_5 \cdot RT \]

This model includes a predictor for “short” vs. “long” responses, a continuous, linear effect for duration, with a different slope depending on the choice the participant made in that trial, and a continuous linear effect for response time (see further controls on the effect of response times in the “response time control analysis” subsection). The predictor decision_{short} was 1 for all trials the participant categorized as “short”, and 0 otherwise. The predictor decision_{long} was 1 for all trials the participant categorized as “long”, and 0 otherwise. This type of categorical variables coding is sometimes called singular parametrization, because it leads to singular design matrices. This is solved in one of two ways: Either by changing the design matrix inversion algorithm to a pseudo-inverse (as is done in LIMO), or by estimating contrasts rather than the parameters themselves (as is done, for example, when using effect coding in typical ANOVAs). However, using one way or another does not change the results of the tests.13 Stimulus durations were standardized per participant (and temporal context, in experiment 3), such that duration 0 is the mean duration for trials in that temporal context. The model parameters, then, have the following interpretation:

1. \( \beta_1 \) – the effect of a “short” decision at the mean duration.
2. \( \beta_2 \) – the effect of a “long” decision at the mean duration.
3. \( \beta_3 \) – the effect of increasing the duration of the stimulus by 1 standard deviation for trials that were categorized as “short.”
4. \( \beta_4 \) – the effect of increasing the duration of the stimulus by 1 standard deviation for trials that were categorized as “long.”
5. \( \beta_5 \) – the effect of increasing the response time by 1 second.
6. \( \beta_0 \) – the predicted EEG amplitude after regressing out all other effects.

In total, 60 X 257 (15,420, one for each electrode and timepoint combination) models were fitted to the data of each participant. The model parameters were estimated using a weighted least squares method for each participant separately. After the model parameters were estimated for each timepoint and electrode for each participant, we computed two contrasts (linear combinations of parameters that capture comparisons of interest). The main contrast of interest was the interaction, which we computed as the difference \( \beta_3 - \beta_4 \) (The difference between the slope of the regression against stimulus duration for “short” vs. “long” decisions). To capture the effect of decision we computed the difference \( \beta_2 - \beta_1 \) (“long” minus “short” responses when the stimulus duration equals the mean duration for that dataset). Last, since the interaction term was always significant, there is no simple contrast that would quantify the overall effect of stimulus duration. Instead, we chose to test the effect of stimulus duration separately for “short” \( \beta_3 \) and “long” \( \beta_4 \) trials. As the contrasts were computed at each electrode and timepoint for each participant, a single participant is represented by 4 spatiotemporal 60 X 257 matrices, one per effect. Every element in each matrix, that is an electrode-timepoint pair, is a single observation in the statistical significance test of a specific effect.

Significance testing also followed LIMO conventions.13 The significance of the contrasts is estimated at the second (or group) level, using a clustering approach to correct for multiple comparisons. First, the contrasts are transformed into robust t-values: At each element of each matrix, the 80% trimmed mean is calculated across participants and then divided by the degrees of freedom. Then, clusters were formed by thresholding the single observations at the t value corresponding to an uncorrected alpha of 0.95 and summing neighboring observations. Next, we calculated the significance of the clusters using bootstrap. First, the trimmed mean t-value at each observation was removed from the result of each participant, to shift the observations to the null distribution. Then, participants were resampled with replacement, and transformed to robust t-values. Next, clusters were formed at each iteration of the bootstrap procedure. The summed absolute t-values of the largest cluster at each iteration were taken to form the distribution of clusters under the null hypothesis, and the observed clusters were compared to this distribution to calculate their p value.

For the two contrasts (the interaction and the decision effect) we used an alpha of 0.05. For the effects of duration in “short” and “long” trials we used an alpha of 0.025, as they were post-hoc given the significant interaction.

Cognitive process model

The cognitive process model we use in this paper is an adaptation of a drift-diffusion model (DDM) proposed previously to account for behavior in temporal bisection.7 Although our model is closely related to this previous model, there are a few differences arising from our simplification. Our model includes a single DDM with a single threshold. At stimulus onset the drift-diffusion process starts. If it reaches the threshold during stimulus presentation, the stimulus is categorized as “long”. If it did not, the stimulus is categorized as “short”. In contrast, the original model includes two DDMs, one with a single threshold and the second with two thresholds, operating in tandem. The first drift-diffusion process starts at stimulus onset. If it reaches the threshold during stimulus presentation, the stimulus is categorized as “long”. Otherwise, the final value of the first process is compared to an internal comparison value. The difference from this internal value is used as the starting point and drift rate of the second DDM. The second process has two decision thresholds, representing either decision (“short” or “long”). The addition of the second DDM, and the numerous extra parameters...
it requires, is necessary to capture response times patterns in the task. However, as our focus is on explaining the EEG patterns, which we find to be generally unrelated to response times, we simplified the model to include a single drift-diffusion process.

The mathematical formulation of our model is as follows (exact formulae are from Simen et al.): A drift-diffusion process, representing the momentary perceived duration (accumulated temporal evidence), starts at the stimulus onset. The dynamics of the accumulated evidence is given by:

\[
dx = A \cdot dt + c \cdot dB
\]

Where \(x\) is the accumulated evidence, \(A\) is the drift rate (how fast accumulated evidence grows over time on average), \(c \cdot dB\) represents the variability in momentary perceived duration (the amount of noise in the accumulated evidence).

We additionally define a threshold \(z\), which is used to represent the decision threshold. If the accumulator reaches the threshold during stimulus presentation, the stimulus is categorized as “long,” and the accumulator terminates. Otherwise, the stimulus is categorized as “short.” The probability of the accumulator reaching the threshold after \(t\) seconds for the first time is described by an inverse Gaussian distribution:

\[
\rho(t, \eta) = \left(\frac{\eta}{2\pi t^3}\right)^{1/2} \exp\left(-\frac{\eta(t - \mu)^2}{2\mu^2 t}\right)
\]

\(\mu\) is the threshold to drift rate ratio \(\frac{\mu}{A}\) and \(\eta\) is the threshold to noise ratio \(\frac{z}{\sigma}\). The probability of categorizing a stimulus of duration \(t\) as “long” is equal to the probability of the accumulator reaching the threshold before this stimulus is over, i.e., the cumulative inverse Gaussian distribution with parameters \(\mu\) and \(\eta\) evaluated at \(t\). Calculating this quantity for all durations used in a temporal bisection experiment gives the psychometric function, which is described by 2 free parameters, \(\mu\) and \(\eta\).

Fitting the model to behavior and simulating the EEG response

We used the model to fit the behavioral data, both at the single participant and group levels. The two free parameters of the model (\(\mu\) and \(\eta\)) were estimated by a maximum likelihood procedure, using the simplex search algorithm implemented in the Palamedes toolbox. This procedure was done both for individual participants, to find the duration that was closest to the bisection point (the duration which the participant is predicted to categorize as “long” with a 0.5 probability), and for the pooled data of all participants.

We carried out an additional analysis, meant to reveal the EEG response curve of the model, and assess its qualitative fit to the grand average EEG data. For this we extracted the mean distance of the accumulator from the threshold over time. This was done by using the parameters that were estimated using the group behavioral data to simulate the accumulator 50000 times with the Euler-Maruyama method and a step size of 0.5 msec. As the behavior only allows fitting 2 parameters, while the full drift diffusion model requires 3 (drift rate, noise and threshold), we arbitrarily fixed the rate at 1. This means that the threshold in the drift diffusion model is equal to \(\mu\), and the noise is \(\frac{z}{\sigma}\). Each simulation started at time 0 and lasted until the longest duration for that dataset. After the traces were simulated, if a trace reached the threshold at any point, we replaced this and all following points with the threshold (implementing a sticky boundary, corresponding to decision that are made before stimulus offset). Next, we calculated the distance from the threshold for each trace and time point and averaged across traces to produce the predicted response for all durations. Finally, to fit the predicted responses to the observed EEG we extracted the points closest to the true durations, the distance from the threshold for each trace and time point and averaged across traces to produce the predicted response for all durations. Finally, to fit the predicted responses to the observed EEG we extracted the points closest to the true durations, and used them as predictors in a linear regression model, following previous examples. This is necessary to control for irrelevant differences in scale and baseline between the model (which produces responses in arbitrary units) and the EEG (which is measured in \(\muV\)). We took the \(R^2\) of the model as the goodness of fit. In addition to using the parameters values estimated using the behavioral results, we tested the ability of neighboring values of the parameters to fit the data and found that the behavioral parameters provide near-maximal fits within that range (Figure S4). We note that adding this second layer to the fit (correlating the simulated and observed EEG) makes the cognitive parameters less sensitive, as changes in their values can often be compensated by changing the scale and baseline.

For both the behavioral parameters and the additional offset and scale parameters used to fit the behavior to the EEG, we estimated the standard deviation of fits using a bootstrap approach. First, we resampled the participants with repetition to obtain a sample of the same size as the original dataset. We then estimated the behavioral parameters and the offset and scale parameters of this random pool of participants. This procedure was repeated 1000 times to create the bootstrap distributions, one for each of the four parameters. The standard deviation of each distribution was taken as the estimate of parameter spread.

Estimating the correlation between the EEG and psychometric curves at the level of single participants

The cognitive model described in this work predicts that participants with shorter bisection points (i.e. with a smaller ratio between their decision threshold and accumulation rate) should exhibit curves that saturate at shorter durations compared to participants with longer bisection points. Estimating the cognitive model parameters by fitting the EEG requires additional scale and baseline parameters and therefore is indirect. These added parameters trade-off strongly with the cognitive parameters themselves (i.e., increasing the cognitive parameter related to slope can be compensated by decreasing the scale parameter of the EEG). This trade-off is also known from previous work about fitting reinforcement learning models to BOLD signals. This challenge can be circumvented by quantifying the duration at which the EEG response reaches a plateau directly. We did this by fitting a piecewise linear function
with two lines: one line corresponding to the initial decrease in EEG amplitude, and a second line with a fixed slope of 0. This model has 3 free parameters (Initial ramp slope, knee and a general baseline). As can be seen in Figure S4, participants with shorter bisection points have EEG curves that saturate at shorter durations and vice versa (Figures S4B and S4C). We used the bisection points of individual participants to divide them into three tertiles. We then took the lower and upper tertile and calculated the mean psychometric curve and EEG curve for each group. In addition, for all experiments, and all participants we correlated the estimated knee in the EEG with the behavioral-bisection point. This analysis revealed a significant correlation between the EEG knee and the behavioral bisection point for experiments 1 and 2 (Figure S4D). Spearman correlations: Experiment 1: rho = 0.44, p < 3 \times 10^{-3}, experiment 2: rho = 0.47, p < 2 \times 10^{-2}, experiment 3a: rho = 0.02, p = 0.91, experiment 3b: rho = -0.03, p = 0.84. We attribute the difference between experiments 1 and 2 versus 3a and 3b we found smaller bisection point variability between participants. Second, those experiments used 7 durations compared to the 12 durations tested in Experiments 1 and 2. Sampling less time points results in a poorer sensitivity to potential inter-subject variability in the data.

Motor preparation control analysis
Participants in experiment 3 used different hands to report their categorization decisions, to control for the involvement of motor preparation in the EEG responses. Half of the participants used their left hand to report stimuli as “long,” and their right hand to report stimuli as “short,” and half used the opposite response mapping. Fitting the cognitive model to the offset responses of each mapping group separately suggests that the offset response documented in this paper is not impacted by motor preparation. A reanalysis of the participants grouped by response mapping led to the same pattern of results (short-durations block: Right hand “long” $R^2 = 0.9$, left hand “long” $R^2 = 0.89$, long-durations block: Right hand “long” $R^2 = 0.96$, left hand “long” $R^2 = 0.97$).

Response time control analysis
Since response times in temporal bisection depend on the stimulus duration and participants’ responses, we carried out several analyses to control for its possible contribution to the EEG pattern reported in this paper. First, we point out that while the EEG amplitude is non-linearly monotonically decreasing (i.e., longer durations within a dataset are associated with smaller amplitudes, saturating around the bisection point (Figure S3). Second, we found that for durations close to the bisection point, “short” decision trials are associated with larger stimulus-offset response compared to “long” decision trials in all 4 experiments (as seen in Figures 2C, 3C, 4C, and 4G), while RTs are significantly different only in Experiments 1 and 3b (Experiment 1: Mean “short”-“long”: 77 msec, $t_{44}=-2.70$, p=0.001, Experiment 2: Mean “short”-“long”: 8 msec, $t_{28}=-0.13$, p=0.901, Experiment 3a: Mean “short”-“long”: 1 msec, $t_{39}=0.02$, p=0.981, Experiment 3b: Mean “short”-“long”: 72.9 msec, $t_{39}=-2.32$, p=0.026). Finally, the pattern of the EEG results remains the same, even when examining the keypress-locked ERP (Figure S1). Future experiments can control the effect of RT directly by instructing participants to withhold their response until an appropriately delayed cue.