

# UNDERSTANDING THE ROLE OF TIME-DIMENSION IN THE BRAIN INFORMATION PROCESSING

EDITED BY : Daya Shankar Gupta and Hugo Merchant  
PUBLISHED IN : Frontiers in Psychology





# frontiers

## Frontiers Copyright Statement

© Copyright 2007-2017 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-149-4

DOI 10.3389/978-2-88945-149-4

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [researchtopics@frontiersin.org](mailto:researchtopics@frontiersin.org)

# UNDERSTANDING THE ROLE OF TIME-DIMENSION IN THE BRAIN INFORMATION PROCESSING

Topic Editors:

**Daya Shankar Gupta**, Camden County College, USA

**Hugo Merchant**, Instituto de Neurobiología, UNAM, Mexico



Temporal Unit  
Image by Nikolai Sorokin

Optimized interaction of the brain with environment requires the four-dimensional representation of space-time in the neuronal circuits. Information processing is an important part of this interaction, which is critically dependent on time-dimension. Information processing has played an important role in the evolution of mammals, and has reached a level of critical importance in the lives of primates, particularly the humans.

The entanglement of time-dimension with information processing in the brain is not clearly understood at present. Time-dimension in physical world – the environment of an organism – can be represented by the interval of a pendulum swing (the cover page depicts temporal unit with the help of a swinging pendulum). Temporal units in neural processes are represented by regular activities of pacemaker neurons, tonic regular activities of proprioceptors and periodic fluctuations in the excitability of neurons underlying brain oscillations. Moreover, temporal units may be representationally associated with time-bins containing bits of information (see the Editorial), which may be studied to understand the entanglement of time-dimension with neural information processing.

The optimized interaction of the brain with environment requires the calibration of neural temporal units. Neural temporal units are calibrated as a result of feedback processes occurring during the interaction of an organism with environment.

Understanding the role of time-dimension in the brain information processing requires a multidisciplinary approach, which would include psychophysics, single cell studies and brain recordings. Although this Special Issue has helped us move forward on some fronts, including theoretical understanding of calibration of time-information in neural circuits, and the role of brain oscillations in timing functions and integration of asynchronous sensory information, further advancements are needed by developing correct computational tools to resolve the relationship between dynamic, hierarchical neural oscillatory structures that form during the brain's interaction with environment.

**Citation:** Gupta, D. S., Merchant, H., eds. (2017). Understanding the Role of Time-Dimension in the Brain Information Processing. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-149-4



# Table of Contents

- 05 Editorial: Understanding the Role of the Time Dimension in the Brain Information Processing**  
Daya S. Gupta and Hugo Merchant
- 09 In Search of Oscillatory Traces of the Internal Clock**  
Tadeusz W. Kononowicz and Virginie van Wassenhove
- 14 Unpredicted Pitch Modulates Beta Oscillatory Power during Rhythmic Entrainment to a Tone Sequence**  
Andrew Chang, Dan J. Bosnyak and Laurel J. Trainor
- 27 Large Scale Functional Brain Networks Underlying Temporal Integration of Audio-Visual Speech Perception: An EEG Study**  
G. Vinodh Kumar, Tamesh Halder, Amit K. Jaiswal, Abhishek Mukherjee, Dipanjan Roy and Arpan Banerjee
- 40 Modulation of Alpha and Beta Oscillations during an n-back Task with Varying Temporal Memory Load**  
Youguo Chen and Xiting Huang
- 50 Corticostriatal Field Potentials Are Modulated at Delta and Theta Frequencies during Interval-Timing Task in Rodents**  
Eric B. Emmons, Rafael N. Ruggiero, Ryan M. Kelley, Krystal L. Parker and Nandakumar S. Narayanan
- 59 Rubberband Effect in Temporal Control of Mismatch Negativity**  
Lingyan Wang, Xiaoxiong Lin, Bin Zhou, Ernst Pöppel and Yan Bao
- 69 Detecting Temporal Change in Dynamic Sounds: On the Role of Stimulus Duration, Speed, and Emotion**  
Annett Schirmer, Nicolas Escoffier, Xiaoqin Cheng, Yenju Feng and Trevor B. Penney
- 81 A Roving Dual-Presentation Simultaneity-Judgment Task to Estimate the Point of Subjective Simultaneity**  
Kielan Yarrow, Sian E. Martin, Steven Di Costa, Joshua A. Solomon and Derek H. Arnold
- 100 Monkeys Share the Human Ability to Internally Maintain a Temporal Rhythm**  
Otto García-Garibay, Jaime Cadena-Valencia, Hugo Merchant and Victor de Lafuente
- 112 Meta-Analysis of Functional Neuroimaging and Cognitive Control Studies in Schizophrenia: Preliminary Elucidation of a Core Dysfunctional Timing Network**  
Irene Alústiza, Joaquim Radua, Anton Albajes-Eizagirre, Manuel Domínguez, Enrique Aubá and Felipe Ortuño
- 124 Commentary: Effects of psilocybin on time perception and temporal control of behavior in humans**  
Katarina L. Shebloski and James M. Broadway
- 127 When and How-Long: A Unified Approach for Time Perception**  
Michail Maniadas and Panos Trahanias



# Editorial: Understanding the Role of the Time Dimension in the Brain Information Processing

Daya S. Gupta<sup>1\*</sup> and Hugo Merchant<sup>2\*</sup>

<sup>1</sup> Department of Biology, Camden County College, Blackwood, NJ, USA, <sup>2</sup> Department of Behavioral and Cognitive Neurobiology, Instituto de Neurobiología, UNAM, Querétaro, Mexico

**Keywords:** temporal processing of sensory information, neural oscillations, Schizophrenia, timing and time perception, sensory information processing

## Editorial on the Research Topic

### Understanding the Role of the Time Dimension in the Brain Information Processing

An accurate representation of time-dimension in the neuronal circuits is required for a successful interaction of the brain with the four-dimensional physical world. Time-dimension, unlike other three dimensions of our physical universe, is never perceived as a novelty, but only reported as the flow of time. As there are no known neurological or psychiatric disorders that are associated with the loss of the sense of flow of time, this suggests that the functions of the brain involve processing of temporal information (Merchant et al., 2013). Moreover, psychological flow of time is likely the result of the perception of the physical nature of the time-dimension.

The information about a stimulus coded by neural circuits can be understood in terms of Shannon information, which is the arrangement of spikes (an absence or presence) in time-bins of specific size along the time-dimension (Gupta and Chen, 2016a,b). Thus, Shannon information inherently incorporates time-bin as the time-dimension in information processing. Encoded stimulus characteristic, can be decoded or utilized in brain circuits by processing this information, referred as the temporal processing of information. Thus, it is implicit that the information processing, underlying various cognitive functions of the brain, is coupled with the invariant time-dimension.

Several novel findings are reported in this Special Issue, which bring us closer to understanding the role of the time-dimension in the brain information processing. These include the representation of the physical time in neural circuits, temporal processing of information, the role of prior information in the internal representation of rhythmic time, and neural oscillations in timing behavior and perception.

## BRAIN OSCILLATIONS IN TIMING BEHAVIOR AND PERCEPTION

Brain oscillations are a key element on information processing and play a crucial role in the communication between and within different cortical and subcortical areas (Buzsáki, 2006). Neural oscillations have been linked to different high cognitive functions of which timing and time perception constitute one of the most studied (Treisman, 1963; Matell and Meck, 2004; Gupta, 2014; Kononowicz and van Rijn, 2014; Merchant et al., 2015a).

Kononowicz and van Wassenhove in this issue, review different oscillatory models that explain how the brain may subserve interval timing in millisecond range, but they focus mainly on the Striatal Beat Frequency (SBF) model and the new evidence that supports it. In contrast to other models, the SBF is more biologically plausible and implies the existence of cortical oscillators of

## OPEN ACCESS

### Edited and reviewed by:

Rufin VanRullen,  
Paul Sabatier University, France

### \*Correspondence:

Daya S. Gupta  
dayagup@gmail.com  
Hugo Merchant  
hugomerchant@unam.mx

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 28 January 2017

**Accepted:** 07 February 2017

**Published:** 23 February 2017

### Citation:

Gupta DS and Merchant H (2017)  
Editorial: Understanding the Role of  
the Time Dimension in the Brain  
Information Processing.  
Front. Psychol. 8:240.  
doi: 10.3389/fpsyg.2017.00240

various frequencies. At the onset of a timed interval, cortical oscillators are phase-reset and, at the offset of the interval, the state of these cortical oscillators is read by the medium spiny neurons of the striatum. Due to the fact that spiny neurons are coincidence detectors of the cortical input, the output of these striatal neurons could in principle tell time. Now, the authors suggest that the cortical input on the spiny neurons could be any stable neural pattern including neural avalanches (Crowe et al., 2014; Merchant et al., 2015b) or population state activity (Merchant et al., 2014; Mello et al., 2015). The authors also emphasize the notion that oscillation modulation for the SBF should be phase-reset at the beginning of a time interval, and that the frequency of the oscillation should be modulated by dopaminergic agents. None of these properties have been properly tested yet. In addition, this opinion letter underlines the fact that recent studies have shown that beta not only alpha oscillations are deeply involved in prediction of events during rhythmic tasks. Last, this paper suggests that the coherence between cortical and striatal signals is a fundamental element for generating a realistic model that includes both structures.

The paper by Chang et al. investigates the changes in beta oscillations induced by unpredicted changes in pitch, during an oddball EEG experiment where human subjects passively listened to isochronous auditory sequences with occasional unpredicted deviant pitches. These researchers tested the notion that if the induced beta power only reflects predictive timing, the occasional unpredicted pitch changes should not affect the ongoing beta entrainment behavior, given that the pitch deviants are presented at the predicted rhythmic time points. In contrast, their experimental results indicate that induced (non-phase locked) beta power was modulated by the unpredicted deviant pitches, suggesting that beta power is associated with predictive perceptual processing for the stimuli what (the pitch) and when (the tempo of the isochronous beat). The authors interpret these findings as the evidence that predictions for what and when are dynamically processed through attentional networks, and that beta oscillations in auditory cortex reflect the functional significance of sensory prediction and prediction error processes.

Kumar et al. report findings from a study of the McGurk effect, in which semantically-incongruent visual information modulates auditory perception. Authors employed incongruent audio-visual (AV) pair (audio/*pa*/ superimposed on the video of the face articulating */ka/*) to induce the cross-modal percept/*ta*/. They find that for asynchronous AV (audio-visual) stimuli, a broadband enhancement, in the global coherence at theta, alpha, beta, and gamma bands, aids the cross-modal perception (percept/*ta*/. Long-range oscillations (alpha and beta bands) aid in the multidimensional processing of information of asynchronous AV stimuli by providing temporal window of integration, which is a specific phase of long-range oscillations (Gupta and Chen, 2016a). Asynchronous AV stimuli can be integrated in the same phase—temporal window of integration—of a long-range oscillation, which is possible due to the difference in the delays prior to their arrival at respective processing circuits (Gupta and Chen, 2016a). This difference in delays helps to eliminate the difference in the time of the initial presentation of asynchronous AV stimuli. In contrast to asynchronous AV

stimuli, the processing of synchronous AV stimuli is temporally coupled to same coordinates on the time-axis, and therefore, it does not require the coupling to the same phase of a long-range oscillation to achieve simultaneous processing. Consistent with this argument, Kumar et al. observed desynchronization at alpha and beta-bands for AV stimuli. However, long-range oscillations may interfere with the integration of synchronous AV stimuli. This study highlights the important differences in the multisensory integration of synchronous and asynchronous AV stimuli.

The next paper by Chen and Huang studied alpha and beta modulations in a temporal version of n-back working memory task. Their findings reveal that while posterior alpha band reflects inhibition of task-irrelevant information, temporal region-distributed beta band activity is important for the active maintenance of temporal duration in the working memory.

Finally, Emmons et al. investigated the LFP oscillatory changes in the medial prefrontal cortex and the striatum during an interval timing task, where rats produced a 3 or 12 s interval. The results showed significant changes in delta and/or theta bands in the two areas during the following epochs of the task: after the cue that signaled the beginning of the time interval, throughout the interval, prior to the response that define the end of the interval, and after reward delivery. These findings support the notion that oscillatory activity between both areas of the motor cortico-basal-ganglia-thalamo-cortical circuit (Merchant et al., 2015a) is engaged in the temporal control of action in rodents.

## MISMATCH NEGATIVITY (MMN): REPRESENTS MECHANISMS FOR EXTRACTING PHYSICAL TIME INFORMATION

MMN is an event-related potential (ERP) wave, which reflects neuronal processes underlying the brain's automatic reaction to novel or deviant as well as unattended sensory stimuli (Näätänen et al., 2007). In a work submitted to this Research Topic, Wang et al. extracted and correlated several different parameters—temporal parameters (onset, offset, and peak latency) and wave shape parameters (amplitude, average amplitude, upslope, downslope)—characterizing the MMN waves produced by deviant sound. Their results revealed only one important correlation: a positive correlation between the MMN amplitude and the slope of decaying phase, also called downslope (Wang et al.). The authors argue that this represents an efficient feedback process, which allows MMN to return to the baseline within a predefined time-window. This also suggests a coupling between the neuronal processes associated with deviant stimuli and a representation of the physical time-axis in the brain. This coupling may subserve the mechanism to input the physical time information into brain circuits, which would calibrate endogenous oscillators in a distributed modular clock model (Gupta, 2014).

In another study Schirmer et al., using MMN paradigm, deviant stimulus was created by subjecting one surprised and one neutrally spoken “Ah” to a speech manipulation procedure

creating a 378 ms (short) and a 600 ms (long) exemplar. In both emotional conditions, short or long exemplars were used as standard or deviant stimuli. When short exemplar served as the standard, long exemplar was used as the deviant stimulus, and the vice versa. Authors observed a MMN-like negativity—climbing negativity that plateaued. Greater negativity for deviants than standards emerged shortly after the deviant onset, before the standard or deviant duration had lapsed. This suggested that listeners implicitly tracked sound speed and detected speed changes (Schirmer et al.). Continuous detection of the changes in the speed of unattended sounds would play a role in the calibration of endogenous oscillators in modular clock mechanisms (Gupta, 2014), including those that play a role in speech production.

## SIMULTANEITY JUDGMENT OF TEMPORAL EVENTS

The paper by Yarrow et al. tested different paradigms to optimally determine the relative judgment of two or more simultaneous events. Specifically, they argue that the dual presentation simultaneity judgment (2 x SJ) task is the most desirable. In this task subjects are asked to discriminate which of two pairs of stimuli presented consecutively was the most synchronous. They develop an appropriate signal detection theory model to analyze the 2xSJ data, and finally, they compare the data from the novel task with more conventional simultaneity tasks. Compared to classical tasks such as the temporal order judgment task, the 2 x SJ provides more constrained estimates of sensory noise, which indicates a more straightforward decision process. In fact, the 2 x SJ requires explicitly to decide which alternative timing relationship is most synchronous on any given trial, rather than revealing what range of relationships are perceived as synchronous. Consequently, 2 x SJ will serve as a crucial complement to existing methods for investigating subjective timing.

## MONKEYS AND HUMANS SHARE THE ABILITY TO INTERNALLY MAINTAIN A TEMPORAL RHYTHM

García-Garibay et al. demonstrate the ability of the rhesus monkeys and humans to perceive and maintain rhythms of different pace in the absence of sensory cues or motor actions. They use a visuospatial task in which subjects observe and then internally track a visual stimulus that periodically changed its location along a circular path. The proportion of trials in which subjects correctly estimated the position of the stimulus, along with other variables were determined in this study. Both species showed variability, consistent with Weber Law, where time independent variability increased as a function of timed duration (Zarco et al., 2009). In a different version of this task tested in humans, which reveals patterns of timing errors, shows that subjects tend to lag in fast rhythms and to get ahead in slow ones. The authors argue that a mean tempo might be incorporated as prior information, helping to reduce the effect of

noise in time estimation and production tasks (García-Garibay et al.).

## ABNORMAL TEMPORAL PROCESSING OF INFORMATION: IN SCHIZOPHRENIA AND PSILOCYBIN-INDUCED STATES

A meta-analysis of functional MRI studies in schizophrenia, comparing brain structures, activated or inactivated by time perception task and increasing levels of cognitive difficulty, revealed bilateral overlapping of cortical and subcortical regions, particularly frontal areas (mainly right BA 6), as well as parietal regions and the basal ganglia (Alústiza et al.). The overlapping regions, which are primarily in the right hemisphere, showed reduced rather than increased activity in schizophrenic patients relative to control subjects, not only by time perception tasks but also by an increase in the level of difficulty of non-temporal tasks (Alústiza et al.). Reduced activity of various brain structures is consistent with the prevailing view that there is an impaired functional connectivity of brain regions in schizophrenia (Hutchison et al., 2013). Thus, dysconnectivity affects common networks in schizophrenia, which are engaged by both the increasing task difficulty and time perception tasks.

In a commentary, Shebloski and Broadway discuss a paper by Wittmann et al. (2007). The study by Wittmann et al. (2007) showed a decreased ability to accurately produce intervals longer than 3 s and synchronize finger-tapping to auditory beats separated by more than 2 s under the influence of psilocybin (Wittmann et al., 2007). The effects on timing performance were accompanied by working-memory deficits and subjective changes in conscious state. Shebloski and Broadway also noted that schizophrenia, which is associated with similar changes in subjective state, such as hallucinations, is also associated with timing deficits in sub- and supra-second range. They further point out that slowing of perceived time induced by psilocybin and schizophrenia may share certain common mechanisms, such as 5-HT<sub>2A</sub> receptor activities. Thus, Shebloski and Broadway propose that commonalities across pharmacological treatments and psychiatric disorders should be explored within a common experimental paradigm.

It should be also noted that in contrast to the effects of psilocybin administration, which are pharmacological, the pathophysiology underlying schizophrenia involves defects at many levels, such as circuit, molecular and morphological levels. Therefore, to interpret the results of experiments in terms of underlying pathophysiology will involve many challenges.

## STUDY OF MODULAR INTERACTIONS BETWEEN BRAIN REGIONS USING ARTIFICIAL SYSTEMS

Maniadakis and Trahanias test a model of artificial cognitive system, which has the ability to sense when events have occurred and how long they have lasted. Authors employ a set of neural networks in their model, to synthesize modules, similar to the modular parts of the human brain. Inspired by the striatal



beat frequency model of interval timing (Matell and Meck, 2004; Meck et al., 2008), authors incorporated a module in their artificial system, which transforms oscillatory inputs into a composite time flow representation.

Authors used a coevolutionary scheme (Maniadakis and Trahanias, 2008) to train the model, and improve the collaboration between component neural networks, forming modules. The coevolutionary procedure, after 500 generations, produced a modular system that memorizes the duration and time of occurrence of events. Such methods can be a useful computational tool for the study of modular interactions in brain networks.

Various papers in this special issue describe that modulations of beta-range oscillations play an important role in the timing behavior. Beta power increased as working memory load increased in a temporal version of n-back working memory task (Chen and Huang), which suggests that beta oscillations play an important role in the functioning of

brain networks serving the levels of cognitive effort and time perception that is likely affected in schizophrenia (Alustiza et al.) Future studies should look more closely at the role of the representation of the time-dimension in the temporal processing of information in the brain, which may be affected in psychiatric illnesses.

## AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

## ACKNOWLEDGMENTS

We DG and HM, are grateful to the authors for their excellent contributions to this Frontiers Research Topic. This work was supported by CONACYT: 236836, CONACYT: 196, and PAPIIT: IN202317 grants to HM.

## REFERENCES

- Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford; New York, NY: Oxford University Press.
- Crowe, D. A., Zarco, W., Bartolo, R., and Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *J. Neurosci.* 34, 11972–11983. doi: 10.1523/JNEUROSCI.2177-14.2014
- Gupta, D. S. (2014). Processing of sub- and supra-second intervals in the primate brain results from the calibration of neuronal oscillators via sensory, motor, and feedback processes. *Front. Psychol.* 5:816. doi: 10.3389/fpsyg.2014.00816
- Gupta, D. S., and Chen, L. (2016a). Brain oscillations in perception, timing and action. *Curr. Opin. Behav. Sci.* 8, 161–166. doi: 10.1016/j.cobeha.2016.02.021
- Gupta, D. S., and Chen, L. (2016b). *Brain Oscillations in Perception, Timing and Action*. San Diego, CA: Society for Neuroscience.
- Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., et al. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* 80, 360–378. doi: 10.1016/j.neuroimage.2013.05.079
- 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. doi: 10.1523/JNEUROSCI.2523-13.2014
- Maniadakis, M., and Trahanias, P. (2008). Hierarchical co-evolution of cooperating agents acting in the brain-arena. *Adapt. Behav.* 16, 221–245. doi: 10.1177/1059712308087285
- Matell, M. S., and Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Brain Res. Cogn. Brain Res.* 21, 139–170. doi: 10.3389/fpsyg.2016.01558
- Meck, W. H., Penney, T. B., and Pouthas, V. (2008). Cortico-striatal representation of time in animals and humans. *Curr. Opin. Neurobiol.* 18, 145–152. doi: 10.1016/j.conb.2008.08.002
- Mello, G. B., Soares, S., and Paton, J. J. (2015). A scalable population code for time in the striatum. *Curr. Biol.* 25, 1113–1122. doi: 10.1016/j.cub.2015.02.036
- Merchant, H., Bartolo, R., Pérez, O., Méndez, J. C., Mendoza, G., Gámez, J., et al. (2014). Neurophysiology of timing in the hundreds of milliseconds: multiple layers of neuronal clocks in the medial premotor areas. *Adv. Exp. Med. Biol.* 829, 143–154. doi: 10.1007/978-1-4939-1782-2\_8
- Merchant, H., Grahm, J., Trainor, L., Rohrmeier, M., and Fitch, W. T. (2015a). Finding the beat: a neural perspective across humans and non-human primates. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 370:20140093. doi: 10.1098/rstb.2014.0093
- 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. doi: 10.1146/annurev-neuro-062012-170349
- Merchant, H., Pérez, O., Bartolo, R., Méndez, J. C., Mendoza, G., Gámez, J., et al. (2015b). Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. *Eur. J. Neurosci.* 41, 586–602. doi: 10.1111/ejn.12811
- Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi: 10.1016/j.clinph.2007.04.026
- Treisman, M. (1963). Temporal discrimination and the indifference interval. Implications for a model of the “internal clock.” *Psychol. Monogr.* 77, 1–31. doi: 10.1037/h0093864
- Wittmann, M., Carter, O., Hasler, F., Cahn, B. R., Grimberg, U., Spring, P., et al. (2007). Effects of psilocybin on time perception and temporal control of behaviour in humans. *J. Psychopharmacol.* 21, 50–64. doi: 10.1177/0269881106065859
- Zarco, W., Merchant, H., Prado, L., and Mendez, J. C. (2009). Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *J. Neurophysiol.* 102, 3191–3202. doi: 10.1152/jn.00066.2009

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Gupta and Merchant. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# In Search of Oscillatory Traces of the Internal Clock

Tadeusz W. Kononowicz\* and Virginie van Wassenhove\*

Cognitive Neuroimaging Unit, CEA DSV/I2BM, INSERM, Université Paris-Sud, Université Paris-Saclay, NeuroSpin Center, Paris, France

**Keywords:** time perception, interval timing, internal clock, oscillations, striatal beta frequency

## INTERVAL TIMING, PACEMAKER(S), AND NEURAL OSCILLATIONS

Neural oscillations are ubiquitous in the mammalian brain and they are typically classified according to their specific frequency responses (Buzsáki, 2006). Neural oscillations are hypothesized to organize communication within and between brain networks (e.g., Fries, 2015). Neural oscillations have increasingly been associated with various cognitive functions such as attention (Klimesch, 2012), working memory (Gulbinaite et al., 2014a; Haegens et al., 2014), and cognitive control (Cavanagh et al., 2009; Gulbinaite et al., 2014b) but also temporal expectation (Praagstra et al., 2006; Cravo et al., 2011; Rohenkohl and Nobre, 2011) and timing (Treisman, 1963; van Wassenhove, 2009; Kösem et al., 2014; Kononowicz and van Rijn, 2014). One quest in cognitive neuroscience is to explain how neural oscillations can subserve complex cognitive processes. Here, we mainly focus on the role of spontaneous rhythms in interval timing (also see van Wassenhove, in press); however, some hypotheses are supported by the literature on rhythmic entrainment.

One of the possible cognitive abilities neural oscillations may support is interval timing (Treisman et al., 1994), which is the ability to perceive, store, encode, and reproduce temporal intervals ranging from few 100 milliseconds to minutes. Over decades, experimental psychologists have proposed the existence of a cognitive mechanism akin to an internal clock. In search for the neural bases of the internal clock(s), it may be tempting to draw an analogy between ticking clocks and oscillating neuronal networks, as one of the reference papers in neurosciences states, “Clocks tick, bridges, and skyscrapers vibrate, neuronal networks oscillate” (Buzsáki and Draguhn, 2004, pp. 1926). Indeed, some interval timing theories have followed through this idea: Treisman (1963) suggested, that the internal clock could consist of a pacemaker which at the beginning of the to-be-timed interval would start sending pulses, that are then stored in the accumulator. The pulse count could serve as a subjective estimate of time. Furthermore, to implement this model into biologically plausible mechanisms, Treisman proposed that the pulse rate of the pacemaker would be driven by neural oscillations in the alpha range (8–12 Hz, see **Figure 1A**): faster alpha rhythms would thus result in longer estimates of time than slower alpha rhythms considering, that more pulses would accumulate during the same physical time interval (Treisman et al., 1994). As no simple relationships have been found between the rate of visual flicker, neural oscillations, and the subjective perception of duration when using oscillatory entrainment (Herbst et al., 2013, 2014; although see Johnston et al., 2006), it remains plausible that spontaneous fluctuations of alpha peaks could modulate perceived duration. For example, Haegens et al. (2014) have shown, that alpha peak frequency changed as a function of cognitive load in a N-back working memory (WM) task such that the larger the WM load, the higher the alpha peak frequency. These results indicate that subjectively longer durations could be associated with larger alpha peak if WM is implicated in the estimation of duration (Gu et al., 2015; van Wassenhove, in press).

Despite mechanistic attempts to link oscillatory processes with internal clock models, direct implementations of internal clock models still lack solid neural foundations whereas, more

### OPEN ACCESS

#### Edited by:

Hugo Merchant,  
Universidad Nacional Autónoma de  
México, Mexico

#### Reviewed by:

Peter Keller,  
University of Western Sydney,  
Australia  
Nicolas Escoffier,  
National University of Singapore,  
Singapore

#### \*Correspondence:

Tadeusz W. Kononowicz  
t.w.kononowicz@icloud.com;  
Virginie van Wassenhove  
virginie.van.wassenhove@gmail.com

#### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

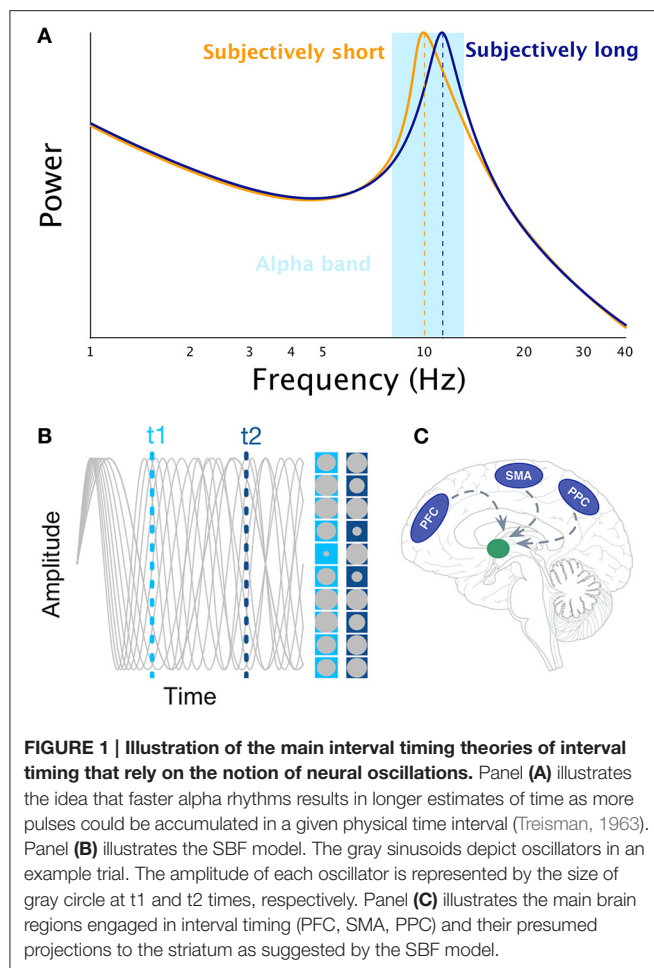
**Received:** 08 November 2015

**Accepted:** 03 February 2016

**Published:** 23 February 2016

#### Citation:

Kononowicz TW and van  
Wassenhove V (2016) In Search of  
Oscillatory Traces of the Internal  
Clock. *Front. Psychol.* 7:224.  
doi: 10.3389/fpsyg.2016.00224



biologically grounded frameworks have been more plausible (Buhusi and Meck, 2005).

## THE (STRIATAL) BEAT FREQUENCY MODEL

The most prominent neurobiologically plausible model of interval timing is the Striatal Beat Frequency (SBF) model (Matell and Meck, 2000, 2004; Buhusi and Meck, 2005) developed on the basis of the beat frequency model (Miall, 1989). One major assumption of SBF (Figure 1B) is the existence of cortical oscillators of various frequency responses most likely located in the Pre-Frontal Cortex (PFC) which is part of the mesocortical pathway. However, other cortical regions cannot be excluded [e.g., Supplementary Motor Area (SMA), Posterior Parietal Cortex (PPC), or sensory cortices, Figure 1C]. At the onset of an interval to be timed, the model posits, that cortical oscillators are phase-reset and, at the offset of the interval, the state of these cortical oscillators is read out by medium spiny neurons located in the striatum. Hence, the SBF model considers, that the phase of cortical oscillators gives rise to a unique activation pattern over time (Buhusi and Meck, 2005; Oprisan and Buhusi, 2011, 2014) and, that spiny neurons are coincidence detectors reading out the

state of these cortical oscillators. Note that in SBF, the pattern of activation is identical whether one reads the phase or the amplitude of the oscillators. Although, cortical oscillators seem to be a key element of the SBF model, only little evidence currently supports the existence of a dedicated set of cortical oscillators for interval timing (e.g., Matell, 2014). It is also unclear whether cortical oscillators are really necessary for the SBF model, as any stable pattern of neural activation (Crowe et al., 2014; Merchant et al., 2014; Mello et al., 2015) within to-be-timed intervals but variable across to-be-timed intervals would be sufficient as input to insure reliable coincidence detection (also see Meck et al., 2013).

## QUANTIFYING THE ROLE OF CORTICAL OSCILLATORS

When considering oscillatory processes in the context of the SBF model, at least two important predictions regarding neural oscillators have to be taken into account. The first prediction is, that in order to provide a meaningful pattern, cortical oscillators have to be phase-reset, such that they always start from the same fixed state. For example, the results by Parker et al. (2014) suggest, that more precise phase reset of ongoing theta oscillations in the medial frontal cortex results in better timing accuracy (Kononowicz, 2015), something that would be in line with the SBF model. This hypothesis awaits future tests and more compelling evidence have to be provided.

The second prediction is linked to the idea, that the speed of internal clock can be modulated by the speed of cortical oscillators (Oprisan and Buhusi, 2014), which are modulated by tonic levels of dopamine (Oprisan and Buhusi, 2011). It is very often assumed, that the clock speed could be represented by the alpha band regime (Treisman et al., 1990, 1994) as it is the most prevalent spontaneous rhythm in the mammalian brain (Oprisan and Buhusi, 2014). However, as previously discussed, the relationship between alpha peak power and fluctuations in subjective timing has not been clearly established; direct attempts to test this hypothesis have not succeeded (Treisman et al., 1990, 1994). The power of alpha is a good marker of temporal expectation (Praagstra et al., 2006; Cravo et al., 2011; Rohenkohl and Nobre, 2011), which is in line with the hypothesized role of alpha as a selective coordinator implicated in the temporal prioritization of sensory events (Jensen et al., 2014). Hence, one possible departure from the early proposals could be that a single dominating frequency may not be necessary to represent the clock speed as neural oscillations outside of the alpha range have been implicated in interval timing (Busch et al., 2004; Kaiser et al., 2007; Sperduti et al., 2011), raising the possibility that other rhythms could serve as “pacemakers.” For instance, recent studies suggest a significant role of beta oscillations in timing (Iversen et al., 2009; Fujioka et al., 2012, 2015; Bartolo et al., 2014; Teki, 2014; Kononowicz and van Rijn, 2014; Wiener and Kanai, 2016) and the phase characteristics of low-frequency oscillators can predict subjective timing (Cravo et al., 2011; Kösem et al., 2014), suggesting, that different neural oscillations have the potentiality to track time.

Therefore, instead of focusing on one single neural oscillation, future studies should explore local trial-to-trial fluctuations across frequency bands and how subdominant frequencies vary as a function of subjectively perceived time intervals. Complementary to this, addressing the implications of such markers at different time scales and across sensory modalities may be desirable.

Interestingly, a recent review by Gu et al. (2015) proposes to unify interval timing and working memory models. Specifically, these authors proposed, that working memory and interval timing can originate from the same oscillatory processes such as gamma and theta oscillations, and phase-amplitude coupling between these frequency bands (Lisman, 2010). The proposed model largely focuses on oscillatory processes that could be shared between working memory and SBF. Nonetheless, the empirical ways to assess the principles of SBF model are still lacking. As the gist of the SBF lies in the notion of communication between cortical areas and the striatum, here we discuss the possibility of testing this hypothesis by investigating functional connectivity between the striatum and PFC.

## STRIATUM-PFC COUPLING AND THE SBF MODEL

Striatal neurons are ideal candidates for coincidence detection as they receive direct inputs from cortical neurons. Through coincidence detection of spiking activity from two or more cortical regions, the same striatal neuron will discharge within a given time window. For instance, Matell et al. (2003) showed, that neural activity in the striatum and the anterior cingulate cortex varied before 10 and 40 s when the reinforcement was presented at one of these two time points, suggesting, that neuronal populations respond to particular time intervals. However, this pattern although predicted by SBF could largely be confounded by motor activity of lever pressing. Nevertheless, note, that Riehle et al. (1997) observed transient synchronization of neurons in motor cortex when stimuli were expected, but failed to appear. Although, this work is very important it only shows pattern of activity that fits into the SBF model under certain conditions. Given, that an important premise of the SBF model is the communication between striatal neuronal ensembles and cortical neurons, we propose, that investigating functional connectivity between subcortical and cortical structures can serve as an important step extending the results of Matell et al. (2003) and giving further support to the SBF model. For example, Antzoulatos and Miller (2014) found, that perceptual (non-temporal) category learning was accompanied by increased synchronization within beta band range (12–30 Hz) between the PFC and striatum, demonstrating the role of functional connectivity in learning. Specifically, synchronization was larger for correct trials. On the basis of the SBF model, a change of

cortico-striatal synaptic weights through learning is predicted to reflect a memory mechanism such as the one implemented in the Scalar Expectancy Theory (Gibbon, 1977). Taken together, striatal neurons are predicted to become more sensitive to firing as a function of specific PFC neurons, and these learning effects should be visible during training of temporal discrimination as a change in inter-areal synchronization. Moreover, according to the SBF model and in line with the results of Antzoulatos and Miller (2014), inter-areal synchronization should be enhanced in “correct” trials (Kononowicz, 2015). Particularly, the striatal-PFC synchrony enhancement should emerge at the time of a standard interval, for example in the task where subjects compare a comparison interval, that could vary in length to a fixed standard interval. That is because striatal and PFC structures should become transiently synchronous due to previous learning enhancing sensitivity/tuning of striatum to the particular neural pattern exhibited at the time of standard interval.

The synchronization of neural oscillations has been associated with neuronal mechanisms such as coincidence detection, neural plasticity though long term potentiation/depression mechanism, and neuronal communication (Fell and Axmacher, 2011). These processes seem like a plausible candidate to coordinate striatum-PFC communication in recognition for specific patterns considered by SBF model. Specifically, the simplest scenario would predict an increase in coherence or spike-filed coherence for accurately timed trials. Coherence was proposed to reflect facilitated communication between brain regions (e.g., Fries, 2015). Effective communication should be linked to the successful timing performance if indeed communication between the striatum and PFC is a key component of timing system as proposed in the SBF model. This cortico-striatal spike-field coherence should be specifically enhanced at the time of standard interval if striatal neurons recognize cortical pattern (see Kononowicz and van Rijn, 2015). Furthermore, the role of cortico-cortical coherence has been shown in passive rhythmical stimulation paradigms, in which an increase in coherence coincided with the next tone occurrence (Fujioka et al., 2012). These results do support the hypothesis sketched in this paper and also suggest cortico-cortical analysis. Moreover, recent progress in neuroscientific methods allows to address this questions in animals and humans using MEG/EEG modeling (David et al., 2011), but also deep brain recordings.

## AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

## ACKNOWLEDGMENTS

This work has been supported by ERC-YSt-263584 to VW.

## REFERENCES

- Antzoulatos, E. G., and Miller, E. K. (2014). Increases in functional connectivity between prefrontal cortex and striatum during category learning. *Neuron* 83, 216–225.
- Bartolo, R., Prado, L., and Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *J. Neurosci.* 34, 3910–3923. doi: 10.1523/JNEUROSCI.2679-13.2014
- Buhusi, C. V., and Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6, 755–765. doi: 10.1038/nrn1764
- Busch, N. A., Debener, S., Kranczioch, C., Engel, A. K., and Herrmann, C. S. (2004). Size matters: effects of stimulus size, duration and eccentricity on the visual gamma-band response. *Clin. Neurophysiol.* 115, 1810–1820. doi: 10.1016/j.clinph.2004.03.015
- Buzsáki, G. (2006). *Rhythms of the Brain*. New York, NY: Oxford University Press.
- Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745
- Cavanagh, J. F., Cohen, M. X., and Allen, J. J. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J. Neurosci.* 29, 98–105. doi: 10.1523/JNEUROSCI.4137-08.2009
- Cravo, A. M., Rohenkohl, G., Wyart, V., and Nobre, A. C. (2011). Endogenous modulation of low frequency oscillations by temporal expectations. *J. Neurophysiol.* 106, 2964–2972. doi: 10.1152/jn.00157.2011
- Crowe, D. A., Zarco, W., Bartolo, R., and Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *J. Neurosci.* 34, 11972–11983. doi: 10.1523/JNEUROSCI.2177-14.2014
- David, O., Maess, B., Eckstein, K., and Friederici, A. D. (2011). Dynamic causal modeling of subcortical connectivity of language. *J. Neurosci.* 31, 2712–2717. doi: 10.1523/JNEUROSCI.3433-10.2011
- Fell, J., and Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nat. Rev. Neurosci.* 12, 105–118. doi: 10.1038/nrn2979
- Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034
- Fujioka, T., Ross, B., and Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *J. Neurosci.* 35, 15187–15198. doi: 10.1523/JNEUROSCI.2397-15.2015
- Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *J. Neurosci.* 32, 1791–1802. doi: 10.1523/JNEUROSCI.4107-11.2012
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* 84:279. doi: 10.1037/0033-295X.84.3.279
- Gu, B. M., van Rijn, H., and Meck, W. H. (2015). Oscillatory multiplexing of neural population codes for interval timing and working memory. *Neurosci. Biobehav. Rev.* 48, 160–185. doi: 10.1016/j.neubiorev.2014.10.008
- Gulbinaite, R., Johnson, A., de Jong, R., Morey, C. C., and van Rijn, H. (2014a). Dissociable mechanisms underlying individual differences in visual working memory capacity. *Neuroimage* 99, 197–206. doi: 10.1016/j.neuroimage.2014.05.060
- Gulbinaite, R., van Rijn, H., and Cohen, M. X. (2014b). Fronto-parietal network oscillations reveal relationship between working memory capacity and cognitive control. *Front. Hum. Neurosci.* 8:761. doi: 10.3389/fnhum.2014.00761
- Haegens, S., Cousijn, H., Wallis, G., Harrison, P. J., and Nobre, A. C. (2014). Inter- and intra-individual variability in alpha peak frequency. *Neuroimage* 92, 46–55. doi: 10.1016/j.neuroimage.2014.01.049
- Herbst, S. K., Chaumon, M., Penney, T. B., and Busch, N. A. (2014). Flicker-induced time dilation does not modulate EEG correlates of temporal encoding. *Brain Topogr.* 1–11. doi: 10.1007/s10548-014-0389-z
- Herbst, S. K., Javadi, A. H., van der Meer, E., and Busch, N. A. (2013). How long depends on how fast—perceived flicker dilates subjective duration. *PLoS ONE* 8:e76074. doi: 10.1371/journal.pone.0076074
- Iversen, J. R., Repp, B. H., and Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Ann. N.Y. Acad. Sci.* 1169, 58–73. doi: 10.1111/j.1749-6632.2009.04579.x
- Jensen, O., Gips, B., Bergmann, T. O., and Bonnefond, M. (2014). Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends Neurosci.* 37, 357–369. doi: 10.1016/j.tins.2014.04.001
- Johnston, A., Arnold, D. H., and Nishida, S. (2006). Spatially localized distortions of event time. *Curr. Biol.* 16, 472–479. doi: 10.1016/j.cub.2006.01.032
- Kaiser, J., Leiberg, S., Rust, H., and Lutzenberger, W. (2007). Prefrontal gamma-band activity distinguishes between sound durations. *Brain Res.* 1139, 153–162. doi: 10.1016/j.brainres.2006.12.085
- Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn. Sci.* 16, 606–617. doi: 10.1016/j.tics.2012.10.007
- Kononowicz, T. W. (2015). Dopamine-dependent oscillations in frontal cortex index 'start-gun' signal in interval timing. *Front. Hum. Neurosci.* 9, 331. doi: 10.3389/fnhum.2015.00331
- 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. doi: 10.1523/JNEUROSCI.2523-13.2014
- Kononowicz, T. W., and van Rijn, H. (2015). Single trial beta oscillations index time estimation. *Neuropsychologia* 75, 381–389. doi: 10.1016/j.neuropsychologia.2015.06.014
- Kösem, A., Gramfort, A., and van Wassenhove, V. (2014). Encoding of event timing in the phase of neural oscillations. *Neuroimage* 92, 274–284. doi: 10.1016/j.neuroimage.2014.02.010
- Lisman, J. (2010). Working memory: the importance of theta and gamma oscillations. *Curr. Biol.* 20, 490–492. doi: 10.1016/j.cub.2010.04.011
- Matell, M. S. (2014). "Searching for the Holy Grail: temporally informative firing patterns in the Rat," in *Neurobiology of Interval Timing*, ed H. Merchant (New York, NY: Springer), 209–234.
- Matell, M. S., and Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *Bioessays* 22, 94–103. doi: 10.1002/(SICI)1521-1878(200001)22:1<94::AID-BIES14>3.0.CO;2-E
- Matell, M. S., and Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cogn. Brain Res.* 21, 139–170. doi: 10.1016/j.cogbrainres.2004.06.012
- Matell, M. S., Meck, W. H., and Nicolelis, M. A. (2003). Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behav. Neurosci.* 117, 760–773. doi: 10.1037/0735-7044.117.4.760
- Meck, W. H., Church, R. M., and Matell, M. S. (2013). Hippocampus, time, and memory—A retrospective analysis. *Behav. Neurosci.* 127, 642–654. doi: 10.1037/a0034201
- Mello, G. B., Soares, S., and Paton, J. J. (2015). A scalable population code for time in the striatum. *Curr. Biol.* 25, 1113–1122. doi: 10.1016/j.cub.2015.02.036
- Merchant, H., Bartolo, R., Pérez, O., Méndez, J. C., Mendoza, G., Gámez, J., et al. (2014). "Neurophysiology of timing in the hundreds of milliseconds: multiple layers of neuronal clocks in the medial premotor areas," in *Neurobiology of Interval Timing*, ed V. de Lafuente (New York, NY: Springer), 143–154.
- Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Comput.* 1, 359–371. doi: 10.1162/neco.1989.1.3.359
- Oprisan, S. A., and Buhusi, C. V. (2011). Modeling pharmacological clock and memory patterns of interval timing in a striatal beat-frequency model with realistic, noisy neurons. *Front. Integr. Neurosci.* 5:52. doi: 10.3389/fnint.2011.00052
- Oprisan, S. A., and Buhusi, C. V. (2014). What is all the noise about in interval timing? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369, 20120459. doi: 10.1098/rstb.2012.0459
- Parker, K. L., Chen, K. H., Kingyon, J. R., Cavanagh, J. F., and Narayanan, N. S. (2014). D1-dependent 4 Hz oscillations and ramping activity in rodent medial frontal cortex during interval timing. *J. Neurosci.* 34, 16774–16783. doi: 10.1523/JNEUROSCI.2772-14.2014
- Praamstra, P., Kourtis, D., Kwok, H. F., and Oostenveld, R. (2006). Neurophysiology of implicit timing in serial choice reaction-time performance. *J. Neurosci.* 26, 5448–5455. doi: 10.1523/JNEUROSCI.0440-06.2006
- Riehle, A., Grün, S., Diesmann, M., and Aertsen, A. (1997). Spike synchronization and rate modulation differentially involved in motor cortical function. *Science* 278, 1950–1953. doi: 10.1126/science.278.5345.1950
- Rohenkohl, G., and Nobre, A. C. (2011). Alpha oscillations related to anticipatory attention follow temporal expectations. *J. Neurosci.* 31, 14076–14084. doi: 10.1523/JNEUROSCI.3387-11.2011



- Sperduti, M., Tallon-Baudry, C., Hugueville, L., and Pouthas, V. (2011). Time is more than a sensory feature: attending to duration triggers specific anticipatory activity. *Cogn. Neurosci.* 2, 11–18. doi: 10.1080/17588928.2010.513433
- Teki, S. (2014). Beta drives brain beats. *Front. Syst. Neurosci.* 8:155. doi: 10.3389/fnsys.2014.00155
- Treisman, M. (1963). Temporal discrimination and the indifference interval: implications for a model of the “internal clock.” *Psychol. Monogr. Gen. Appl.* 77, 1–31. doi: 10.1037/h0093864
- Treisman, M., Cook, N., Naish, P. L., and MacCrone, J. K. (1994). The internal clock: electroencephalographic evidence for oscillatory processes underlying time perception. *Q. J. Exp. Psychol.* 47, 241–289. doi: 10.1080/14640749408401112
- Treisman, M., Faulkner, A., Naish, P. L., and Brogan, D. (1990). The internal clock: evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception* 19, 705–743. doi: 10.1068/p190705
- van Wassenhove, V. (2009). Minding time in an amodal representational space. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1815–1830. doi: 10.1098/rstb.2009.0023
- van Wassenhove, V. (in press). Temporal Cognition and Neural Oscillations. *Curr. Opin. Behav. Sci.*
- Wiener, M., and Kanai, R. (2016). Frequency tuning for temporal perception and prediction. *Curr. Opin. Behav. Sci.* 8, 1–6. doi: 10.1016/j.cobeha.2016.01.001
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Kononowicz and van Wassenhove. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Unpredicted Pitch Modulates Beta Oscillatory Power during Rhythmic Entrainment to a Tone Sequence

Andrew Chang<sup>1</sup>, Dan J. Bosnyak<sup>1,2</sup> and Laurel J. Trainor<sup>1,2,3\*</sup>

<sup>1</sup> Department of Psychology, Neuroscience and Behaviour, McMaster University, Hamilton, ON, Canada, <sup>2</sup> McMaster Institute for Music and the Mind, McMaster University, Hamilton, ON, Canada, <sup>3</sup> Rotman Research Institute, Baycrest Hospital, Toronto, ON, Canada

## OPEN ACCESS

### Edited by:

Hugo Merchant,  
Universidad Nacional Autónoma  
de México, Mexico

### Reviewed by:

Michael Schwartz,  
Maastricht University, Netherlands  
John Rehner Iversen,  
University of California-San Diego,  
USA

### \*Correspondence:

Laurel J. Trainor  
ljt@mcmaster.ca

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 25 November 2015

**Accepted:** 21 February 2016

**Published:** 09 March 2016

### Citation:

Chang A, Bosnyak DJ and Trainor LJ  
(2016) Unpredicted Pitch Modulates  
Beta Oscillatory Power during  
Rhythmic Entrainment to a Tone  
Sequence. *Front. Psychol.* 7:327.  
doi: 10.3389/fpsyg.2016.00327

Extracting temporal regularities in external stimuli in order to predict upcoming events is an essential aspect of perception. Fluctuations in induced power of beta band (15–25 Hz) oscillations in auditory cortex are involved in predictive timing during rhythmic entrainment, but whether such fluctuations are affected by prediction in the spectral (frequency/pitch) domain remains unclear. We tested whether unpredicted (i.e., unexpected) pitches in a rhythmic tone sequence modulate beta band activity by recording EEG while participants passively listened to isochronous auditory oddball sequences with occasional unpredicted deviant pitches at two different presentation rates. The results showed that the power in low-beta (15–20 Hz) was larger around 200–300 ms following deviant tones compared to standard tones, and this effect was larger when the deviant tones were less predicted. Our results suggest that the induced beta power activities in auditory cortex are consistent with a role in sensory prediction of both “when” (timing) upcoming sounds will occur as well as the prediction precision error of “what” (spectral content in this case). We suggest, further, that both timing and content predictions may co-modulate beta oscillations via attention. These findings extend earlier work on neural oscillations by investigating the functional significance of beta oscillations for sensory prediction. The findings help elucidate the functional significance of beta oscillations in perception.

**Keywords:** sensory prediction, beta band, EEG oscillations, rhythmic entrainment, pitch, attention, auditory cortex, oddball

## INTRODUCTION

Perceptual systems extract regularities from the stream of continuous sensory input, and form internal representations for predicting future events. Predictive timing is the sensory prediction (or expectation) of when an event will occur (Nobre et al., 2007; Schroeder and Lakatos, 2009). Such predictions are hypothesized to be essential for many human behaviors, including understanding speech and music (Ding et al., 2015; Doelling and Poeppel, 2015), and synchronizing movements (Jenkinson and Brown, 2011; Fujioka et al., 2012, 2015; Kilavik et al., 2013). Predictive timing can be studied at a basic level in that an isochronous stream of metronome clicks sets up a strong prediction for when the next click will occur.

Entrainment is the process of internal neural oscillations becoming synchronized with temporal regularities in an external auditory rhythmic input stream, and it provides a mechanism for

predicting future events in time (Jones, 2010). Such entrainment appears to be accomplished in the brain by neural oscillatory activity, which has been shown to represent temporal regularities in the sensory input, as well as the prediction of upcoming sensory events (Friston, 2005; Jones, 2010; Arnal and Giraud, 2012; Fujioka et al., 2012, 2015; Henry and Herrmann, 2014; Morillon and Schroeder, 2015; Herrmann et al., 2016). While time domain event-related potential (ERP) analyses of electroencephalogram (EEG) waveforms in response to unpredicted stimuli have revealed aspects of neural processes underlying sensory prediction (e.g., Costa-Faidella et al., 2011; Schwartze and Kotz, 2013; Schröger et al., 2015), recent studies indicate that neural oscillatory activities obtained by decomposing EEG signals into frequency-specific bands reveal processes of communication between neural ensembles (Buzsáki, 2006) that are essential to sensory prediction (Arnal and Giraud, 2012).

Oscillatory activities in sensory cortices in both delta (1–3 Hz) and beta (15–25 Hz) bands are associated with temporal prediction (Henry and Herrmann, 2014). The phase of the delta oscillation shows entrainment to rhythmic sequences and it is reset by the onset of a stimulus and predicted (imagined) onset of a future stimulus. On this basis, it has been suggested that delta phase reflects an oscillatory time frame for parsing a continuous sensory stream into meaningful chunks for subsequent perceptual processing (Schroeder and Lakatos, 2009; Calderone et al., 2014). Neural responses to sensory inputs that occur at the time of the excitation phase of delta oscillations are enhanced compared to those that coincide with the inhibition phase (Schroeder and Lakatos, 2009). Local field potential recordings in primary visual and auditory cortices of macaque monkeys show that the delta phase entrains to the onsets of stimuli in rhythmic stimulus streams (Lakatos et al., 2008, 2013), consistent with intracranial electrocortical and surface EEG recordings in humans (Besle et al., 2011; Gomez-Ramirez et al., 2011; Henry and Obleser, 2012; Herrmann et al., 2016), and it can be endogenously directed by selectively attending to one or the other of two simultaneous stimulus streams (Lakatos et al., 2008, 2013; Calderone et al., 2014).

The amplitude fluctuation dynamics of induced (non-phase-locked) beta band power also entrain to the tempo of events in an auditory input stream, as well as reflecting temporal prediction. EEG and MEG recordings of isochronous auditory sequences show that induced beta power decreases following each tone onset, and increases again prior to the onset time of the next tone, with the timing of the increase varying with tempo in a predictive manner (Snyder and Large, 2005; Fujioka et al., 2009, 2012, 2015; Iversen et al., 2009; Cirelli et al., 2014; **Figure 1**). Both delta phase angle and beta power in auditory and motor areas in the pre-stimulus onset period predict the accuracy of detecting a temporal delay in the stimulus (Arnal et al., 2015). Furthermore, in primary motor cortex, beta power is modulated by attention, and aligned with the delta phase, suggesting that beta power might reflect attentional fluctuation in time and delta phase an entrained internal clock that aids in the execution of a motor task (Saleh et al., 2010).

Although delta phase and induced beta power are both associated with temporal prediction, compared to the compelling evidence for delta oscillations, the functional significance of beta oscillations in perceptual processing remains less clear. We hypothesized that the entrainment of induced beta power in auditory cortex to an external stimulus might reflect more than predictive timing. Specifically, given that auditory cortex is sensitive to both spectral and temporal dimensions of the input (Fritz et al., 2003; Griffiths and Warren, 2004; King and Nelken, 2009), and auditory evoked ERP components can be interactively modulated by predictions of both pitch and time (Costa-Faidella et al., 2011), beta oscillations might also reflect predictive coding for specific content, such as pitch. In order to examine this hypothesis, we conducted two experiments in which we presented isochronous auditory oddball sequences containing occasional deviations in pitch at different presentation rates. If the induced beta power only reflects predictive timing, the occasional unpredicted pitch changes should not affect the ongoing beta entrainment behavior, given that the pitch deviants are presented at the predicted rhythmic time points. On the other hand, if the induced beta power is affected by the unpredicted deviant pitches, it would suggest that beta power is associated with predictive perceptual processing for both what and when. In the case that induced beta power is affected by unpredicted deviant pitches, we examine further whether it is modulated by response to novelty (rare events in the preceding local context) or prediction error (the probability of encountering a deviant pitch under the statistical conditions of the context).

## MATERIALS AND METHODS

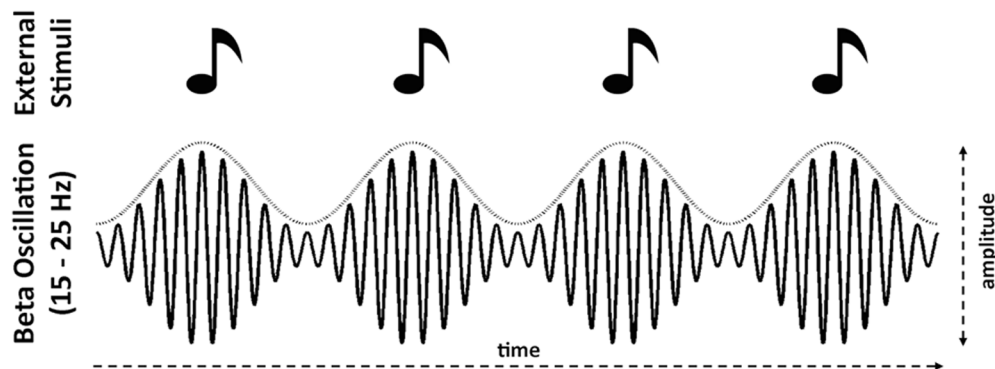
### Stimuli

Two recorded piano tones, C4 (262 Hz) and B4 (494 Hz), from the University of Iowa Musical Instrument Samples were used. The amplitude envelopes of the piano tones were percussive with 10 ms rise times. Tones were truncated to be 200 ms in duration, and a linear decay to zero was applied over the entire excerpt to remove offset artifact. The DC shift was removed for each tone. Sounds were converted into a monaural stream at 71 dB (C weighted), measured through an artificial ear (type 4152, Brüel & Kjær) with sound level meter (type 2270, Brüel & Kjær).

### Procedure

The experiment was conducted in a sound-attenuated room. Each participant was presented with a continuous sequence of tones in two sessions, each lasting 30 min, while they watched a silent movie on a computer screen. Participants took a 3-min break between sessions. Sounds were delivered binaurally via ear inserts (Etymotic Research ER-2). All stimulus sequences were presented under the control of a digital signal processor (Tucker Davis RP2.1).

The tones were presented in an oddball sequence. The C4 tone was used as the standard and the B4 tone as the deviant. For the first group of participants, the inter-onset interval (IOI) was fixed at 500 ms. There were 3600 tones presented in each session, and the deviance occurrence rate was 10% in one session and



**FIGURE 1 | Schematic illustrations of power modulation in induced (non-phase-locked) beta (15–25 Hz) entraining to the tempo of the stimuli.**

Specifically, power decreases following isochronous onsets and increases that predict the onset time of the next stimulus (e.g., Fujioka et al., 2012; Cirelli et al., 2014). The dotted curve above the beta waveform envelope represents this power modulation.

20% in the other session, with an equal number of participants completing the 10% or 20% session first. Within each session, tone order was pseudorandomized with the constraint that two deviant tones could not be presented sequentially, and each session started with five consecutive standard tones. Participants were instructed to sit comfortably and remain as still as possible during the experiment while watching a silent movie. They were not required to make any responses.

In order to replicate and to generalize the findings to a different presentation rate, for a second group of participants, we employed a longer IOI of 610 ms in an isochronous oddball sequence with the 10% deviant tones condition. Otherwise, the procedure for group two was the same as that for group one.

For convenience, we refer to the 500 ms IOI experimental sessions (10% and 20% deviance occurrence rates) as the Fast Experiment, and the 610 ms IOI experimental session (10% deviance occurrence rate only) as the Slow Experiment.

## Participants

Sixteen participants (17–22 years old, mean age  $18.93 \pm 1.39$ ; 12 female) for the Fast Experiment and a different thirteen participants (17–21 years old, mean age  $18.62 \pm 1.33$ , 10 female) for the Slow Experiment were recruited from the McMaster University community. Participants were screened by a self-report survey to ensure they had normal hearing, were neurologically healthy and were right-handed. Signed informed consent was obtained from each participant. The McMaster University Research Ethics Board approved all procedures. Participants received course credit or reimbursement for completing the study.

## Electroencephalographic Recording

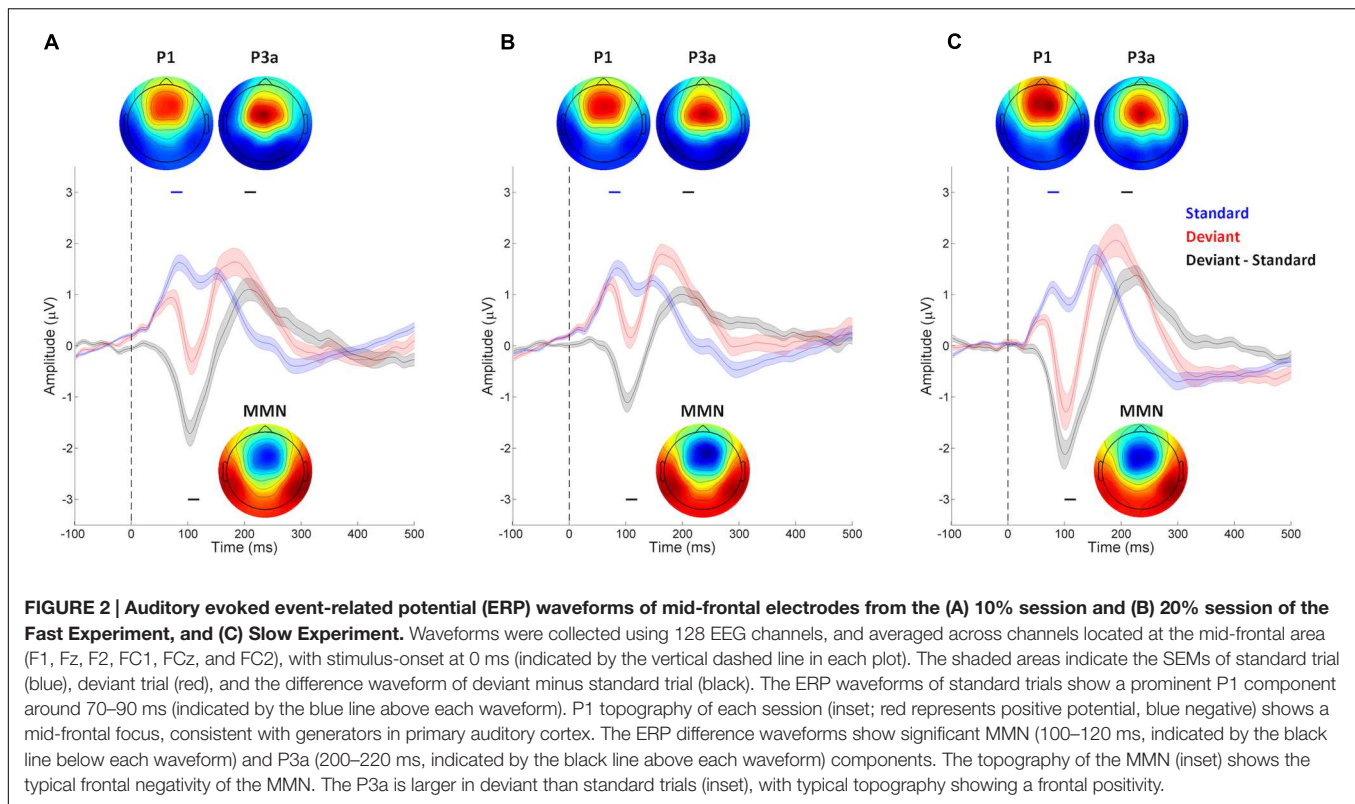
The EEG was sampled at 2048 Hz (filtered DC to 417 Hz) using a 128-channel Biosemi Active Two amplifier (Biosemi B.V., Amsterdam). The electrode array was digitized for each participant (Polhemus Fastrak) prior to recording. EEG data were stored as continuous data files referenced to the vertex electrode.

## Signal Processing of the EEG Data

Three stages of signal processing were conducted in order to examine the behavior of auditory evoked and induced oscillations in bilateral auditory cortices. In the first stage, we obtained a dipole source model based on auditory evoked responses, following Fujioka et al. (2012). The second stage segmented and categorized the source waveform into epochs based on the relative order of the presented auditory sequence. In the third stage, epochs containing excessive artifacts were rejected.

### Stage 1: Dipole Source Modeling

The continuous EEG data was band-pass filtered 0.3–100 Hz for each participant for each session, and then segmented into epochs covering the time period -100 to 300 ms, time locked to stimulus onset. Epochs containing standard tones that preceded and followed other standard tones with amplitudes exceeding  $150 \mu\text{V}$  were rejected as artifacts. The surviving standard epochs ( $89.6\% \pm 5.1\%$  for 10% session and  $89.5\% \pm 5.1\%$  for 20% session of Fast Experiment, and  $88.4\% \pm 5.5\%$  of Slow Experiment) were averaged into ERP waveforms and band pass filtered between 1 and 20 Hz (Figure 2). To confirm that our oddball context was set up appropriately, a similar procedure was performed on the deviant epochs, and the average of the standard epochs subtracted from the average of the deviant epochs in order to produce difference waves. As can be seen in Figure 2, both mismatch negativity (MMN) and P3a responses can be observed, consistent with the literature on ERP responses in oddball contexts (Friedman et al., 2001). Paired *t*-tests, performed on the average of channels in the mid-frontal area (F1, Fz, F2, FC1, FCz, and FC2), confirmed the presence of an MMN component between 100 and 120 ms; specifically, deviant trials were significantly more negative than standard trials in this time window in all sessions of both Fast and Slow Experiments ( $p < 0.001$ ). There was also a P3a component between 200 and 220 ms: deviant trials were significantly more positive than standard trials in this time window in all sessions of both Fast and Slow Experiments ( $p < 0.001$ ). It is worth noting that although the latencies of MMN and P3a observed in the current study were



earlier than are sometimes reported (e.g., MMN: 150–250 ms, P3a: 250–300 ms; Friedman et al., 2001; Näätänen et al., 2007; Polich, 2007), our results are consistent with several previous studies showing that the latencies of MMN and P3a are as short as around 100 and 200 ms, respectively, when the stimuli are presented in a rhythmic context with IOIs less than or equal to 700 ms (e.g., Regnault et al., 2001; Jongsma et al., 2004; Pablos Martin et al., 2007; Matsuda et al., 2013).

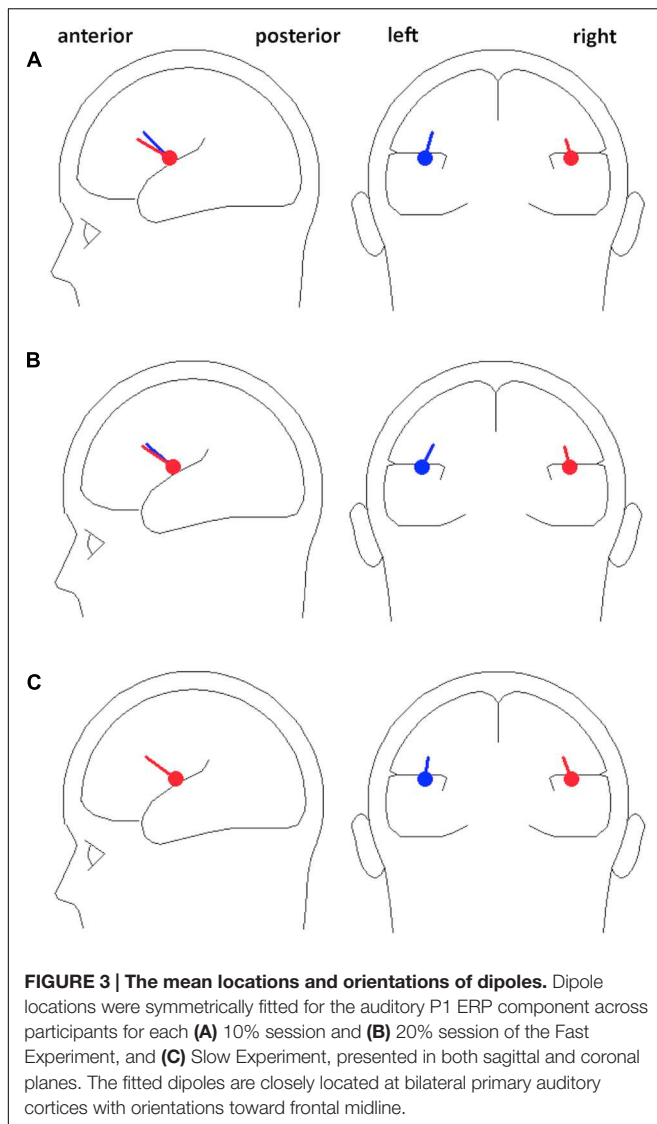
We employed a dipole source model as a spatial filter for increasing the signal-to-noise ratio of the EEG signal generated from left and right auditory cortices for subsequent analyses. A previous study showed that beta activities generated in both auditory and motor cortices entrained to external auditory rhythms when participants passively listened to isochronous sequence of tones (Fujioka et al., 2012). In the present study, we were primarily interested in responses from auditory areas, so we analyzed the EEG signals in source space rather than from surface channels, to extract the oscillatory signals generated from auditory cortex while attenuating signals generated from other brain regions. The source modeling was performed on each participant's mean standard ERP waveform using the multiple source probe scan algorithm and the four-shell ellipsoid model included in the Brain Electrical Source Analysis (BESA) software package. Two auditory cortex sources were estimated for each participant for the auditory evoked P1 (60–100 ms; Figure 2) with the dipoles constrained to be symmetric across hemispheres in location but not orientation. P1 was chosen because it is the dominant peak at fast presentation rates (N1 peaks are strongly reduced at fast rates; Näätänen and Picton, 1987), and is

generated primarily from primary auditory cortex (Godey et al., 2001). The mean locations of fitted dipoles across participants were at Talairach coordinates  $-45.0, -3.2, 16.2$  with orientation  $(0.2, 0.6, 0.8)$  and  $45.0, -3.2, 16.2$  with orientation  $(-0.1, 0.7, 0.7)$  in the 10% session of the Fast Experiment; and at  $-45.4, -3.1, 17.2$  with orientation  $(0.3, 0.7, 0.7)$  and  $45.4, -3.1, 17.2$  with orientation  $(-0.1, 0.8, 0.6)$  in the 20% session of the Fast Experiment; and  $-44.9, -4.7, 16.4$  with orientation  $(0.1, 0.7, 0.7)$  and  $44.9, -4.7, 16.4$  with orientation  $(-0.2, 0.7, 0.7)$  in the Slow Experiment, which are all closely located at bilateral primary auditory cortices with orientations toward the mid-frontal surface area (Figure 3). The residual variances of the source fittings for each session for each participant were between 5% and 10%.

## Stage 2: Epoching

Based on individual participant dipole model fits for each session, the source activities of single trials in auditory cortices were extracted for all epoch types using signal space projection following Fujioka et al. (2012). Because we were interested in the inter-stimulus neural responses, and to avoid edge effects in subsequent time-frequency analysis, the unfiltered EEG data of each session were segmented into relatively long -500 to 1000 ms epochs, where 0 ms represents a stimulus onset. The epochs were categorized based on the relative position of tones presented in the experiment, including standard (standard tones between two standard tones), deviant (deviant tones between two standard tones) and SpreD (standard tones preceding a deviant tone and following a standard tone). The individual source waveform





epochs as well as raw channel EEG data were exported from BESA to MATLAB for further processing.

### Stage 3: Artifact Rejection

Another artifact rejection procedure was applied to the raw 128-channel data. Epochs identified to have artifacts were noted, and the corresponding source waveform epochs were eliminated from further analysis. Thus we made sure the source waveform epochs entered into the time-frequency analysis in the next stage were artifact-reduced and unfiltered, to maximize the signal-to-noise ratio. Because we aimed to reject epochs containing EOG or EMG responses, each raw channel EEG epoch was filtered by a third-order Butterworth band pass filter (1–60 Hz). The filtered EEG epochs that exceeded a threshold (40  $\mu$ V, compared to the baseline mean voltage of -100–0 ms) for more than 10% of the epoch at any channel were excluded from further analysis. An additional seven participants' data were not included in the current data set because more than 50% of their epochs did not

pass the criteria at this stage. For the remaining participants  $66.18\% \pm 8.68\%$  of the epochs in the Fast Experiment and  $71.57\% \pm 10.54\%$  in the Slow Experiment were accepted for further analysis.

### Time-Frequency Decompositions

Time-frequency decompositions were calculated for each participant on each single-epoch source waveform in left and right auditory cortices and for each stimulus condition using a Morlet wavelet transform (Bertrand et al., 1994) for beta frequency band.

In order to remove the evoked (phase-locked) responses from the epoch and thereby obtain the induced (non-phase-locked) responses for subsequent analyses on beta band, we averaged the source waveform for each trial type (evoked response estimate), and then subtracted it from each source waveform epoch (Kalcher and Pfurtscheller, 1995; Fujioka et al., 2012).

The Morlet wavelet transformation was calculated for each time point for each induced epoch with 32 logarithmically spaced frequency bins between 15 and 25 Hz. The wavelet was designed such that the half-maximum width was equal to 3.25 periods of the lowest frequency while the width was equal to 3.56 periods of the highest frequency, linearly interpolated for each frequency bin in between. Subsequently, 300 ms at the beginning and ending of the epoch were eliminated to avoid edge effects. The induced oscillatory mean signal power was calculated by averaging the magnitude of each time-frequency point of wavelet coefficients across trials. Normalizing this to the mean value of the standard epochs across the whole epoch for each frequency resulted in relative signal power changes expressed as a percentage (Fujioka et al., 2012), and all types of epochs within the same session were compared to the same baseline (mean power in the averaged standard epoch between 0 and 500 ms). The fluctuation in power for each type of epoch at each frequency was visualized as a function of time and frequency in color-coded maps of event-related synchronization and desynchronization (Pfurtscheller and Lopes da Silva, 1999).

### Discrete Fourier Transform for Neural Oscillation Entrainment

In order to examine whether the observed neural oscillation activity entrained to the presented stimulus rate, we analyzed the time series of each participant's normalized mean induced beta power (derived as above) via discrete Fourier transforms (DFT). For each participant, we took the -200 to 700 ms epoch for the averaged induced beta power from the wavelet transform, zero-padded to 5 s in order to increase the frequency resolution of the DFT to a bin size 0.2 Hz. For each of the beta power time series, the power spectrums revealed by the DFTs were averaged across participants at each of the left and right auditory cortices.

### Data Analysis and Statistics

In order to examine whether the deviant tone affected the beta band induced power (1) we compared the standard and deviant trials for each individual participant for both the 10% and 20% deviance sessions to identify deviant-elicited prediction error



responses, and (2) we compared this difference of “standard - deviant” between the 10% and 20% deviance rate sessions to investigate the effect of prediction precision, as deviants in the 10% session are less predicted than those in the 20%. We analyzed the window 0–500 ms for the Fast Experiment and 0–610 ms for the Slow Experiment, time-locked to stimulus onset. The standard and deviant trials of individual participants were then used for random effects analysis.

To assess the statistical differences between the induced beta band powers while controlling for multiple comparisons, we performed cluster-based permutation analyses on the two-dimensional time-frequency maps (Maris and Oostenveld, 2007). First, we used a Wilcoxon signed-rank test, a non-parametric paired difference test, to examine the mean power difference in the beta band between each paired time-frequency sample from 0 to 500 ms for the Fast Experiment or 0–610 ms for the Slow Experiment. Second, we grouped the time-frequency adjacent samples reaching a threshold of  $p < 0.05$  into single clusters. Third, we summed the test statistics within each cluster into a cluster-level statistic, which became the observed value. Fourth, to build a permutation distribution, we randomly interchanged the experimental conditions for each participant, repeated the previous three steps 5000 times, and extracted the largest cluster-level statistics for each repetition. The final p-value was calculated by comparing the observed value of each cluster with the permutation distribution.

## RESULTS

We first tested whether the induced beta power entrainment phenomenon reported by Fujioka et al. (2012) was replicated in the standard trials. In the Fast Experiment, the induced power in the beta band of the standard trials showed a clear entrainment to the IOI rate (2.0 Hz). Specifically, the DFT analysis on induced beta band power showed the strongest power at 2.0 Hz for both the 10% and 20% sessions at both left and right auditory cortices (**Figures 4A–D**). In the Slow Experiment, the induced power in the beta band of the standard trials showed a clear entrainment to the slower IOI rate (~1.6 Hz) with the DFT analysis showing the strongest power at 1.6 Hz at both left and right auditory cortices (**Figures 4E,F**). These results replicate previous studies showing that induced beta band power entrains to the IOI of isochronous stimulus sequences (Fujioka et al., 2009, 2012, 2015; Cirelli et al., 2014).

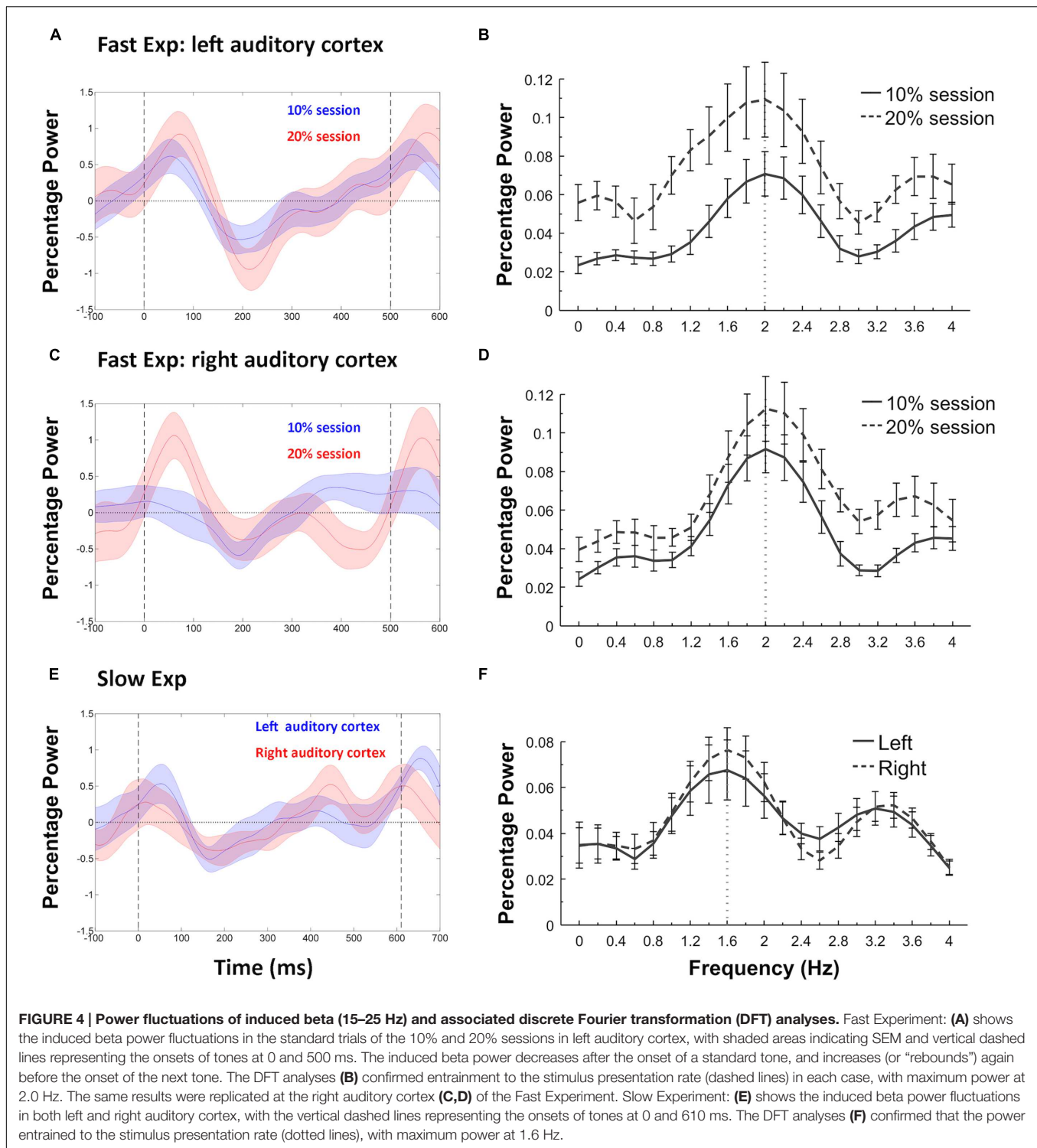
We then examined whether trial type (deviant vs. standard) and session (deviant rate) modulate the induced beta power, in addition to the entrainment activities. In the Fast Experiment, the cluster-based permutation test identified one significant cluster in the 10% session at right auditory cortex, in which the mean induced power at 16–20 Hz, within the range of low-beta band (15–20 Hz), around 200–300 ms after stimulus onset was larger in the deviant trials than the standard trials ( $p = 0.044$ ; **Figure 5A**) with a large effect size (rank correlation = 0.67). We did not identify any significant cluster at left auditory cortex. We examined the same contrast for the 20% session. Although we failed to identify any significant cluster at either left or

right auditory cortex, the power difference of “deviant–standard” trials peaked around 200–300 ms in the low-beta band at right auditory cortex (**Figure 5B**), which is consistent with the results of the 10% session. We further compared the power difference of “deviant–standard” trials between the 10% and 20% sessions at the previously identified cluster. The Wilcoxon signed-rank test showed that the power difference was significantly larger in the 10% session than in the 20% session ( $p = 0.026$ ), with a large effect size (rank correlation = 0.56). Taken together, this indicates that the induced power in low-beta band around 200–300 ms after stimulus onset was higher in deviant trials than in standard trials, and that this effect was larger in the 10% session than in the 20% session.

The results of the Slow Experiment replicated the results of the Fast Experiment. A cluster-based permutation test showed only one significant cluster around 200–300 ms after stimulus onset at 15–19 Hz at right auditory cortex ( $p = 0.026$ ; **Figure 5C**), in which the mean induced power was larger in the deviant trials than the standard trials with a large effect size (rank correlation = 0.79).

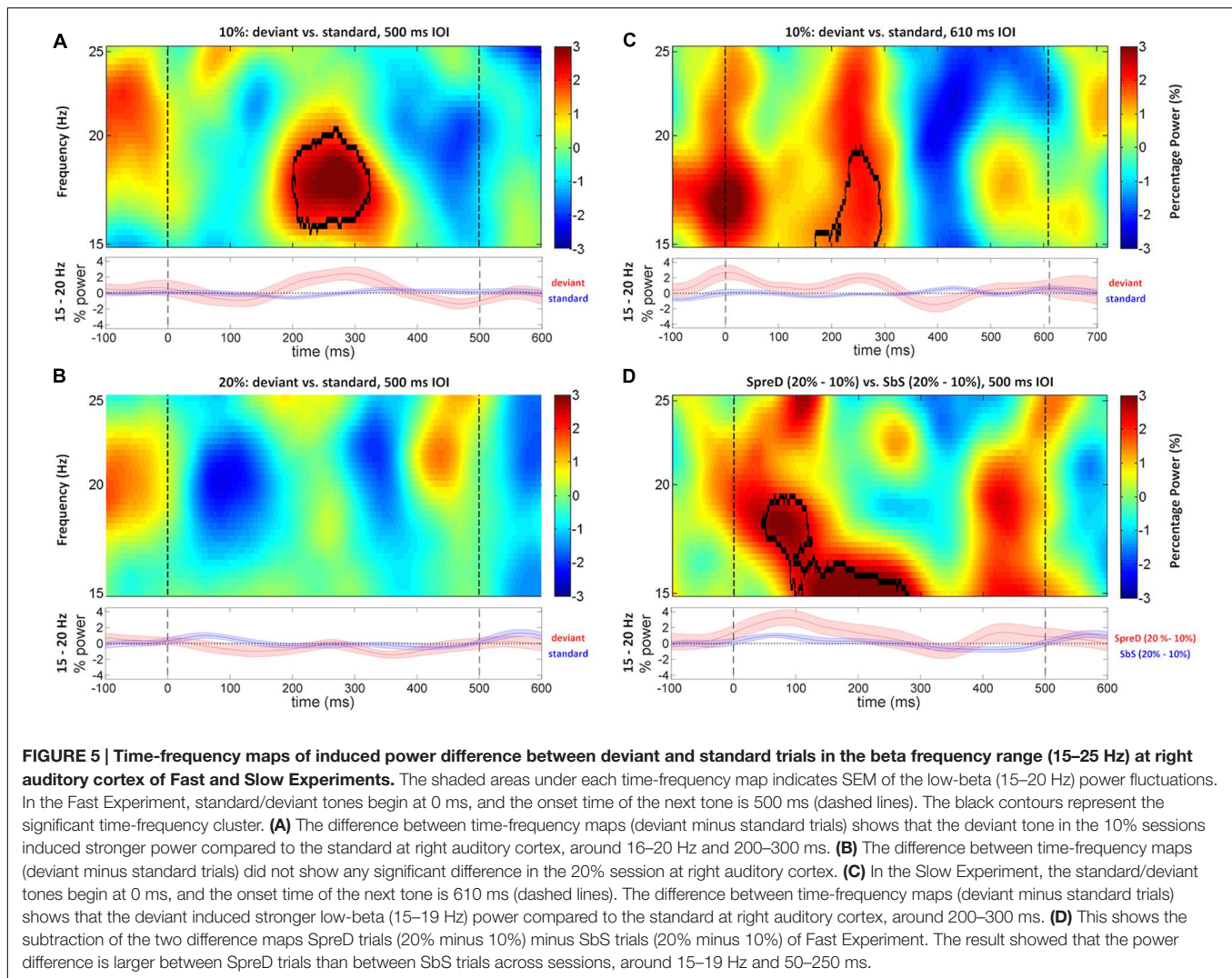
To further distinguish whether the deviant-induced responses in low-beta band are associated with prediction error or response to novelty (rare events in the preceding local context), given that both processes can be engaged by deviant stimuli in an oddball context (Friedman et al., 2001), we performed an additional analysis for standard tones occurring in different places in the sequence. This was based on the idea that in an oddball sequence, not only can the presentation of a deviant tone violate a prediction for a standard tone, but also the presentation of a standard tone that follows several standard tones in a row can violate an expectation (prediction) for a deviant tone. Specifically, the more standards that occur in a row, the more likely it is that a deviant will occur next, given a fixed overall probability of a deviant. On the other hand, a standard occurring after several standards in a row would not elicit a novelty response, as there is no change in the stimulus. If the beta band response that we measured reflects prediction error and not response to novelty, then the response to standard tones should depend on how many standards occurred prior to the standard of interest (as each successive standard builds prediction for an eventual deviant), whereas if the response simply associates with novelty, there should be a larger response to standards in the 20% than 10% condition, but no effect of how many standards occur in a row. Given that a deviant tone must occur eventually along the time line (Luce, 1986; Nobre et al., 2007), the conditional likelihood of encountering a standard tone decreases with the number of repetitions of the standard tone in a row, and thus, on average, the prediction of standard tones preceding a deviant tone will be lower in the 10% than in the 20% session since there are on average more standards in a row before each deviant in the 10% condition.

We can compare responses to standards between 10% and 20% sessions that occur either immediately before a deviant (SpreD) or between two other standards in the sequence (here referred to as SbS). SbS trials occur earlier on average in the sequence compared to SpreD trials. This allows a test of the two alternative hypotheses. Specifically, if the induced low-beta power



response at right auditory cortex results from prediction error, the power difference between SpreD trials (20% session–10% session) should be larger than the difference between SbS trials (20% session–10% session), because the prediction error (mismatch between standard and deviant tone) is modulated by conditional likelihood (the position of standard tones in a stimulus sequence).

On the other hand, if the induced low-beta power response is modulated by the novelty in the preceding context, the power difference between SpreD trials (20% session–10% session) should be equal to the difference between SbS trials (20% session–10% session), because the conditional likelihood does not matter. Indeed, if anything, the SbS trials would be predicted to show a



larger induced low-beta power difference than the SpreD trials because the SbS trials constitute a deviation from a more recently presented deviant tone whereas SpreD trials follow a larger number of standard trials. A cluster-based permutation test in low-beta band at right auditory cortex showed that the SpreD trials had a larger induced power difference than the SbS trials ( $p = 0.045$ ; **Figure 5D**) around 50–250 ms at 15–19 Hz with a large effect size (rank correlation = 0.74). This suggests that the increased induced low-beta power is elicited by prediction error, modulated by conditional likelihood, rather than response to novelty, modulated by rareness of a pitch in the preceding context.

Another additional analysis was performed to investigate whether the current results were associated with the mechanism of auditory stimulus-specific adaptation (SSA) rather than sensory prediction. Auditory SSA refers to the phenomenon that the neural response to the same tone decreases as the number of times it is repeated increases, and raises the possibility that responses to rare tones in an oddball context reflect release from adaptation rather than prediction or response to novelty (e.g.,

Butler, 1968; Näätänen et al., 1988; Lanting et al., 2013). In the present study, it is possible that the magnitude of the low-beta response to pitch deviants reflects a release from adaptation to the repeated standard tones in our oddball context. Further, the finding that the low-beta power response was stronger on deviant trials in the 10% than 20% session might be due to the fact that there were on average more repeated standard tones preceding a deviant trial in the former case. In order to investigate whether the low-beta response was modulated by a predictive process, we compared conditions where the effect of SSA was constant, but prediction differed. Specifically, to accomplish this, we compared 10% and 20% sessions of the Fast Experiment where the number of standards since the previous deviant was held constant. Thus, we averaged separately deviant effects where there were two standards, three standards, four standards, five standards, or six standards since the last deviant. In each case, we took the low-beta power difference of deviant minus standard trials and compared between the 10% and 20% sessions. The critical point is that, for a given number of standard trials preceding a deviant, the sensory prediction hypothesis indicates that deviants are



more expected in the 20% than 10% session because there is a generally higher probability of a deviant in the 20% condition. Specifically, the conditional likelihoods of encountering a deviant tone can be estimated by summing up the empirical occurrence rates of a deviant tone in the all the locations in a sequence following a deviant trial, until the current location (Figure 6A). We performed a cluster-based permutation test on the low-beta band at right auditory cortex. We did not find any cluster to be significant, but there was a trend for the power difference at the cluster at 200 to 300 ms to be larger in the 10% session than in the 20% session (Figure 6B) as predicted by the sensory prediction hypothesis. The fact that it did not reach conventional significance levels is likely due to the small number of trials (in the 10% session,  $141.0 \pm 19.0$  deviant trials were included in the current analysis, compared to the  $244.6 \pm 37.8$  trials that were included in previous analyses). We compared the maximum deviant minus standard power difference of the averaged low-beta frequency band between 10% and 20% sessions in the time window 130–370 ms for each participant, time-locked to stimulus onset (Figure 6C). The Wilcoxon signed-rank test showed that the maximum low-beta power difference between deviant and standard trials was significantly larger in the 10% session than in the 20% session ( $2.96 \pm 1.09$  vs.  $0.32 \pm 0.45$ ,  $p = 0.040$ ) with a medium effect size (rank correlation = 0.53). This suggests that the increased induced low-beta power is associated with the degree of prediction error when we controlled the effect of SSA to be the same in both sessions.

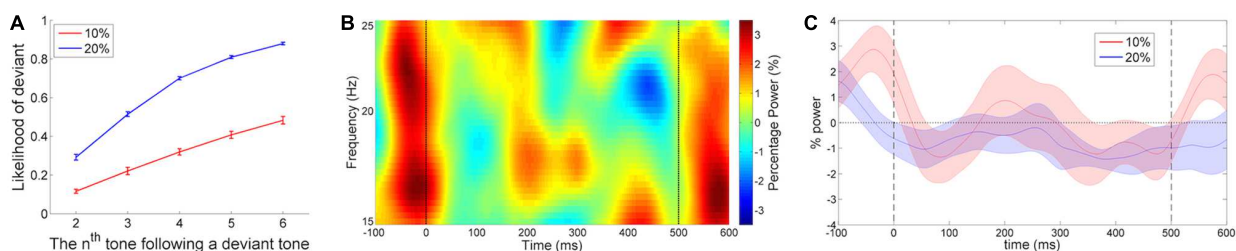
In sum, we showed that the deviant tone induced an increase in power in the low-beta band around 200–300 ms following tone onset in right auditory cortex, regardless of the presentation rate. Also, the effect was stronger when the deviance occurrence rate was lower. Furthermore, two additional analyses suggest that the induced low-beta power was higher for standard tones that violated a stronger prediction for a deviant tone, confirming that the low-beta response is more likely to reflect prediction error than response to novelty. Also, the induced low-beta power response was larger on deviant trials when they were less

predictable, even when the effects of SSA were controlled, again suggesting that the low-beta response to deviant tones reflected processes associate with prediction.

## DISCUSSION

We sought to understand the roles of beta oscillations in entrainment to rhythmically predictable sequences by introducing occasional unpredictable pitch deviants. We replicated previous findings related to timing entrainment in induced beta power (Snyder and Large, 2005; Fujioka et al., 2009, 2012, 2015; Iversen et al., 2009; Cirelli et al., 2014), showing that fluctuations in beta power entrained to the rate of presented isochronous auditory stimulus sequences in both left and right auditory cortices. In addition, we found that induced beta band power at right auditory cortex increased around 200–300 ms after the onsets of deviant tones compared to standard tones, especially in the low-beta range (15–20 Hz). This effect was larger when the deviant pitch was less likely to occur (10% vs. 20%), suggesting it is related to prediction processes. The right lateralization of the beta response to pitch deviants is consistent with the idea that the right auditory cortex is more sensitive for processing spectral information than its left counterpart (e.g., Zatorre et al., 1992, 2002). To the best of our knowledge, this is the first study to show that induced beta power in auditory cortex is sensitive to an unpredicted pitch change, even when it is presented at the predicted time. This suggests that induced beta power plays a role in sensory prediction for both *what* will occur as well as *when* it will occur.

The increased beta response with decreased likelihood of deviance occurrence indicates that beta oscillations may associate with precision-weighted prediction error. It has been suggested that while prediction error signals do not necessarily involve attention, high precision-weighted prediction errors act through attention to increase the gain of neural responses, acting as teaching signals for subsequent prediction updating (Friston,



**FIGURE 6 | The cumulative conditional likelihoods of encountering a deviant tone, and the time-frequency maps of induced difference (deviant minus standard) responses on matched trial locations in the beta frequency range (15–25 Hz) at right auditory cortex between the 10% and 20% sessions of Fast Experiment. (A)** The cumulative conditional likelihoods of encountering a deviant tone as a function of the  $n^{\text{th}}$  location following a deviant trial in 10% session (red) and 20% session (blue) with error bar indicating SEM. This was calculated by summing up the empirical occurrence rates of deviant tones at the current location and all preceding locations in the experiment. The likelihood of a deviant tone being presented at the  $n^{\text{th}}$  location is the accumulation of the occurrence rate from the first to  $n^{\text{th}}$  location following the previous deviant trial. **(B)** The subtraction of the two difference maps in the 10% session (deviant minus standard) minus the 20% session (deviant minus standard) at the second to the sixth trial following a deviant tone. Although the cluster-based permutation test did not find any cluster to be significantly different, the maximum of low-beta power difference (deviant minus standard) within the 130 to 370 ms window, time-locked to stimulus onset, was significantly larger in the 10% session than in the 20% session. **(C)** The shaded areas indicate SEM of the averaged low-beta (15–20 Hz) power difference (deviant minus standard) fluctuations of 10% session (red) and 20% session (blue).

2009; den Ouden et al., 2012; Hohwy, 2012; Schröger et al., 2015). According to predictive coding theory, prediction error is defined as the sensory mismatch between the predicted and perceived stimuli, and precision is the inverse of the input variance of the context which determines whether or not to deploy attention for updating future predictions (den Ouden et al., 2012). For example, prediction precision is higher for standard tones in the 10% than 20% session, because on average there are fewer deviant tones intermixed in the same length of sequence in 10% than 20% sessions. Thus, larger beta power responses to deviants in the 10% compared to 20% session might indicate that the process involved is one of prediction precision. That beta oscillations are associated with deploying attention for improving perceptual performance is supported by attentional blink studies showing that enhanced phase synchronization in low-beta band among frontal-parietal-temporal regions involved in the attentional network is associated with improved behavioral performance for targets with abrupt onsets (Gross et al., 2004; Kranczioch et al., 2007). Further, it has also been suggested that gamma oscillations (> 30 Hz) reflect feed forward prediction error signals (Herrmann et al., 2004) while beta oscillations represent a subsequent feed back processing stage for updating prediction (Arnal and Giraud, 2012), again consistent with the idea presented here that low-beta is sensitive to the precision of prediction, and associates with attention and prediction updating.

The latency of the low-beta response also implies that it is likely associated with attention and prediction updating. The low-beta response to pitch deviants in our data was around 200–300 ms after tone onset, which was later than the well-studied MMN prediction error response in the time waveform ERP, which was around 100 to 120 ms (**Figure 2**), consistent with other studies employing rhythmic sequences with relatively fast IOIs (Näätänen et al., 2007; Pablos Martin et al., 2007; Fujioka et al., 2008; Matsuda et al., 2013; Hove et al., 2014). This suggests that the low-beta response reflects a processing stage that is later than detecting prediction error. Interestingly, the 200–300 ms timing of the beta band power response occurs around the same time as P3a (Regnault et al., 2001; Jongsma et al., 2004, see **Figure 2** for P3a latency), which is known to reflect exogenous attentional orienting and attentional updating (Friedman et al., 2001; Polich, 2007). The P3a and induced low-beta power likely reflect distinct neural responses because the P3a is phase-locked to stimulus onset and originates in the anterior cingulate cortex and related structures (Polich, 2007) while, in contrast, the induced low-beta power response is not phase locked to stimulus onset and is observed with a spatial filter located in auditory regions. However, the overlapped response latencies are consistent with the idea that attentional processing in frontal areas, reflected by P3a, interacts with prediction precision, and is associated with induced beta power in auditory cortex.

To further evaluate the idea that beta is associated with precision-weighted prediction error, it is important to consider the alternative possibility that the beta band power increases we observed following pitch deviants are simply a response to novelty in the preceding local context rather than prediction

error. Indeed, a number of studies in humans and other animals have shown effects of rare stimuli on both induced and evoked beta oscillations (Haenschel et al., 2000; Kisley and Cornwell, 2006; Hong et al., 2008; Fujioka et al., 2009; Pearce et al., 2010; Kopell et al., 2011). Our results strongly favor the idea that induced beta power associates with prediction rather than a simple response to rareness for two reasons. First, the induced power fluctuations of beta oscillation entrain to external isochronous tone sequences in the absence of deviants (Fujioka et al., 2012), which suggests that a primary function of induced beta power concerns temporal prediction rather than detecting rare events. Second, our analyses of standard tones showed that induced low-beta power responses were stronger after the onset of standard tones that were less likely to occur (i.e., the last standard tone occurring after an uninterrupted series of sequential standard tones, SpreD trials) than standard tones that were more likely to occur (i.e., standards occurring earlier in a sequence of standards, Sbs trials). This confirms that increased induced low-beta power after tone onset reflects a process that is sensitive to the precision of prediction error.

Our results also suggest that the low-beta response is associated with precision-weighted prediction error while controlling possible effects of SSA. Previous studies on adaptation show that the neural response decreases to repeated tones, and that an increased response to the presentation of a new (rare) tone in an oddball context could reflect a release from this adaptation (e.g., Butler, 1968; Näätänen et al., 1988; Lanting et al., 2013). By selecting the deviant trials in the 10% and 20% sessions that had a matched number of standard tones preceding them we equated any effects of SSA between sessions. The results showed that the low-beta response to a deviant tone was larger in the 10% session than in the 20% session even after SSA was equated. Thus, the lower conditional likelihoods of encountering a deviant tone in the 10% than 20% session associate with a larger low-beta response on deviant trials. This analysis suggests that the low-beta response associates with precision-weighted prediction error although there may also have been a smaller effect of stimulus adaptation. Further research is needed on this question (e.g., see Herrmann et al., 2013, 2014, 2015).

A remaining question concerns the relation between prediction of rhythmic timing (Fujioka et al., 2012, 2015) and prediction precision for pitch, given that induced beta power is interactively modulated by both factors. Here we propose that timing and content (*when* and *what*) interact through attentional processing. Dynamic attending theory proposes that internal rhythmic entrainment to external temporal regularities is accomplished by a combination of self-sustained neural oscillation and the dynamic allocation of attention in the temporal dimension (Jones and Boltz, 1989; Large and Jones, 1999; Jones, 2010). The self-sustained oscillation acts as a time frame, and adapts its rate and phase to the external auditory rhythm. Attention increases at important time points such as the onset of beats, which is guided by the temporal prediction of the oscillatory time frame, and reflects temporal prediction for upcoming events during rhythmic entrainment. This attentional rhythmic entrainment is characterized as exogenous orienting



(Jones et al., 2006; Nobre et al., 2007; Coull and Nobre, 2008; Jones, 2010), which is involuntary and automatic (Rohenkohl et al., 2011; Triviño et al., 2011; Correa et al., 2014). Further, an MEG study has shown that the mathematical model of dynamic attending theory predicts delta power activities generated in auditory cortex (Herrmann et al., 2016), suggesting that rhythmic attending modulates oscillatory activities in auditory cortex. In this way, it is possible that rhythmic beta power fluctuations representing attention to events with temporal regularity increase perceptual processing of the content of the input stream at predictable time points, such as beat onsets. The idea that beta oscillations reflect temporal attention is also consistent with converging evidence that similar processes occur in the motor system, where rhythmic temporal structure also plays a critical role (e.g., Nobre et al., 2007; Coull and Nobre, 2008; Morillon et al., 2015). This is particularly interesting given that an auditory rhythm sets up beta power oscillations not only in auditory cortex, but also in motor areas even though movement is not involved. Thus, beta power oscillations in response to a rhythmic auditory input have also been interpreted as reflecting communication between auditory and motor system in the cortex (Jenkinson and Brown, 2011; Fujioka et al., 2012, 2015; Kilavik et al., 2013).

A lack of concurrent behavioral measurements to confirm whether induced beta power modulates perceptual sensitivity is a limitation of the current study. Further experiments are needed to examine this directly. However, the evidence to date shows that increased beta power before a stimulus onset reflects enhanced predictive readiness and improves perceptual performance. Studies using an auditory spatial temporal order judgments task (Bernasconi et al., 2011), an auditory temporal delay detection task (Arnal et al., 2015), intensity detection task (Herrmann et al., 2016), pitch distortion detection task during music listening (Doelling and Poeppel, 2015), or an audiovisual temporal integration task (Geerligs and Akyürek, 2012), all show that when the beta band power happened to be larger in the pre-stimulus period, participants made more accurate judgments or had enhanced audiovisual integration compared to when beta power was smaller. Together, the results of these studies are consistent with our speculation that beta oscillations reflect attention (Wróbel, 2000; Buschman and Miller, 2007, 2009).

## REFERENCES

- Arnal, L. H., Doelling, K. B., and Poeppel, D. (2015). Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cereb. Cortex* 25, 3077–3085. doi: 10.1093/cercor/bhu103
- Arnal, L. H., and Giraud, A. L. (2012). Cortical oscillations and sensory predictions. *Trends Cogn. Sci.* 16, 390–398. doi: 10.1016/j.tics.2012.05.003
- Bernasconi, F., Manuel, A. L., Murray, M. M., and Spierer, L. (2011). Pre-stimulus beta oscillations within left posterior sylvian regions impact auditory temporal order judgment accuracy. *Int. J. Psychophysiol.* 79, 244–248. doi: 10.1016/j.ijpsycho.2010.10.017
- Bertrand, O., Bohorquez, J., and Pernier, J. (1994). Time-frequency digital filtering based on an invertible wavelet transform: an application to evoked potentials. *IEEE Trans. Biomed. Eng.* 41, 77–88. doi: 10.1109/10.277274

## CONCLUSION

We presented isochronous auditory oddball sequences containing occasional pitch deviants to show that induced beta power is sensitive to the content of the input during rhythmic entrainment. We replicated previous findings that induced beta power entrains to externally presented rhythms. More interestingly, we showed that unpredicted pitch deviants modulate beta power 200–300 ms after deviant tone onsets, and that the magnitude of the modulation reflects the deviant occurrence likelihood (precision-weighted prediction error). Our data show that induced beta power activities in auditory cortex are consistent with a role in sensory prediction for both *what* (pitch) will occur as well as *when* (rhythm) events will occur. The timing and nature of the beta power response to pitch deviants suggests that it reflects an attentional modulation. In conjunction with other research, we propose that predictions for *what* and *when* are dynamically processed through attentional networks, and that beta oscillations in auditory cortex reflect the functional significance of sensory prediction and prediction error processes.

## AUTHOR CONTRIBUTIONS

AC, DB, and LT designed research; AC performed research; AC, DB, and LT contributed unpublished analytic tools; AC and DB analyzed data; AC, DB, and LT wrote the paper.

## FUNDING

This work was supported by a grant from the Canadian Institutes of Health Research (MOP 115043 to LT). AC was supported by a graduate student award from the Natural Sciences and Engineering Research Council of Canada CREATE grant in Auditory Cognitive Neuroscience (371324-2009).

## ACKNOWLEDGMENTS

We thank Dr. Ian C. Bruce for advice on signal processing, Dave Thompson for technical assistance and Alexandra Rice for assisting with data collection.

- Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., McKhann, G. M., et al. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. *J. Neurosci.* 31, 3176–3185. doi: 10.1523/JNEUROSCI.4518-10.2011
- Buschman, T. J., and Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862. doi: 10.1126/science.1138071
- Buschman, T. J., and Miller, E. K. (2009). Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. *Neuron* 63, 386–396. doi: 10.1016/j.neuron.2009.06.020
- Butler, R. A. (1968). Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. *J. Acoust. Soc. Am.* 44, 945–950. doi: 10.1121/1.1911233
- Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press.

- Calderone, D. J., Lakatos, P., Butler, P. D., and Castellanos, F. X. (2014). Entrainment of neural oscillations as a modifiable substrate of attention. *Trends Cogn. Sci.* 18, 300–309. doi: 10.1016/j.tics.2014.02.005
- Cirelli, L. K., Bosnyak, D., Manning, F. C., Spinelli, C., Marie, C., Fujioka, T., et al. (2014). Beat-induced fluctuations in auditory cortical beta-band activity: using EEG to measure age-related changes. *Front. Psychol.* 5:742. doi: 10.3389/fpsyg.2014.00742
- Correa, Á., Cona, G., Arbula, S., Vallesi, A., and Bisiacchi, P. (2014). Neural dissociation of automatic and controlled temporal preparation by transcranial magnetic stimulation. *Neuropsychologia* 65, 131–136. doi: 10.1016/j.neuropsychologia.2014.10.023
- Costa-Faidella, J., Baldeweg, T., Grimm, S., and Escera, C. (2011). Interactions between “what” and “when” in the auditory system: temporal predictability enhances repetition suppression. *J. Neurosci.* 31, 18590–18597. doi: 10.1523/JNEUROSCI.2599-11.2011
- Coull, J. T., and Nobre, A. C. (2008). Dissociating explicit timing from temporal expectation with fMRI. *Curr. Opin. Neurobiol.* 18, 137–144. doi: 10.1016/j.conb.2008.07.011
- den Ouden, H. E., Kok, P., and De Lange, F. P. (2012). How prediction errors shape perception, attention, and motivation. *Front. Psychol.* 3:548. doi: 10.3389/fpsyg.2012.00548
- Ding, N., Melloni, L., Zhang, H., Tian, X., and Poeppel, D. (2015). Cortical tracking of hierarchical linguistic structures in connected speech. *Nat. Neurosci.* 19, 158–164. doi: 10.1038/nn.4186
- Doelling, K. B., and Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proc. Natl. Acad. Sci. U.S.A.* 112, E6233–E6242. doi: 10.1073/pnas.1508431112
- Friedman, D., Cycowicz, Y. M., and Gaeta, H. (2001). The novelty P3: an event-related brain potential (ERP) sign of the brain's evaluation of novelty. *Neurosci. Biobehav. Rev.* 25, 355–373. doi: 10.1016/S0149-7634(01)00019-7
- Friston, K. (2005). A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 360, 815–836. doi: 10.1098/rstb.2005.1622
- Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13, 293–301. doi: 10.1016/j.tics.2009.04.005
- Fritz, J., Shamma, S., Elhilali, M., and Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nat. Neurosci.* 6, 1216–1223. doi: 10.1038/nn1141
- Fujioka, T., Ross, B., and Trainor, L. J. (2015). Beta-band oscillations represent auditory beat and its metrical hierarchy in perception and imagery. *J. Neurosci.* 35, 15187–15198. doi: 10.1523/JNEUROSCI.2397-15.2015
- Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Ann. N. Y. Acad. Sci.* 1169, 89–92. doi: 10.1111/j.1749-6632.2009.04779.x
- Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *J. Neurosci.* 32, 1791–1802. doi: 10.1523/JNEUROSCI.4107-11.2012
- Fujioka, T., Trainor, L. J., and Ross, B. (2008). Simultaneous pitches are encoded separately in auditory cortex: an MMNm study. *Neuroreport* 19, 361–366. doi: 10.1097/WNR.0b013e3282f51d91
- Geerligs, L., and Akyürek, E. G. (2012). Temporal integration depends on increased prestimulus beta band power. *Psychophysiology* 49, 1632–1635. doi: 10.1111/j.1469-8986.2012.01453.x
- Godey, B., Schwartz, D., De Graaf, J. B., Chauvel, P., and Liegeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clin. Neurophysiol.* 112, 1850–1859. doi: 10.1016/S1388-2457(01)00636-8
- Gomez-Ramirez, M., Kelly, S. P., Molholm, S., Sehatpour, P., Schwartz, T. H., and Foxe, J. J. (2011). Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: a human electrocorticographic investigation. *J. Neurosci.* 31, 18556–18567. doi: 10.1523/JNEUROSCI.2164-11.2011
- Griffiths, T. D., and Warren, J. D. (2004). What is an auditory object? *Nat. Rev. Neurosci.* 5, 887–892. doi: 10.1038/nrn1538
- Gross, J., Schmitz, F., Schnitzler, L., Kessler, K., Shapiro, K., Hommel, B., et al. (2004). Modulation of long-range neural synchrony reflects temporal limitations of visual attention in humans. *Proc. Natl. Acad. Sci. U.S.A.* 101, 13050–13055. doi: 10.1073/pnas.0404944101
- Haenschel, C., Baldeweg, T., Croft, R. J., Whittington, M., and Gruzelier, J. (2000). Gamma and beta frequency oscillations in response to novel auditory stimuli: a comparison of human electroencephalogram (EEG) data with in vitro models. *Proc. Natl. Acad. Sci. U.S.A.* 97, 7645–7650. doi: 10.1073/pnas.120162397
- Henry, M. J., and Herrmann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. *Timing Time Percept.* 2, 62–86. doi: 10.1163/22134468-00002011
- Henry, M. J., and Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20095–20100. doi: 10.1073/pnas.1213390109
- Herrmann, B., Henry, M. J., Fromboluti, E. K., McAuley, J. D., and Obleser, J. (2015). Statistical context shapes stimulus-specific adaptation in human auditory cortex. *J. Neurophysiol.* 113, 2582–2591. doi: 10.1152/jn.00634.2014
- Herrmann, B., Henry, M. J., Haegens, S., and Obleser, J. (2016). Temporal expectations and neural amplitude fluctuations in auditory cortex interactively influence perception. *Neuroimage* 124, 487–497. doi: 10.1016/j.neuroimage.2015.09.019
- Herrmann, B., Henry, M. J., and Obleser, J. (2013). Frequency-specific adaptation in human auditory cortex depends on the spectral variance in the acoustic stimulation. *J. Neurophysiol.* 109, 2086–2096. doi: 10.1152/jn.00907.2012
- Herrmann, B., Schlichting, N., and Obleser, J. (2014). Dynamic range adaptation to spectral stimulus statistics in human auditory cortex. *J. Neurosci.* 34, 327–331. doi: 10.1523/JNEUROSCI.3974-13.2014
- Herrmann, C. S., Munk, M. H., and Engel, A. K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends Cogn. Sci.* 8, 347–355. doi: 10.1016/j.tics.2004.06.006
- Hohwy, J. (2012). Attention and conscious perception in the hypothesis testing brain. *Front. Psychol.* 3:96. doi: 10.3389/fpsyg.2012.00096
- Hong, L. E., Buchanan, R. W., Thaker, G. K., Shepard, P. D., and Summerfelt, A. (2008). Beta (~16 Hz) frequency neural oscillations mediate auditory sensory gating in humans. *Psychophysiology* 45, 197–204. doi: 10.1111/j.1469-8986.2007.00624.x
- Hove, M. J., Marie, C., Bruce, I. C., and Trainor, L. J. (2014). Superior time perception for lower musical pitch explains why bass-ranged instruments lay down musical rhythms. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10383–10388. doi: 10.1073/pnas.1402039111
- Iversen, J. R., Repp, B. H., and Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Ann. N. Y. Acad. Sci.* 1169, 58–73. doi: 10.1111/j.1749-6632.2009.04579.x
- Jenkinson, N., and Brown, P. (2011). New insights into the relationship between dopamine, beta oscillations and motor function. *Trends Neurosci.* 34, 611–618. doi: 10.1016/j.tins.2011.09.003
- Jones, M. R. (2010). “Attending to sound patterns and the role of entrainment,” in *Attention and Time*, eds A. C. Nobre and J. T. Coull (New York, NY: Oxford University Press), 317–330.
- Jones, M. R., and Boltz, M. (1989). Dynamic attending and responses to time. *Psychol. Rev.* 96, 459–491. doi: 10.1037/0033-295X.96.3.459
- Jones, M. R., Johnston, H. M., and Puente, J. (2006). Effects of auditory pattern structure on anticipatory and reactive attending. *Cogn. Psychol.* 53, 59–96. doi: 10.1016/j.cogpsych.2006.01.003
- Jongsma, M. L., Desain, P., and Honing, H. (2004). Rhythmic context influences the auditory evoked potentials of musicians and nonmusicians. *Biol. Psychol.* 66, 129–152. doi: 10.1016/j.biopsycho.2003.10.002
- Kalcher, J., and Pfurtscheller, G. (1995). Discrimination between phase-locked and non-phase-locked event-related EEG activity. *Electroencephalogr. Clin. Neurophysiol.* 94, 381–384. doi: 10.1016/0013-4694(95)00040-6
- Kilavik, B. E., Zaepffel, M., Brovelli, A., MacKay, W. A., and Riehle, A. (2013). The ups and downs of beta oscillations in sensorimotor cortex. *Exp. Neurol.* 245, 15–26. doi: 10.1016/j.expneurol.2012.09.014
- King, A. J., and Nelken, I. (2009). Unraveling the principles of auditory cortical processing: can we learn from the visual system? *Nat. Neurosci.* 12, 698–701. doi: 10.1038/nn.2308
- Kisley, M. A., and Cornwell, Z. M. (2006). Gamma and beta neural activity evoked during a sensory gating paradigm: effects of auditory, somatosensory and cross-modal stimulation. *Clin. Neurophysiol.* 117, 2549–2563. doi: 10.1016/j.clinph.2006.08.003

- Kopell, N., Whittington, M. A., and Kramer, M. A. (2011). Neuronal assembly dynamics in the beta1 frequency range permits short-term memory. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3779–3784. doi: 10.1073/pnas.1019676108
- Kranczioch, C., Debener, S., Maye, A., and Engel, A. K. (2007). Temporal dynamics of access to consciousness in the attentional blink. *Neuroimage* 37, 947–955. doi: 10.1016/j.neuroimage.2007.05.044
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., and Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320, 110–113. doi: 10.1126/science.1154735
- Lakatos, P., Musacchia, G., O'Connell, M. N., Falchier, A. Y., Javitt, D. C., and Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron* 77, 750–761. doi: 10.1016/j.neuron.2012.11.034
- Lanting, C. P., Briley, P. M., Sumner, C. J., and Krumbholz, K. (2013). Mechanisms of adaptation in human auditory cortex. *J. Neurophysiol.* 110, 973–983. doi: 10.1152/jn.00547.2012
- Large, E. W., and Jones, M. R. (1999). The dynamics of attending: how people track time-varying events. *Psychol. Rev.* 106, 119–159. doi: 10.1037/0033-295X.106.1.119
- Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. New York, NY: Oxford University Press.
- Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG and MEG-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024
- Matsuda, A., Hara, K., Miyajima, M., Matsushima, E., Ohta, K., and Matsuura, M. (2013). Distinct pre-attentive responses to non-scale notes: an auditory mismatch negativity (MMN) study. *Clin. Neurophysiol.* 124, 1115–1121. doi: 10.1016/j.clinph.2012.12.002
- Morillon, B., Hackett, T. A., Kajikawa, Y., and Schroeder, C. E. (2015). Predictive motor control of sensory dynamics in auditory active sensing. *Curr. Opin. Neurobiol.* 31, 230–238. doi: 10.1016/j.conb.2014.12.005
- Morillon, B., and Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Ann. N. Y. Acad. Sci.* 1337, 26–31. doi: 10.1111/nyas.12629
- Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi: 10.1016/j.clinph.2007.04.026
- Näätänen, R., and Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology* 24, 375–425. doi: 10.1111/j.1469-8986.1987.tb00311.x
- Näätänen, R., Sams, M., Alho, K., Paavilainen, P., Reinikainen, K., and Sokolov, E. N. (1988). Frequency and location specificity of the human vertex N1 wave. *Electroencephalogr. Clin. Neurophysiol.* 69, 523–531. doi: 10.1016/0013-4694(88)90164-2
- Nobre, A. C., Correa, Á., and Coull, J. T. (2007). The hazards of time. *Curr. Opin. Neurobiol.* 17, 465–470. doi: 10.1016/j.conb.2007.07.006
- Pablos Martin, X., Deltenre, P., Hoonhorst, I., Markessis, E., Rossion, B., and Colin, C. (2007). Perceptual biases for rhythm: the Mismatch Negativity latency indexes the privileged status of binary vs. non-binary interval ratios. *Clin. Neurophysiol.* 118, 2709–2715. doi: 10.1016/j.clinph.2007.08.019
- Pearce, M. T., Ruiz, M. H., Kapasi, S., Wiggins, G. A., and Bhattacharya, J. (2010). Unsupervised statistical learning underpins computational, behavioural, and neural manifestations of musical expectation. *Neuroimage* 50, 302–313. doi: 10.1016/j.neuroimage.2009.12.019
- Pfurtscheller, G., and Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–1857. doi: 10.1016/S1388-2457(99)00141-8
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Regnault, P., Bigand, E., and Besson, M. (2001). Different brain mechanisms mediate sensitivity to sensory consonance and harmonic context: evidence from auditory event-related brain potentials. *J. Cogn. Neurosci.* 13, 241–255. doi: 10.1162/089892901564298
- Rohenkohl, G., Coull, J. T., and Nobre, A. C. (2011). Behavioural dissociation between exogenous and endogenous temporal orienting of attention. *PLoS ONE* 6:e14620. doi: 10.1371/journal.pone.0014620
- Saleh, M., Reimer, J., Penn, R., Ojakangas, C. L., and Hatsopoulos, N. G. (2010). Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. *Neuron* 65, 461–471. doi: 10.1016/j.neuron.2010.02.001
- Schroeder, C. E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 32, 9–18. doi: 10.1016/j.tins.2008.09.012
- Schröger, E., Marzecová, A., and SanMiguel, I. (2015). Attention and prediction in human audition: a lesson from cognitive psychophysiology. *Eur. J. Neurosci.* 41, 641–664. doi: 10.1111/ejn.12816
- Schwartz, M., and Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neurosci. Biobehav. Rev.* 37, 2587–2596. doi: 10.1016/j.neubiorev.2013.08.005
- Snyder, J. S., and Large, E. W. (2005). Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Cogn. Brain Res.* 24, 117–126. doi: 10.1016/j.cogbrainres.2004.12.014
- Triviño, M., Arnedo, M., Lupiáñez, J., Chirivella, J., and Correa, Á. (2011). Rhythms can overcome temporal orienting deficit after right frontal damage. *Neuropsychologia* 49, 3917–3930. doi: 10.1016/j.neuropsychologia.2011.10.009
- Wróbel, A. (2000). Beta activity: a carrier for visual attention. *Acta Neurobiol. Exp.* 60, 247–260.
- Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cogn. Sci.* 6, 37–46. doi: 10.1016/S1364-6613(00)01816-7
- Zatorre, R. J., Evans, A. C., Meyer, E., and Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256, 846–849. doi: 10.1126/science.1589767

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Chang, Bosnyak and Trainor. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Large Scale Functional Brain Networks Underlying Temporal Integration of Audio-Visual Speech Perception: An EEG Study

G. Vinodh Kumar<sup>1</sup>, Tamesh Halder<sup>1</sup>, Amit K. Jaiswal<sup>1</sup>, Abhishek Mukherjee<sup>1</sup>,  
Dipanjan Roy<sup>2</sup> and Arpan Banerjee<sup>1\*</sup>

<sup>1</sup> Cognitive Brain Lab, National Brain Research Centre, Gurgaon, India, <sup>2</sup> Centre for Behavioural and Cognitive Sciences, University of Allahabad, Allahabad, India

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College, USA

### Reviewed by:

John Magnotti,  
Baylor College of Medicine, USA  
Julian Keil,  
Charité-Universitätsmedizin Berlin,  
Germany

### \*Correspondence:

Arpan Banerjee  
arpan@nbrn.ac.in

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 11 February 2016

**Accepted:** 23 September 2016

**Published:** 13 October 2016

### Citation:

Kumar GV, Halder T, Jaiswal AK,  
Mukherjee A, Roy D and Banerjee A  
(2016) Large Scale Functional Brain  
Networks Underlying Temporal  
Integration of Audio-Visual Speech  
Perception: An EEG Study.  
Front. Psychol. 7:1558.  
doi: 10.3389/fpsyg.2016.01558

Observable lip movements of the speaker influence perception of auditory speech. A classical example of this influence is reported by listeners who perceive an illusory (cross-modal) speech sound (McGurk-effect) when presented with incongruent audio-visual (AV) speech stimuli. Recent neuroimaging studies of AV speech perception accentuate the role of frontal, parietal, and the integrative brain sites in the vicinity of the superior temporal sulcus (STS) for multisensory speech perception. However, if and how does the network across the whole brain participates during multisensory perception processing remains an open question. We posit that a large-scale functional connectivity among the neural population situated in distributed brain sites may provide valuable insights involved in processing and fusing of AV speech. Varying the psychophysical parameters in tandem with electroencephalogram (EEG) recordings, we exploited the trial-by-trial perceptual variability of incongruent audio-visual (AV) speech stimuli to identify the characteristics of the large-scale cortical network that facilitates multisensory perception during synchronous and asynchronous AV speech. We evaluated the spectral landscape of EEG signals during multisensory speech perception at varying AV lags. Functional connectivity dynamics for all sensor pairs was computed using the time-frequency global coherence, the vector sum of pairwise coherence changes over time. During synchronous AV speech, we observed enhanced global gamma-band coherence and decreased alpha and beta-band coherence underlying cross-modal (illusory) perception compared to unisensory perception around a temporal window of 300–600 ms following onset of stimuli. During asynchronous speech stimuli, a global broadband coherence was observed during cross-modal perception at earlier times along with pre-stimulus decreases of lower frequency power, e.g., alpha rhythms for positive AV lags and theta rhythms for negative AV lags. Thus, our study indicates that the temporal integration underlying multisensory speech perception requires to be understood in the framework of large-scale functional brain network mechanisms in addition to the established cortical loci of multisensory speech perception.

**Keywords:** EEG, AV, multisensory, perception, functional connectivity, coherence, temporal synchrony, integration



## INTRODUCTION

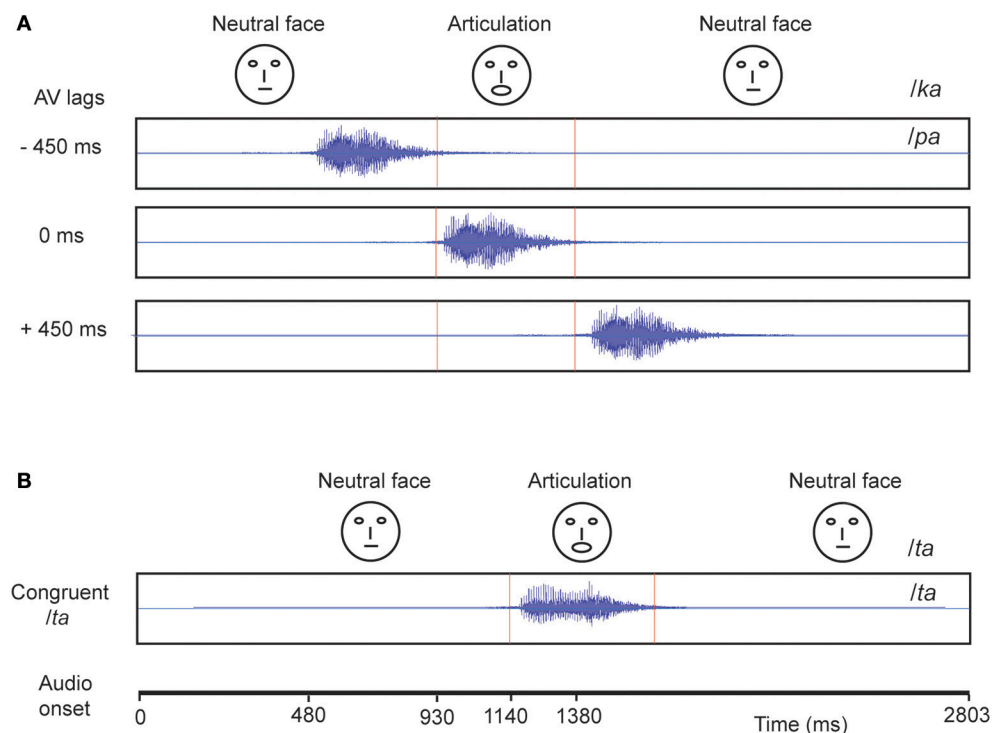
Perception of the external world involves the efficient integration of information over multiple sensory systems (Wallace et al., 1993). During speech perception, visual cues from the speaker's face enhances the intelligibility of auditory signal (Sumbly, 1954; Helfer, 1997; Bulkin and Groh, 2006). Also, the incidence of specific *semantically-incongruent* visual information modulates auditory perception, for example, an auditory speech sound /ba/ superimposed with a speaker's lip movement of /ga/, gives rise to a perception of /da/ (McGurk and Macdonald, 1976). Similarly, an incongruent AV combination of /pa/-/ka/ elicits an 'illusory' (cross-modal) percept /ta/ (McGurk and Macdonald, 1976; MacDonald and McGurk, 1978; van Wassenhove et al., 2007). However, such multisensory-mediated effects are influenced by the relative timing of the auditory and visual inputs (Stein et al., 1989; Munhall et al., 1996; Sekuler et al., 1997; van Atteveldt et al., 2007; van Wassenhove et al., 2007). Consequently, the temporal processing of the incoming multiple sensory (auditory and visual) information and their integration to yield a crossmodal percept is pivotal for speech perception (Deroy et al., 2014). Where and how the underlying information processing takes place is subject of several research studies which we review in the following paragraph. Cortical and sub-cortical regions and functional brain networks with specific patterns of connectivity becomes the prime target for these investigations. In a nutshell, characterization of the multi-scale representational space of temporal processing underlying multisensory stimuli is an open question to the community.

As we discuss in the following paragraph, a dominant strategy in multisensory research is the search for loci comprising of brain areas that are responsible for triggering the multisensory experience (Jones and Callan, 2003; Beauchamp, 2010; Nath and Beauchamp, 2013). However, from the perspective of functional integration (Bressler, 1995; Bressler and Menon, 2010) understanding the large-scale network organization underlying the temporal processes is a critical component of formulating a comprehensive theory of multisensory speech perception. Numerous neuroimaging and electrophysiological studies have explored the neural mechanism that underpins audio-visual integration employing McGurk effect (Wallace et al., 1993; Jones and Callan, 2003; Sekiyama et al., 2003; Kaiser, 2004; van Wassenhove et al., 2005; Hasson et al., 2007; Saint-Amour et al., 2007; Skipper et al., 2007; Stevenson et al., 2010; Keil et al., 2012; Nath and Beauchamp, 2013). A majority of these studies accentuate the role of primary auditory and visual cortices, multisensory areas such as posterior superior temporal sulcus (pSTS) (Jones and Callan, 2003; Sekiyama et al., 2003; Nath and Beauchamp, 2011, 2013) and other brain regions including frontal and parietal areas (Callan et al., 2003; Skipper et al., 2007) in the perception of the illusion. In particular, the electrophysiological evidences primarily emphasizes the significance of beta (Keil et al., 2012; Roa Romero et al., 2015) and gamma band activity (Kaiser, 2004) toward illusory (cross-modal) perceptual experience. Source-level functional connectivity among brain areas employing phase synchrony measures, reveal interactions among cortical regions of interest

(left Superior Temporal Gyrus) and the whole brain that correlates with cross-modal perception (Keil et al., 2012). These studies either reveal the activations in the cortical loci or the functional connectedness to particular cortical regions of interest that are elemental for the illusory percept. On the other hand, the role of timing between auditory and visual components in AV speech stimuli has been studied from the perspective of the main modules in multisensory processing (Jones and Callan, 2003). Recently, we have addressed this issue using a dynamical systems model to study the interactive effects between AV lags and underlying neural connectivity onto perception (Thakur et al., 2016). Interestingly, how these network are functionally connected in the context of behavioral performance or perceptual experience are increasingly being revealed (Nath and Beauchamp, 2011; Keil et al., 2012). Nonetheless, the identification and systematic characterization of these networks under cross-modal and unimodal perception is an open question.

A traditional measure of large-scale functional connectivity in EEG is the sensor-level global coherence (Cimenser et al., 2011; Balazs et al., 2015; Fonseca et al., 2015; Alba et al., 2016; Clarke et al., 2016). Global coherence can be described as either the normalized vector sum of all pairwise coherences between sensor combinations, the frequency domain representation of cross-correlation between two time-series (Lachaux et al., 1999; Cimenser et al., 2011) or the ratio of the largest eigenvalue of the cross-spectral matrix to the sum of its eigenvalues (Mitra and Bokil, 2008). An increased global coherence confirms the presence of a spatially extended network that spans over several EEG sensors, since local pairwise coherence would not survive statistical threshold after averaging. To the best of our knowledge, global coherence has not been used in the domain of audio-visual (AV) speech perception to evaluate the presence of whole brain networks. Furthermore, characterization of the differences in whole brain network organization underlying cross-modal vs. unimodal perceptual experience vis-à-vis the timing of sensory signals will be critical to understanding the neurobiology of multisensory perception.

In the current study, we used the incongruent McGurk pair (audio /pa/ superimposed on the video of the face articulating /ka/) to induce the illusory percept /ta/. Further, we generated a temporal asynchrony in the onset of audio and visual events of the McGurk pair to diminish the rate of cross-modal responses. Subsequently, we exploited the inter-trial perceptual variability to study integration both at behavioral levels by accounting perceptual response and eye-tracking as well as neural levels using EEG. We considered subjects' /pa/ responses as unimodal perception since it represents only one sensory stream and /ta/ responses as cross-modal perception since it represents an experience resulting from integrating features from two modalities (Deroy et al., 2014). We studied the spectral landscape of perceptual categorization as function of AV timing and found patterns that matched with previous reports. Finally, we evaluate the large-scale brain network organization dynamics using time-frequency global coherence analysis for studying perceptual categorization underlying different temporal processing scenarios at various AV lags. In the process, we



**FIGURE 1 | Stimuli** Each block represents a video. **(A)** The McGurk stimuli: Audio /pa/ superimposed on visual (lip movement) /ka/ was presented under different audio-visual (AV) lag scenarios. The location of onset of audio is varied with respect to a person's initiation of lip-movement /ka/ at -450, 0, and 450 ms. **(B)** In congruent /ta/ condition, audio /ta/ is presented synchronously with onset of lip movement /ta/.

reveal the complex spectro-temporal organization of networks underlying multisensory perception.

## MATERIALS AND METHODS

### Participants

Nineteen [10 males and 9 females, ranging from 22–29, (mean age 25;  $SD = 2$ )] healthy volunteers participated in the study. No participant had neurological or audiological problems. They all had normal or corrected-to-normal vision and were right handed. The study was carried out following the ethical guidelines and prior approval of Institutional Review Board of National Brain Research Centre, India.

### Stimuli and Trials

The experiment consisted of 360 trials overall in which we showed the videos of a male actor pronouncing the syllables /ta/ and /ka/ (Figure 1). One-fourth of the trials consisted of congruent video (visual /ta/ auditory /ta/) and the remaining trials comprised incongruent videos (visual /ka/ auditory /pa/) presented in three audio-visual lags: -450 ms (audio lead), 0 ms (synchronous), +450 ms (audio lag), each comprising one-fourth of the overall trials. The stimuli were rendered into a  $800 \times 600$  pixels movie with a digitization rate of 29.97 frames per second. Stereo soundtracks were digitized at 48 kHz with 32 bit resolution. The stimuli were presented via Presentation software (Neurobehavioral System Inc.). The video was presented using a

17" LED monitor. Sounds were delivered at an overall intensity of  $\sim 60$  dB through sound tubes.

The experiment was carried out in three blocks each block consisting of 120 trials. Inter-trial intervals were pseudo-randomly varied between 1200 and 2800 ms. Each block comprised the four stimuli types (30 trials of each): Congruent video and three incongruent videos with the AV lags. The subjects were instructed to report what they heard while watching the articulator using a set of three keys. The three choices were /pa/, /ta/ and "anything else" (Other).

Post EEG scan, the participants further performed a behavioral task. The task comprised of 60 trials, comprising 30 trials each of auditory syllables /pa/ and /ta/. Participants were instructed to report their perception using a set of two keys while listening to syllables. The choices were /pa/ and /ta/.

### Data Acquisition and Analysis

#### EEG

EEG recordings were obtained using a Neuroscan system (Compumedics NeuroScan, SynAmps2) with 64 Ag/AgCl sintered electrodes mounted on an elastic cap of Neuroscan in a 10–20 montage. Data were acquired continuously in AC mode (sampling rate, 1 kHz). Reference electrodes were linked mastoids, grounded to AFz. Channel impedances were kept at  $< 5$  k $\Omega$ . All subsequent analysis was performed in adherence to guidelines set by Keil et al. (2014).

## Eye Tracking

Gaze fixations of participants on the computer screen were recorded by EyeTribe eye tracking camera with resolution 30 Hz (<https://theeyetribe.com/>). The gaze data were analyzed using customized MATLAB codes. The image frame of the speaker video was divided into 3 parts, the head, the nose and the mouth (Figure 2A). The gaze locations at these quadrants over the duration of stimulus presentation were converted into percentage measures for further statistical analysis.

## Pre-processing of EEG Signals

The collected EEG data were subsequently filtered using a bandpass of 0.2–45 Hz. Epochs of 400 and 900 ms before and after the onset of first stimuli (sound or articulation) were extracted and sorted based on the responses, /ta/, /pa/, and “other” respectively. Epochs were baseline corrected by removing the temporal mean of the EEG signal on an epoch-by-epoch basis. Epochs with maximum signal amplitude above 100  $\mu$ V or minimum below  $-100 \mu$ V were removed from all the electrodes to eliminate the response contamination from ocular and muscle-related activities. Approximately 70–75 % (~250 trials) trials of each subject were preserved after artifact rejection. In the final data analysis, a mean of 24 ( $SD = 9$ ), 18 ( $SD = 9$ ), and 25 ( $SD = 13$ ) incongruent trials at  $-450, 0, +450$  ms AV lags respectively in which the participants responded /pa/ were included. Similarly, a mean of 32 ( $SD = 15$ ), 42 ( $SD = 13$ ), and 32 ( $SD = 14$ ) incongruent trials at  $-450, 0, +450$  ms AV lags respectively in which the participants responded /ta/ were included in the final analyses. Approximately 2–6% of trials were excluded from each of the aforementioned trial categories. The response category with lowest number of occurrences was /pa/ at 0 ms AV lag with 270 hits from a total of 1350 trials across all volunteers ( $15 \times 90$ ). Subsequently, we randomly resampled 270 trials from /ta/ responses at 0 ms AV lag, and /pa/ and /ta/ responses at other AV lags. Thus, for each AV lag condition, 270 trials chosen randomly from the respective sorted response epochs (/pa/ or /ta/) entered the final analyses.

## Spectral Analysis

Power spectra of the preprocessed EEG signals at each electrode were computed on a single trial basis. We computed the spectral power at different frequencies using customized MATLAB ([www.mathworks.com](http://www.mathworks.com)) codes and the Chronux toolbox ([www.chronux.org](http://www.chronux.org)). Time bandwidth product and number of tapers were set at 3 and 5 respectively while using the Chronux function `mtspecgramc.m` to compute the power spectrum of the sorted time series in EEG data. Subsequently, the differences in the power during /ta/ and /pa/ responses at each AV lag were statistically compared by means of a cluster-based permutation test (Maris and Oostenveld, 2007) using the fieldtrip toolbox ([www.fieldtriptoolbox.org](http://www.fieldtriptoolbox.org)). The fieldtrip function `ft_freqstatistics.m` was used to perform the cluster computation. During the statistical comparison, an observed test statistic value below the threshold of 0.05 in at least 2 of the neighborhood channels were set for being considered in the cluster computation. Furthermore, 1000 iterations of trial randomization were carried out for generating the permutation

distribution at a frequency band. Subsequently, a two tailed test with a threshold of 0.025 was used for evaluating the sensors that exhibit significant difference in power. Statistical analysis was carried out separately for alpha (8–12 Hz), beta (13–30 Hz), and gamma (30–45 Hz) frequency ranges.

## Large-Scale Network Analysis

For deciphering the coordinated oscillatory brain network underlying the AV integration, we employed global coherence analyses (Bressler et al., 1993; Lachaux et al., 1999; Maris et al., 2007; Cimenser et al., 2011) on the perceptual categories (/ta/ and /pa/). A higher value of this measure will indicate the presence of strong large-scale functional networks. We computed the global coherence by decomposing information from the cross-spectral matrix employing the eigenvalue method (Mitra and Bokil, 2008). The cross-spectrum value at a frequency  $f$  between sensor pair  $i$  and  $j$  was computed as:

$$C_{ij}^X(f) = \frac{1}{K} \sum_{k=1}^K X_i^k(f) X_j^k(f)^* \quad (1)$$

where  $X_i^k$  and  $X_j^k$  are tapered Fourier transforms of the time series from the sensors  $i$  and  $j$  respectively, at the frequency  $f$ . A  $62 \times 62$  matrix of cross spectra, that represents all pairwise sensor combination, was computed in our case. Conversely, to characterize the dynamics of coordinated activity over time, we evaluated the time-frequency global coherogram. We employed the Chronux function `cohgramc.m` to obtain the time-frequency cross-spectral matrix for all the sensor combinations. Subsequently, for each trial we obtained the global coherence at each time point and frequency bin by computing the ratio of the largest eigenvalue of the cross-spectral matrix to the sum of the eigenvalues employing the following equation:

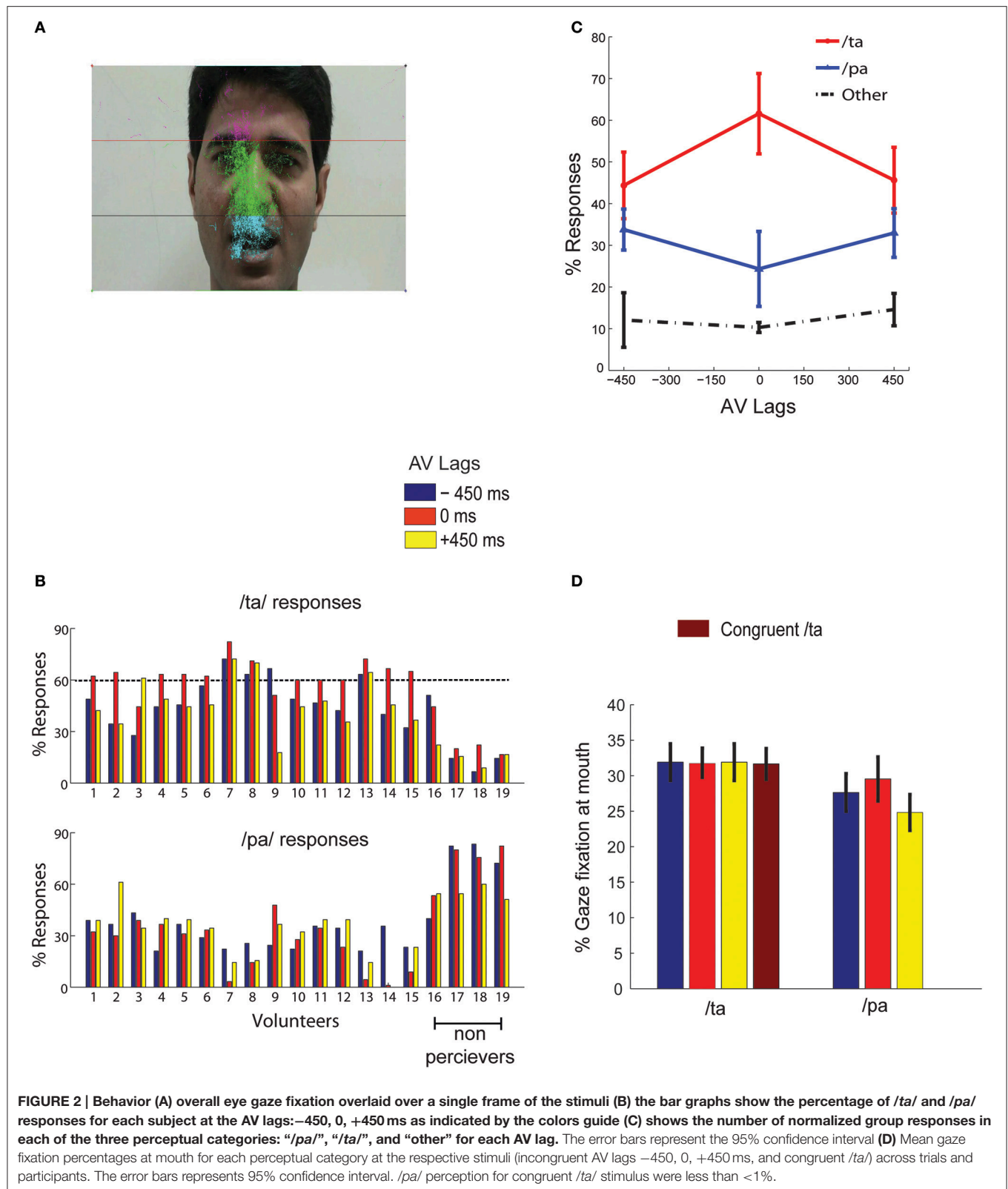
$$C_{Global}(f) = \frac{S_1^Y(f)}{\sum_{i=1}^n S_i^Y(f)} \quad (2)$$

where  $C_{Global}(f)$  is the global coherence,  $S_1^Y(f)$  is the largest eigenvalue and the denominator  $\sum_{i=1}^n S_i^Y(f)$  represents the sum of eigenvalues of the cross-spectral matrix (Cimenser et al., 2011). Time-frequency global coherogram computed for /ta/ and /pa/ responses were further compared at each time point for significant difference in different frequency bands (alpha, beta, and gamma) by means of cluster-based permutation test (Maris et al., 2007).

For every frequency bin at each time point, the coherence difference between /ta/ and /pa/ was evaluated using the Fisher's  $Z$  transformation

$$Z(f) = \frac{\tanh^{-1}(C_1(f)) - \tanh^{-1}(C_2(f)) - (\frac{1}{2m_1-2} - \frac{1}{2m_2-2})}{\sqrt{\frac{1}{2m_1-2} + \frac{1}{2m_2-2}}} \quad (3)$$

where  $2m_1, 2m_2$  = degrees of freedom;  $Z(f) \approx N(0, 1)$  a unit normal distribution; and  $C_1$  and  $C_2$  are the coherences at frequency  $f$ .



The coherence Z-statistic matrix obtained from the above computation formed the observed Z-statistics. Subsequently, from the distribution of observed Z-statistics, 5th and the 95th

quantile values were chosen as upper and lower threshold i.e., the values below and above the threshold values respectively were considered in the cluster computation. Based on spectral



adjacency (4–7 Hz, theta; 8–12 Hz, alpha; 13–30 Hz, beta; 30–45 Hz, gamma), clusters were selected at each time point. Consequently, cluster-level statistics were computed by taking the sum of positive and negative values within a cluster separately. Following the computation of the cluster-level statistics of the observed Z-statistics, 1000 iterations of trial randomization were carried out. For every iteration, cluster-level statistic was computed on the randomized trials to generate the permutation distribution. Subsequently, the values of observed cluster-level statistics were compared with the 2.5th and the 97.5th quantile values of the respective permutation distribution. The observed cluster-level statistics value that were below 2.5th and above 97.5th quantile consequently for two time points formed the negative and positive clusters respectively.

## RESULTS

### Behavior

Behavioral responses corresponding to McGurk stimuli with the AV lags were converted to percentage measures for each perceptual category (/pa/, /ta/, or “other”) from all subjects. We set a minimum threshold of 60% of /ta/ response in any AV lag, −450, 0, and +450 ms to qualify a participant as an illusory perceiver. 15 participants passed this threshold and 4 participants failed to perceive above the set threshold (see **Figure 2B**). Data from only 15 perceivers were used for further group level analysis. We observed that maximum percentage of illusory (/ta/) responses occurred at 0 ms AV lag when the lip movement of the speaker was synchronous with the onset of auditory stimulus (**Figure 2C**). Also, the percentage of /pa/ responses was minimum at 0 ms AV lag. We ran one-way ANOVAs on the percentage responses for /pa/, /ta/, and “other” with AV lags as the variable. We observed that AV lags influenced the percentage of /ta/ [ $F_{(2, 44)} = 27.68, p < 0.0001$ ] and /pa/ [ $F_{(2, 44)} = 5.89, p = 0.0056$ ] responses. However, there was no influence of AV lags on “other” responses [ $F_{(2, 44)} = 0.36, p = 0.700$ ]. We also performed paired Student’s *t*-test on the percentage of responses (/ta/ and /pa/) at each AV lag. Insignificant differences of 10.20–11.40% were observed between /ta/ and /pa/ responses at −450 ms AV lag [ $t_{(14)} = 0.63, p = 0.27$ ] and +450 ms AV lag [ $t_{(14)} = 0.45, p = 0.67$ ] respectively. However, at 0 ms AV lag we observed the percentage of /ta/ responses were significantly higher by 36.58% than the percentage of /pa/ responses,  $t_{(14)} = 10.20, p < 0.0001$ . Furthermore, the hit rate of /ta/ responses during congruent /ta/ was observed to be 0.97. Also, the hit rate of /ta/ and /pa/ during auditory alone conditions were observed to be 0.96 and 0.98 respectively.

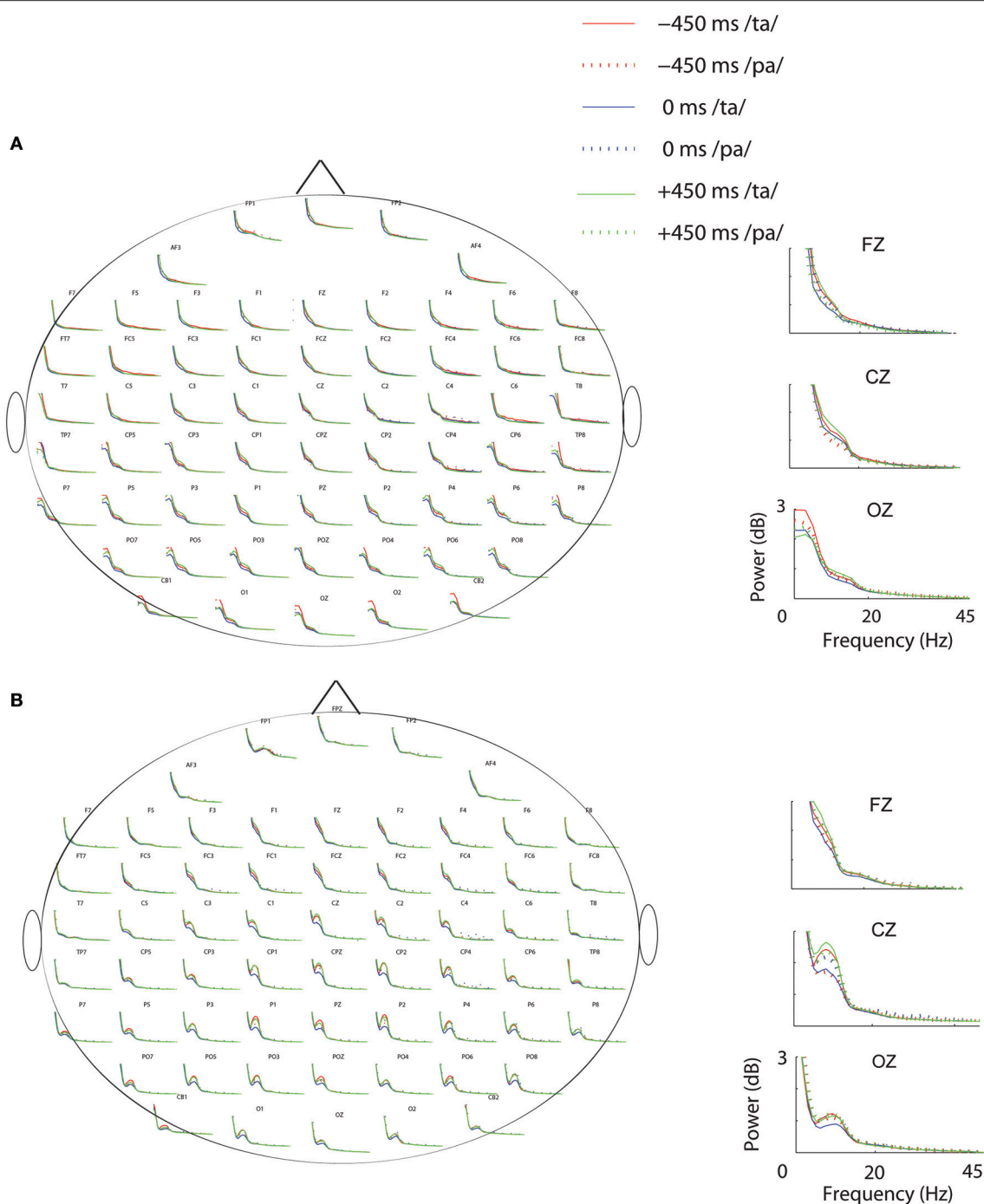
Gaze fixations at different locations on the speaker’s, head, nose and mouth areas were converted into percentage measures trial-by-trial for each subject and stimuli conditions. **Figure 2A** indicates that most of the gaze fixations were around head, nose, and mouth areas only. We ran a repeated measures 2-way ANOVA on mean gaze fixation percentages across trials at mouth areas with lags and perceived objects (/pa/ or /ta/) as variables. No significant differences were found for gaze fixations across lags [ $F_{(2, 89)} = 0, p = 0.95$ ] and perceptual categorization [ $F_{(1, 89)} = 1.33, p = 0.27$ ] as well as their interactions [ $F_{(2, 89)} = 0.01, p =$

0.85]. Number of /pa/ responses for congruent /ta/ stimulus was negligible (<1%), to do meaningful statistical comparisons. We also performed paired Student’s *t*-tests on the mean gaze fixation percentages for /pa/ and /ta/ responses at each lag. Increases in gaze fixation at mouth during /ta/ perception by 15.5 % at −450 ms AV lag [ $t_{(14)} = 0.90, p = 0.38$ ], 7.2 % at 0 ms AV lag [ $t_{(14)} = 0.90, p = 0.38$ ] and 28.54% at +450 ms AV lag [ $t_{(14)} = -0.32, p = 0.74$ ] (see **Figure 2D** for the mean values) were not statistically significant.

### Oscillatory Activity

Subsequent to replicating the perceptual (Munhall et al., 1996; van Wassenhove et al., 2007) and the eye gaze behavior (Gurler et al., 2015) results as reported earlier, the focus of interest was what differentiates the two perceptual states (/ta/ and /pa/) in terms of brain oscillations and large-scale functional brain networks. Therefore, spectral power at different frequency bands during /ta/ and /pa/ perception were compared at different AV lags. Power spectra at each sensor computed in the time window before (see **Figure 3A**) and after (see **Figure 3B**) the onset of first stimuli showed distinct changes in power for the two states. Cluster-based permutation tests employed for comparing the spectral power between the perceptual states show that /ta/ perception is associated with an overall suppression in power for all AV lags (see **Figure 4**). The magenta “\*” on the topoplots highlight the position of the negative clusters showing a significant suppression at 95% confidence levels in power. The blue areas on the scalp map highlight the regions that show decrease in the spectral power and the orange and red regions highlight the regions that show an increase in the spectral power. During the pre-stimulus period, one significant negative cluster [ $t_{(269)} = -2.04, p = 0.02$ ] over temporo-occipital sensors, two over frontal and occipital sensors [ $t_{(269)} = -3.57, p = 0.002$  and  $t_{(269)} = -3.14, p = 0.0002$ ] and one over occipital sensors [ $t_{(269)} = -2.18, p = 0.01$ ] were observed for alpha, beta, and gamma bands respectively in 0 ms AV lag (see **Figure 4A**). Also, one significant negative cluster over fronto-temporal and occipital sensors [ $t_{(269)} = -2.65, p = 0.004$ ], one over frontal and occipital sensors [ $t_{(269)} = -2.31, p = 0.01$ ] were observed at alpha and beta bands respectively during +450 ms AV lag (see **Figure 4C**). However, no significant difference was found during −450 ms AV lag.

Furthermore, during post-stimulus onset period, the /ta/-/pa/ comparison revealed one significant negative cluster over all sensors [ $t_{(269)} = -1.93, p = 0.02$ ], one over frontal, parietal, and occipital sensors [ $t_{(269)} = -2.70, p = 0.004$ ] and one over occipital sensors [ $t_{(269)} = -2.54, p = 0.006$ ] at alpha, beta, and gamma bands respectively during −450 ms AV lag (see **Figure 4C**). During 0 ms AV lag, one significant negative cluster [ $t_{(269)} = -2.22, p = 0.01$ ] spanning over all sensors and one over occipital sensors [ $t_{(269)} = -2.10, p = 0.02$ ] was observed at alpha and beta bands respectively (see **Figure 4D**). However, no significant difference in power between /ta/-/pa/ trials was observed during the post-stimulus period at +450 ms AV lag. Overall, significant spectral power was lower during /ta/ than /pa/ as reflected from cluster-based analysis during pre- and post-stimulus periods.

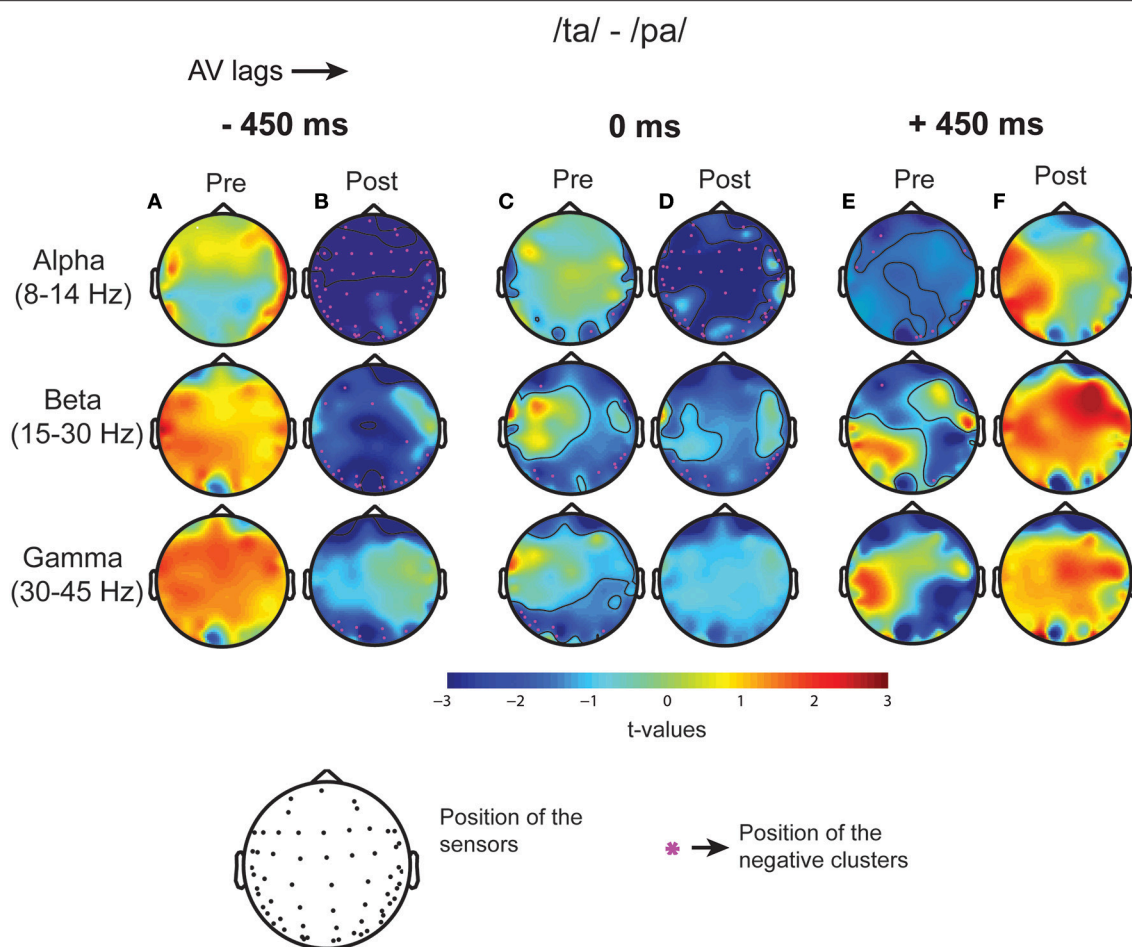


**FIGURE 3 | Power Spectrum.** Spectral-power at each condition and perceptual category during **(A)** Pre-stimulus onset. **(B)** Post-stimulus onset periods. The plots adjacent to the scalp maps show the enlarged plots of the power spectrum at the sensors: Fz, Cz, and Oz.

### Time-Frequency Global Coherogram

Eigenvalue based time-frequency global coherogram (Cimenser et al., 2011) was computed for the epochs of 1.3 s duration (0.4 s pre-stimulus, and 0.9 s post-stimulus segments). The time locking was done to the first sensory component, audio or visual, for  $-450$  and  $+450$  ms AV lag and the onset of AV

stimulus for 0 ms AV lag. The mean coherogram plots for the perceptual categories */ta/* and */pa/* and their difference at AV lags:  $-450$  ms (see **Figures 5A–C**), 0 ms (see **Figures 5D–F**),  $+450$  ms (see **Figures 5G–I**) showed relatively heightened global coherence in the theta band (4–8 Hz) throughout the entire epoch duration. Cluster-based permutation tests employed to



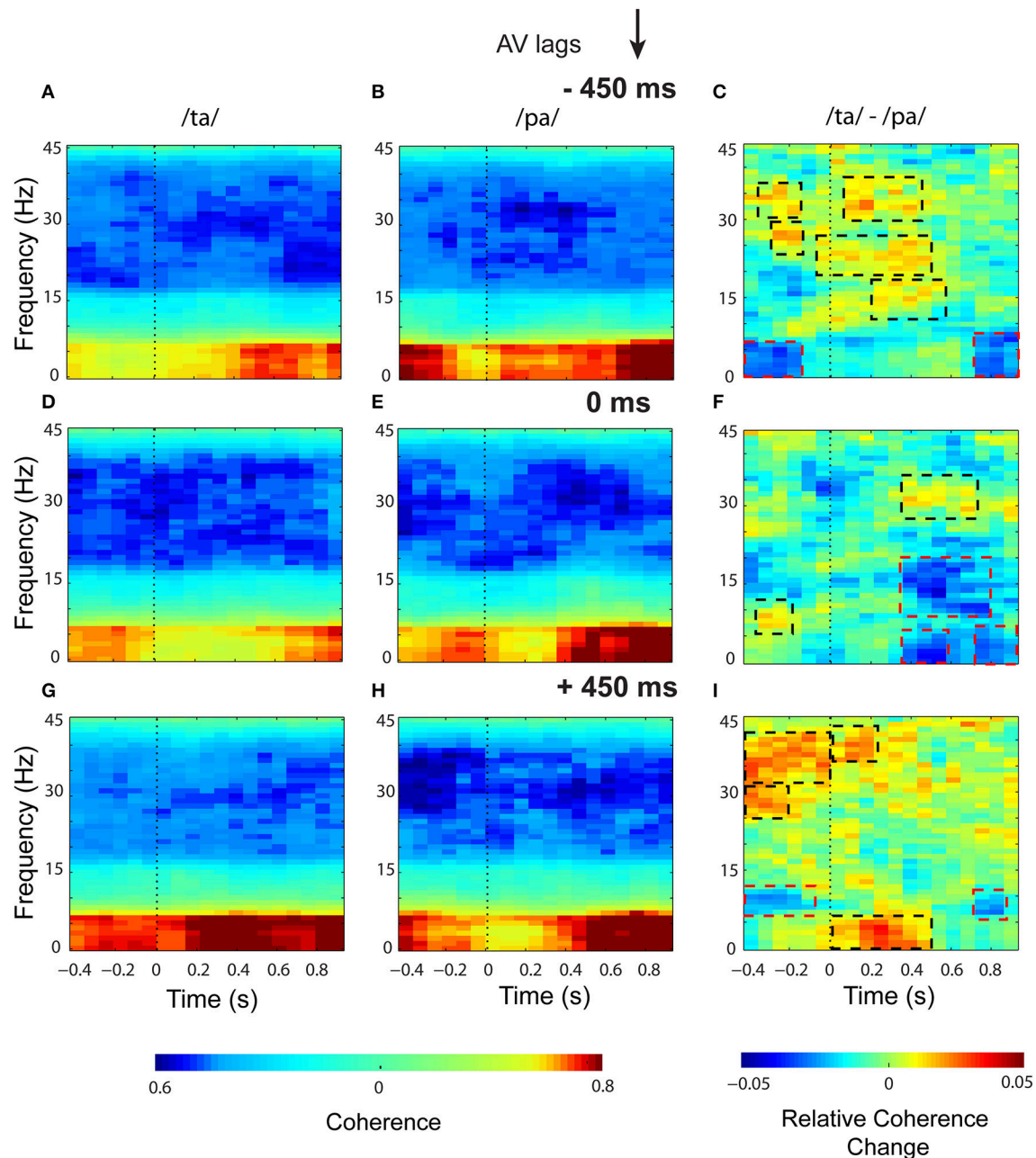
**FIGURE 4 | Spectral Difference.** The topoplots and the magenta “\*” highlight the clusters that show significant difference between the perceptual categories /ta/-/pa/ during the three stimulus conditions: -450 ms AV lag at (A) pre-stimulus onset (B) post-stimulus onset, 0 ms AV lag (C) pre-stimulus onset (D) post-stimulus onset, +450 ms AV lag (E) pre-stimulus onset (F) post-stimulus onset.

compare the mean coherogram for /ta/ and /pa/ at the respective AV lags revealed both positive and negative clusters (see **Figures 5C,E,I**). Positive clusters highlighted in black dashed rectangles signify time-frequency islands of increased synchrony and the negative clusters in red dashed boxes signify islands of decreased synchrony in the global neuronal network.

In the pre-stimulus period, we observed two positive and one negative cluster each during -450 and +450 ms AV lag. The first and second positive clusters during -450 ms AV lag were observed in the frequency bands beta (16–30 Hz) ( $z_{97.5} = 0.29$ ) and gamma (>30 Hz) ( $z_{97.5} = 0.78$ ) respectively and the negative cluster was found in theta band (4–7 Hz) ( $z_{0.025} = -0.29$ ). Here,  $z_{97.5}$  and  $z_{0.025}$  represent the two-tailed thresholds at  $p = 0.05$  set by permutation tests to compute the significantly different cluster (for details, see Methods section and Maris et al., 2007). Similarly during +450 ms AV lag the first and second positive clusters were observed in the frequency bands beta ( $z_{97.5} = 0.26$ ) and gamma ( $z_{97.5} = 0.34$ ) respectively and the negative cluster was found in the alpha band (8–12 Hz) ( $z_{0.025} = -0.78$ ). However, during

0 ms AV lag, only a significant positive cluster was observed in the alpha frequency band ( $z_{97.5} = 0.58$ ).

In the post-stimulus onset period, during -450 ms AV lag (see **Figure 5C**), three positive clusters were observed, (1) in alpha band with temporal range between ~200 and 560 ms ( $z_{97.5} = 0.50$ ), (2) in beta band with temporal range between ~-50 and 500 ms ( $z_{97.5} = 0.29$ ), and (3) in gamma band between ~50 and 400 ms ( $z_{97.5} = 0.78$ ). Also, a negative cluster ( $z_{0.025} = -1.02$ ) was observed in the theta band between ~800 and 900 ms. During +450 ms AV lag (see **Figure 5I**), two positive clusters were observed, one in the theta band ( $z_{97.5} = 0.73$ ) between ~0 and 500 ms and the other one in gamma band ( $z_{97.5} = 0.34$ ) between ~0 and 200 ms. A negative cluster was also observed in the theta band ( $z_{0.025} = -0.68$ ) between ~700 and 850 ms. Interestingly, during 0 ms AV lag (see **Figure 5F**) we observed a positive cluster ( $z_{97.5} = 0.26$ ) precisely in the gamma band (~300 and 700 ms) and three negative clusters ( $p \leq 0.05$ ). Two of the negative clusters ( $z_{0.025} = 0.31$ ) were observed in the theta band around 300 and 600 ms and ~700 and 900 ms



**FIGURE 5 | Time-frequency representations of large-scale functional brain networks.** Mean time frequency coherogram for different perceptual categories time locked to the onset of the first sensory component (A or V) during the three conditions and the mean coherence difference between /ta/ and /pa/ responses at different AV lags: for -450 ms (A) /ta/ (B) /pa/ (C) /ta/-/pa/; for 0 ms (D) /ta/ (E) /pa/ (F) /ta/-/pa/; for 450 ms (G) /ta/ (H) /pa/ (I) /ta/-/pa/.

and the third negative cluster incorporated both alpha and beta bands (9–21 Hz) ( $z_{0.025} = -0.25$ ) and appeared between 300 and 800 ms.

## DISCUSSION

Characterizing the dynamics of the whole brain network is essential for understanding the neurophysiology of multisensory speech perception. We have shown that the spatiotemporal

dynamics of the brain during speech perception can be represented in terms of brain oscillations and large-scale functional brain networks. We explicitly focused on investigating the characteristics of the brain networks that facilitate perception of the McGurk illusion. We exploited the perceptual variability of McGurk stimuli by comparing the oscillatory responses and network characteristics within identical trials. The main findings of the study are: (1) heightened global coherence in the gamma band along with decreased global coherence in the alpha and



theta bands facilitates multisensory perception (2) a broadband enhancement in the global coherence at theta, alpha, beta, and gamma bands aids multisensory perception for asynchronous AV stimuli, as brain engages more energy for multisensory integration. We discuss the behavioral and neural-level findings in following sub-sections.

## Variability of Perceptual Experience

A vast body of literature has reported that under controlled settings one can induce illusory perceptual experience in human participants (McGurk and Macdonald, 1976; MacDonald and McGurk, 1978; van Wassenhove et al., 2007; Nath and Beauchamp, 2011; Keil et al., 2012). Here, we constructed incongruent AV stimuli (auditory /pa/ superimposed onto video of face articulating /ka/) using three different AV lags: -450 ms (audio precede articulatory movements), 0 ms (synchronous onsets of audio and articulatory movements), and +450 ms (articulatory movements precede audio) (see **Figure 1**). We identified that a categorical perceptual difference appeared with variation in AV lags. Synchronous AV stimuli resulted in higher percentage response of crossmodal (/ta/) perception (**Figure 2C**) whereas AV lags of -450 and +450 ms resulted in lowering of the percentage of crossmodal percept and higher occurrence of the unimodal percept /pa/. Furthermore, we observed high hit rate of /ta/ responses both during congruent /ta/ stimuli (>90%) and during our *post-hoc* “auditory alone” behavioral experiment (>95%). Behavioral studies by van Wassenhove et al. (2007) demonstrate 200 ms of asynchrony as the temporal window of bimodal integration. However, electrophysiological studies especially in the domain of preparatory processes demonstrate the elicitation of ERP components up to 600–800 ms in response to a cue followed by a target stimulus (Simson et al., 1977). Extending this line of reasoning to our experimental paradigm, we believe an existence of temporal integration mechanisms beyond 200 ms does not allow the percentage of /pa/ perception to reach the level for congruent multisensory or purely auditory perception. In the current study we focused on the boundaries of stable illusory perception but the temporal boundaries of multisensory integration needs to be tested by future studies.

Interestingly percentage of gaze fixation at the mouth of the speaker for crossmodal response trials did not vary significantly at any AV lags based on *t*-test. Also, the interaction between lags and perceptual categorization was not significant when analyzed with 2-way ANOVA. Even though not statistically significant, the mean gaze fixation percentages at mouth for crossmodal perception were slightly higher than unimodal perception at all AV lags. Therefore, we cannot completely rule out the findings of an earlier study that show that frequent perceivers of McGurk effect fixate more at the mouth of the speaker (Gurler et al., 2015) as well as we were limited by the number of participants to evaluate correlations between the behavioral results and the percentage of gaze fixation at 0 ms AV lag. On the other hand the subjective behavioral response for perceptual categorization clearly showed an interaction effect between AV lags and perceived objects. It is important to note that the identical multisensory stimuli generated varying responses for different trials. All stimuli being multisensory, differential perception

served as an efficient handle to tap into the perceptual processing underlying speech perception. Our behavioral response results are consistent with previous studies on McGurk stimuli (Munhall et al., 1996; van Wassenhove et al., 2007) that demonstrate the influence of AV lags on perceptual experience. Hence, we expected to identify the neurophysiological processes underlying different multisensory perceptual scenarios.

## Spectral Landscape of the Cortical Activity

Non-parametric statistical comparison between the perceptual categories (/ta/-/pa/) showed suppression of the spectral power in alpha, beta, and gamma frequency bands (see **Figure 4**). Suppression of alpha-band power has been associated with attention and language comprehension processes by enabling controlled access to knowledge (Bastiaansen and Hagoort, 2006; Hanslmayr et al., 2011; Klimesch, 2012; Payne et al., 2013). Accordingly, the suppression of alpha-band power observed in our study can be attributed to the attention related network aiding access to stored knowledge and filter redundant information.

Beta-band power was observed to be suppressed at frontoparietal to occipital sensors during -450 ms AV lag and at occipital scalp regions during 0 ms AV lag but no such suppression was observed during +450 ms AV lag. Beta band power has been linked with various cognitive facets including top-down control of attention and cognitive processing (Engel and Fries, 2010). Besides, in the domain of multisensory integration and language processing, suppression of beta-band power has been associated with the occurrence of unexpected stimuli (Bastiaansen and Hagoort, 2006; Weiss and Mueller, 2012). Furthermore, recent studies also show suppression of beta power during the perception of the McGurk illusion (Roa Romero et al., 2015). Extending the line of reasoning from the aforementioned studies, suppression of beta-band power might be associated to the occurrence of an unexpected stimulus and its processing. Visual-lead condition, wherein we observed no significant difference in the beta power, is possibly the most predictable situation and hence significant beta power modulation was not detected. Behaviorally, Munhall et al. (1996), report McGurk illusion is most dominant between an AV lag of 0–200 ms and there is a slight asymmetry toward positive AV lags (visual lead). In fact, our data from a different experiment also replicated this result.

Gamma-band power was observed to be significantly suppressed only during -450 ms AV lag at the occipital scalp regions. Also, in the pre-stimulus period significant reduction in gamma band power was observed at occipital scalp regions during 0 ms AV lag. Existing studies have demonstrated the role of gamma-band oscillations in cognitive functions like visual perception, attention and in the processing of auditory spatial and pattern information (Jochen Kaiser and Lutzenberger, 2005a,b). Also, gamma band activity over sensory areas has been attributed to the detection of changes in AV speech (Kaiser et al., 2006). However, we observed a suppression in gamma band activity which may be linked with preparatory processes over wider network that waits for the expected visual information to arrive. Although, the brain oscillatory responses to multisensory

perception have been extensively studied, a consensus on the mechanisms associated with these oscillations remains elusive. Our study contributes to this vast body of work in conveying that multisensory speech perception requires complex signal processing mechanisms that involves the participation of several brain regions. Therefore, understanding the process requires analyzing the whole brain operating as large scale neurocognitive network. In the subsequent section we discuss the network analysis results.

## Neurocognitive-Network Level Processing Underlying Illusory Perception

Global time-frequency coherogram (see **Figure 5**) computed for the perceptual categories quantifies the extent of coordinated neuronal activity over the whole brain. Global coherence reflects the presence of neuro-cognitive networks in physiological signals (Bressler, 1995). Previous studies posit that neuronal assemblies that represent same perceptual object (von der Malsburg and Schneider, 1986; Engel, 1997; Engel et al., 2001). Besides, going by the communication-through-coherence (CTC) hypothesis, only coherently oscillating neuronal groups communicate effectively as their communication window for spike output and synaptic input are open at the same time (Senkowski et al., 2008; Fries, 2015). Hence, coherent transmission poses a flexible mechanism that facilitates the integration of converging streams in time windows of varying duration. In our analysis we observed a relatively heightened theta-band coherence for both the perceptual categories at all the AV lags (see **Figures 5A,B,D,E,G,H**). Theta band coherence has been associated to cognitive control processes (Cooper et al., 2015). Accordingly, the enhanced theta-band coherence might reflect the control processes preparing for upcoming stimuli.

Non-parametric statistical analysis employed to test the global coherence differences between */ta/* and */pa/* during 0 ms AV lag, revealed a positive cluster, signifying enhanced synchrony specifically at the gamma band (between ~300 ms and 700 ms). Also, we observed negative clusters (between ~300 and 900 ms) in the theta, alpha and beta bands that signify decreased synchrony among the underlying brain regions. Overall temporal congruence of AV stimuli results in a narrow-band coherence whereas lagged AV stimuli seemed to engage a broadband coherence (see **Figure 5C,F,I**). However, we had one limitation because of the nature of our stimuli. A direct statistical comparison across lagged conditions was not meaningful since each lagged condition had a different temporal sequence of audio-visual components.

Inter-areal coherence of oscillatory activity in the beta frequency range (15–30 Hz) has been associated with top-down processing (Wang, 2010). Moreover, top-down processing involves the modulation of the hierarchical sensory and motor systems by pre-frontal and frontal brain areas (Mesulam, 1990). The dense anatomical interconnectivity among these association areas give rise to self-organized large scale neuronal assemblies defined as neuro-cognitive networks (NCNs), with respect to the cognitive demands (Bressler and Richter, 2014). In this context, our finding of increased coherence in the beta band during –450

and +450 ms AV lag is especially relevant as it enables us to hypothesize that synchronization of the beta oscillations provides long range inter-areal linkage of distributed cortical areas in NCNs. Such networks can readily process the retrieval of well learnt audio-visual associations suggested by Albright (2012).

Gamma band coherence are shown to be associated with voluntary eye movements, saccades (Balazs et al., 2015). Besides, stimulus selection by attention also induces local gamma band synchronization (Hipp et al., 2011). Our results show enhanced gamma coherence (positive cluster) at all AV lags. Considering the increased gaze fixation at mouth during */ta/* perception, heightened gamma coherence reflects the recruitment of the visual attention areas. A recent review proposes that gamma band (30–90 Hz) coherence activates postsynaptic neurons effectively by modulating the excitation such that it escapes the following inhibition (Fries, 2015). Besides rendering effective communication, gamma coherence has also been proposed to render communication that are precise and selective (Buzsáki and Schomburg, 2015; Fries, 2015). Importantly, gamma band coherence has also been demonstrated to be implicated in associative learning (Miltner et al., 1999). Thus, our observation of enhanced coherence exclusively at gamma and desynchronization at alpha and beta-bands during 0 ms AV lag portrays an attention network working in harmony with the NCNs most likely linked to associative memory retrieval. This conjecture is also supported by the secondary evidence in case of –450 and 450 ms AV lags, where an additional working memory process is competing for processing and integration of the multisensory stimuli and leading to a broadband enhancement in global coherence. A more detailed delineation of working memory processing and associative memory recall needs to be carried out with other kinds of multisensory stimuli and will be a major focus of our future endeavors.

## AUTHOR CONTRIBUTIONS

GK and AB designed the study; GK, AM, and AB designed the stimulus; GK and AJ recorded the data; GK, TH, DR, and AB analyzed the data; GK, DR, and AB wrote the manuscript; GK, TH, AJ, AM, DR, and AB read and commented on the manuscript.

## ACKNOWLEDGMENTS

The study was supported by NBRC Core funds and by grants Ramalingaswami fellowship, (BT/RLF/Re-entry/31/2011) and Innovative Young Bio-technologist Award (IYBA), (BT/07/IYBA/2013) from the Department of Biotechnology (DBT), Ministry of Science and Technology, Government of India to AB. AB also acknowledges the support of Centre of Excellence in Epilepsy and MEG (BT/01/COE/09/08/2011) from DBT. DR was supported by the Ramalingaswami fellowship (BT/RLF/Re-entry/07/2014) from DBT. We thank Prof. Nandini Chatterjee Singh for helpful comments on an earlier version of this manuscript and Mr. Neeraj Kumar for help in collection of control data.

## REFERENCES

- Alba, G., Pereda, E., Mañas, S., Méndez, L. D., Duque, M. R., González, A., et al. (2016). The variability of EEG functional connectivity of young ADHD subjects in different resting states. *Clin. Neurophysiol.* 127, 1321–1330. doi: 10.1016/j.clinph.2015.09.134
- Albright, T. D. (2012). On the perception of probable things: neural substrates of associative memory, imagery, and perception. *Neuron* 74, 227–245. doi: 10.1016/j.neuron.2012.04.001
- Balazs, S., Kermanshahi, K., Binder, H., Rattay, F., and Bodis-Wollner, I. (2015). Gamma-band modulation and coherence in the EEG by involuntary eye movements in patients in unresponsive wakefulness syndrome. *Clin. EEG Neurosci.* 47, 196–206. doi: 10.1177/1550059415601756
- Bastiaansen, M., and Hagoort, P. (2006). Oscillatory neuronal dynamics during language comprehension. *Prog. Brain Res.* 159, 179–196. doi: 10.1016/S0079-6123(06)59012-0
- Beauchamp, M. S. (2010). fMRI-guided TMS reveals that the STS is a cortical locus of the McGurk effect. *J. Neurosci.* 30, 2414–2417. doi: 10.1523/JNEUROSCI.4865-09.2010.fMRI-guided
- Bressler, S. L. (1995). Large-scale cortical networks and cognition. *Brain Res. Brain Res. Rev.* 20, 288–304.
- Bressler, S. L., Coppola, R., and Nakamura, R. (1993). Episodic multiregional cortical coherence at multiple frequencies during visual task performance. *Nature* 366, 153–156. doi: 10.1038/366153a0
- Bressler, S. L., and Menon, V. (2010). Large-scale brain networks in cognition: emerging methods and principles. *Trends Cogn. Sci.* 14, 277–290. doi: 10.1016/j.tics.2010.04.004
- Bressler, S. L., and Richter, C. G. (2014). Interareal oscillatory synchronization in top-down neocortical processing. *Curr. Opin. Neurobiol.* 31C, 62–66. doi: 10.1016/j.conb.2014.08.010
- Bulkin, D. A., and Groh, J. M. (2006). Seeing sounds: visual and auditory interactions in the brain. *Curr. Opin. Neurobiol.* 16, 415–419. doi: 10.1016/j.conb.2006.06.008
- Buzsáki, G., and Schomburg, E. W. (2015). What does gamma coherence tell us about inter-regional neural communication? *Nat. Neurosci.* 18, 484–489. doi: 10.1038/nn.3952
- Callan, D. E., Jones, J. A., Munhall, K., Callan, A. M., Kroos, C., and Vatikiotis-Bateson, E. (2003). Neural processes underlying perceptual enhancement by visual speech gestures. *Neuroreport* 14, 2213–2218. doi: 10.1097/01.wnr.0000095492.38740.8f
- Cimenser, A., Purdon, P. L., Pierce, E. T., Walsh, J. L., Salazar-Gomez, A. F., Harrell, P. G., et al. (2011). Tracking brain states under general anesthesia by using global coherence analysis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8832–8837. doi: 10.1073/pnas
- Clarke, A. R., Barry, R. J., Indraratna, A., Dupuy, F. E., McCarthy, R., and Selikowitz, M. (2016). EEG activity in children with Asperger's Syndrome. *Clin. Neurophysiol.* 127, 442–451. doi: 10.1016/j.clinph.2015.05.015
- Cooper, P. S., Wong, A. S. W., Fulham, W. R., Thienel, R., Mansfield, E., Michie, P. T., et al. (2015). Theta frontoparietal connectivity associated with proactive and reactive cognitive control processes. *Neuroimage* 108, 354–363. doi: 10.1016/j.neuroimage.2014.12.028
- Deroy, O., Chen, Y.-C., Spence, C., Marcel, A., Angelaki, D., Gu, Y., et al. (2014). Multisensory constraints on awareness. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130207. doi: 10.1098/rstb.2013.0207
- Engel, A. (1997). Role of the temporal domain for response selection and perceptual binding. *Cereb. Cortex* 7, 571–582. doi: 10.1093/cercor/7.6.571
- Engel, A., and Fries, P. (2010). Beta-band oscillations- signalling the status quo? *Curr. Opin. Neurobiol.* 20, 156–165. doi: 10.1016/j.conb.2010.02.015
- Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* 2, 704–716. doi: 10.1038/35094565
- Fonseca, L. C., Tedrus, G. M. A. S., Rezende, A. L. R. A., and Giordano, H. F. (2015). Coherence of brain electrical activity: a quality of life indicator in Alzheimer's disease? *Arq. Neuropsiquiatr.* 73, 396–401. doi: 10.1590/0004-282X20150035
- Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034
- Gurur, D., Doyle, N., Walker, E., Magnotti, J., and Beauchamp, M. (2015). A link between individual differences in multisensory speech perception and eye movements. *Atten. Percept. Psychophys.* 77, 1333–1341. doi: 10.3758/s13414-014-0821-1
- Hanslmayr, S., Gross, J., Klimesch, W., and Shapiro, K. L. (2011). The role of alpha oscillations in temporal attention. *Brain Res. Rev.* 67, 331–343. doi: 10.1016/j.brainresrev.2011.04.002
- Hasson, U., Skipper, J. I., Nusbaum, H. C., and Small, S. L. (2007). Abstract coding of audiovisual speech: beyond sensory representation. *Neuron* 56, 1116–1126. doi: 10.1016/j.neuron.2007.09.037
- Helfer, K. S. (1997). Auditory and auditory-visual perception of clear and conversational speech. *J. Speech Lang. Hear. Res.* 40, 432–443.
- Hipp, J. F., Engel, A. K., and Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron* 69, 387–396. doi: 10.1016/j.neuron.2010.12.027
- Jones, J. A., and Callan, D. E. (2003). Brain activity during audiovisual speech perception: an fMRI study of the McGurk effect. *Neuroreport* 14, 1129–1133. doi: 10.1097/01.wnr.0000074343.81633.2a
- Kaiser, J. (2004). Hearing lips: gamma-band activity during audiovisual speech perception. *Cereb. Cortex* 15, 646–653. doi: 10.1093/cercor/bhh166
- Kaiser, J., Hertrich, I., Ackermann, H., and Lutzenberger, W. (2006). Gamma-band activity over early sensory areas predicts detection of changes in audiovisual speech stimuli. *Neuroimage* 30, 1376–1382. doi: 10.1016/j.neuroimage.2005.10.042
- Kaiser, J., and Lutzenberger, W. (2005a). Cortical oscillatory activity and the dynamics of auditory memory processing. *Rev. Neurosci.* 16, 239–254. doi: 10.1515/REVNEURO.2005.16.3.239
- Kaiser, J., and Lutzenberger, W. (2005b). Human gamma-band activity: a window to cognitive processing. *Neuroreport* 16, 207–211. doi: 10.1097/00001756-200502280-00001
- Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., et al. (2014). Committee report: publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology* 51, 1–21. doi: 10.1111/psyp.12147
- Keil, J., Müller, N., Ihssen, N., and Weisz, N. (2012). On the variability of the McGurk effect: audiovisual integration depends on prestimulus brain states. *Cereb. Cortex* 22, 221–231. doi: 10.1093/cercor/bhr125
- Klimesch, W. (2012). controlled access to stored information. *Trends Cogn. Sci.* 16, 606–617. doi: 10.1016/j.tics.2012.10.007
- Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8, 194–208. doi: 10.1002/(SICI)1097-0193(1999)8:4<194::AID-HBM4>3.0.CO;2-C
- MacDonald, J., and McGurk, H. (1978). Visual influences on speech perception processes. *Percept. Psychophys.* 24, 253–257. doi: 10.3758/BF03206096
- Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024
- Maris, E., Schoffelen, J.-M., and Fries, P. (2007). Nonparametric statistical testing of coherence differences. *J. Neurosci. Methods* 163, 161–175. doi: 10.1016/j.jneumeth.2007.02.011
- McGurk, H., and Macdonald, J. (1976). Hearing lips and seeing voices. *Nature* 264, 691–811. doi: 10.1038/264746a0
- Mesulam, M. M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Ann. Neurol.* 28, 597–613.
- Miltner, W. H. R., Braun, C., Arnold, M., Witte, H., and Taub, E. (1999). Coherence of gamma-band EEG activity as a basis for associative learning. *Nature* 397, 434–436. doi: 10.1038/17126
- Mitra, P., and Bokil, H. (2008). *Observed Brain Dynamics*. New York, NY: Oxford University Press.
- Munhall, K. G., Gribble, P., Sacco, L., and Ward, M. (1996). Temporal constraints on the McGurk effect. *Percept. Psychophys.* 58, 351–362. doi: 10.3758/BF03206811
- Nath, A. R., and Beauchamp, M. S. (2011). Dynamic changes in superior temporal sulcus connectivity during perception of noisy audiovisual speech. *J. Neurosci.* 31, 1704–1714. doi: 10.1523/JNEUROSCI.4853-10.2011
- Nath, A. R., and Beauchamp, M. S. (2013). A neural basis for interindividual differences in the McGurk effect, a multisensory speech illusion. *Neuroimage* 59, 781–787. doi: 10.1016/j.neuroimage.2011.07.024

- Payne, L., Guillory, S., and Sekuler, R. (2013). Attention-modulated alpha-band oscillations protect against intrusion of irrelevant information. *J. Cogn. Neurosci.* 25, 1463–1476. doi: 10.1162/jocn\_a\_00395
- Roa Romero, Y., Senkowski, D., and Keil, J. (2015). Early and late beta-band power reflects audiovisual perception in the McGurk illusion. *J. Neurophysiol.* 113, 2342–2350. doi: 10.1152/jn.00783.2014
- Saint-Amour, D., De Sanctis, P., Molholm, S., Ritter, W., and Foxe, J. J. (2007). Seeing voices: high-density electrical mapping and source-analysis of the multisensory mismatch negativity evoked during the McGurk illusion. *Neuropsychologia* 45, 587–597. doi: 10.1016/j.neuropsychologia.2006.03.036
- Sekiyama, K., Kanno, I., Miura, S., and Sugita, Y. (2003). Auditory-visual speech perception examined by fMRI and PET. *Neurosci. Res.* 47, 277–287. doi: 10.1016/S0168-1012(03)00214-1
- Sekuler, R., Sekuler, A. B., and Lau, R. (1997). Sound alters visual motion perception. *Nature* 385:308. doi: 10.1038/385308a0
- Senkowski, D., Schneider, T. R., Foxe, J. J., and Engel, A. K. (2008). Crossmodal binding through neural coherence: implications for multisensory processing. *Trends Neurosci.* 31, 401–409. doi: 10.1016/j.tins.2008.05.002
- Simson, R., Vaughan, H. G. Jr., and Ritter, W. (1977). The scalp topography of potentials in auditory and visual Go/NoGo tasks. *Electroencephalogr. Clin. Neurophysiol.* 43, 864–875.
- Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., and Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cereb. Cortex* 17, 2387–2399. doi: 10.1093/cercor/bhl147
- Stein, B. E., Meredith, M. A., Huneycutt, W. S., and McDade, L. (1989). Behavioral indices of multisensory integration: orientation to visual cues is affected by auditory stimuli. *J. Cogn. Neurosci.* 1, 12–24. doi: 10.1162/jocn.1989.1.1.12
- Stevenson, R. A., Altieri, N. A., Kim, S., Pisoni, D. B., and James, T. W. (2010). Neural processing of asynchronous audiovisual speech perception. *Neuroimage* 49, 3308–3318. doi: 10.1016/j.neuroimage.2009.12.001
- Sumby, W. H. (1954). Visual contribution to speech intelligibility in noise. *J. Acoust. Soc. Am.* 26, 212. doi: 10.1121/1.1907309
- Thakur, B., Mukherjee, A., Sen, A., and Banerjee, A. (2016). A dynamical framework to relate perceptual variability with multisensory information processing. 6, 1–11. doi: 10.1038/srep31280
- van Atteveldt, N. M., Formisano, E., Blomert, L., and Goebel, R. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. *Cereb. Cortex* 17, 962–974. doi: 10.1093/cercor/bhl007
- van Wassenhove, V., Grant, K. W., and Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1181–1186. doi: 10.1073/pnas.0408949102
- van Wassenhove, V., Grant, K. W., and Poeppel, D. (2007). Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* 45, 598–607. doi: 10.1016/j.neuropsychologia.2006.01.001
- von der Malsburg, C., and Schneider, W. (1986). A neural cocktail-party processor. *Biol. Cybern.* 54, 29–40. doi: 10.1007/BF00337113
- Wallace, M. T., Meredith, M. A., and Stein, B. E. (1993). Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. *J. Neurophysiol.* 69, 1797–809.
- Wang, X. J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiol. Rev.* 1195–1268. doi: 10.1152/physrev.00035.2008
- Weiss, S., and Mueller, H. M. (2012). “Too many betas do not spoil the broth”: the role of beta brain oscillations in language processing. *Front. Psychol.* 3:201. doi: 10.3389/fpsyg.2012.00201

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Kumar, Halder, Jaiswal, Mukherjee, Roy and Banerjee. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Modulation of Alpha and Beta Oscillations during an n-back Task with Varying Temporal Memory Load

Youguo Chen and Xiting Huang\*

Key Laboratory of Cognition and Personality (Ministry of Education), Center of Studies for Psychology and Social Development, Faculty of Psychology, Southwest University, Chongqing, China

## OPEN ACCESS

### Edited by:

Hugo Merchant,  
Universidad Nacional Autónoma  
de México, Mexico

### Reviewed by:

Hedderik Van Rijn,  
University of Groningen, Netherlands  
Ramon Bartolo,  
National Institutes of Health, USA

### \*Correspondence:

Xiting Huang  
ygchen246@gmail.com;  
xthuang@swu.edu.cn

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 19 August 2015

**Accepted:** 21 December 2015

**Published:** 08 January 2016

### Citation:

Chen Y and Huang X (2016)  
Modulation of Alpha and Beta  
Oscillations during an n-back Task  
with Varying Temporal Memory Load.  
Front. Psychol. 6:2031.  
doi: 10.3389/fpsyg.2015.02031

Temporal information can be retained and manipulated in working memory (WM). Neural oscillatory changes in WM were examined by varying temporal WM load. Electroencephalography was obtained from 18 subjects performing a temporal version of the visual n-back WM task ( $n = 1$  or  $2$ ). Electroencephalography revealed that posterior alpha power decreased and temporal region-distributed beta power increased as WM load increased. This result is consistent with previous findings that posterior alpha band reflects inhibition of task-irrelevant information. Furthermore, findings from this study suggest that temporal region-distributed beta band activity is engaged in the active maintenance of temporal duration in WM.

**Keywords:** temporal information, duration, working memory, n-back task, neural oscillation

## INTRODUCTION

Humans can memorize not only attributes of a presented visual stimulus but also its duration of presentation. Working memory (WM) is the system responsible for short-term storage and online manipulation of information, which is necessary for higher-order cognition, such as language, reasoning, and problem-solving (Baddeley, 1992, 2010, 2012). WM constitutes a fundamental aspect of temporal information processing, as encoded stimulus duration is temporally maintained in WM and then transferred into long-term memory. A previously encoded stimulus duration can be retrieved from long-term memory, and held in WM during a task (Gibbon, 1977; Gibbon et al., 1984; Allan, 1998; Coull et al., 2008).

Previous studies have revealed neural substrates that underlie the maintenance of stimulus duration in WM. Acetylcholine in the frontal cortex modulates the speed at which stimulus duration is translated into temporal memory representations (Meck and Church, 1987; Meck, 1996). Both working and reference memory for temporal information are sensitive to choline acetyltransferase inhibition in rats (Meck, 2006). In monkeys, stimulus duration in WM is represented by neuronal activity in prefrontal cortex (Sakurai et al., 2004). Frontal distributed alpha activity is involved in duration maintenance in WM (Chen et al., 2015). A neural network that includes the frontal lobe (left inferior frontal gyrus, right anterior cingulate, pre-supplementary motor area/supplementary motor area, right paracentral lobule, and left precentral gyrus), parietal lobe (left post-central gyrus), temporal lobe (left superior temporal gyrus), limbic system (left insula), and basal ganglia (right and left caudate and putamen) are correlated with maintenance of temporal information (Harrington et al., 2010).

Much research has addressed the maintenance of temporal duration in WM; however, few studies have investigated the manipulation of temporal duration in WM. Only one study has

reported neural substrates underlying the update of temporal information in WM (Gruber et al., 2000). This study used light-emitting diodes (LEDs) that flashed with either a constant inter-stimulus interval (ISI) of 1 s or variable ISIs (0.3–1.7 s, mean = 1 s). In task 1, subjects ignored ISI changes by attempting to detect a hypothetical hidden feature of the LEDs. In task 2, subjects were required to detect ISI changes. Task 1 served as a baseline that controlled for perceptual aspects common to all tasks. During task 2, subjects were required to continually update memorized temporal information. Thus, this task was similar to a one-back WM task including perceptual processing, temporal encoding, memory updating, and comparison. More activation of prefrontal and lateral premotor cortices was observed in task 2 compared with task 1, which may engage in temporal encoding, memory updating, and comparison (Gruber et al., 2000). Typical WM tasks should be adopted to explore the maintenance and manipulation of temporal duration in WM.

The *n*-back task is a representative example of a WM task, because it requires manipulation as well as maintenance of information in WM (Cohen et al., 1997; Meegan et al., 2004; Owen et al., 2005). The *n*-back task requires participants to decide whether a currently presented stimulus matches the stimulus presented *n* trials previously. The load factor *n* can be adjusted to increase or decrease the difficulty level of the task, and to identify the neural substrates underlying WM. Various types of information can be maintained and manipulated in WM, such as letters, words, numbers, shapes, fractals, faces, pictures, locations, and auditory tones (Owen et al., 2005). Neural oscillations during *n*-back tasks have been extensively investigated (Gevins et al., 1997; McEvoy et al., 1998; Pesonen et al., 2007; Krause et al., 2010; Palomaki et al., 2012; Imperatori et al., 2013). Frontal midline theta rhythm (4–7 Hz) has been shown to increase in magnitude as memory load increases (Gevins et al., 1997, 1998; Lei and Roetting, 2011). Studies have shown that theta oscillations play an important role in WM control mechanisms (Schmiedt et al., 2005; Sauseng et al., 2009). In particular, theta oscillations reflect the organization of sequentially ordered items in WM (Hsieh et al., 2011; Roberts et al., 2013; Roux and Uhlhaas, 2014). In contrast, posterior alpha band power (7.5–12 Hz) has been shown to decrease as memory load increases (Gevins et al., 1997, 1998; Lei and Roetting, 2011). Alpha oscillations tend to be attenuated by attention-demanding tasks, reflecting the inhibition of cortical areas that represent task-irrelevant information (Gevins and Smith, 2000; Jokisch and Jensen, 2007; Klimesch et al., 2007; Tuladhar et al., 2007; Manza et al., 2014). The role of the beta band (13–35 Hz) in WM remains under debate. One study found that beta band frequency increases over the parietal region as memory load increases (Deiber et al., 2007). The authors of this study proposed that the beta band is related to item retention and active maintenance for further task requirements. In contrast, other studies have reported that increased WM load is associated with beta desynchronization (i.e., decrease in beta power; Bocková et al., 2007; Pesonen et al., 2007; Krause et al., 2010). It has been proposed that beta oscillations correlate with higher WM performance due to more effective filtering of irrelevant information (Zanto and Gazzaley, 2009).

The present study applied an *n*-back task to investigate neural oscillations that underlie manipulation and maintenance of temporal duration in WM. Neural substrates that underlie an increase in temporal WM load can be identified by parametric changes in *n*. In a temporal version of the *n*-back task, the participant is shown a series of items (e.g., red circles) and asked to decide whether the duration of presentation of the current item matches the duration of the item presented *n* trials back. The task requires manipulation and maintenance of temporal information in WM. As stated previously, theta and alpha bands reflect central executive functions of WM (Sauseng et al., 2005). Specifically, theta band oscillations reflect the organization of sequentially ordered WM items (Schmiedt et al., 2005; Sauseng et al., 2009; Hsieh et al., 2011; Roberts et al., 2013; Roux and Uhlhaas, 2014) and alpha oscillations reflect inhibition of task-irrelevant information (Gevins and Smith, 2000; Jokisch and Jensen, 2007; Klimesch et al., 2007; Tuladhar et al., 2007; Manza et al., 2014). According to the “multiple-component model” by Baddeley and Hitch, unique central executive control mechanisms, such as item organization and inhibition of irrelevant information (Bledowski et al., 2010), are activated for different types of information in WM (Baddeley, 1992, 2010, 2012). We hypothesized that frontal theta would increase and posterior alpha would decrease as temporal WM load increased. As previously stated, the role of the beta band in WM remains under debate. If beta oscillations are related to the maintenance of item information (Deiber et al., 2007), then we would expect beta band power to increase as temporal WM load increases. In contrast, if beta oscillations are like alpha oscillations, which have been associated with inhibition of task-irrelevant information (Zanto and Gazzaley, 2009; Waldhauser et al., 2012), then we would expect to observe a decrease in beta band power (beta desynchronization) as temporal WM load increases.

## MATERIALS AND METHODS

### Participants

Eighteen right-handed undergraduate students (eight male students, 19–24 years of age) were paid for their participation in this experiment. Each participant had normal or corrected-to-normal visual acuity. Participants were not taking any medications and did not suffer from any central nervous system abnormalities or injuries. The study was approved by the local institutional review board. Written informed consent was obtained from each participant. The experimental procedure was conducted in accordance with the Declaration of Helsinki (World Medical Association, 2013).

### Experimental Material and Apparatus

Visual stimuli were displayed on a black background in the center of a computer screen. A 3-cm red circle (2.29°) and a white 2-cm question mark (1.53°) were used as visual stimuli. Four presentation durations were chosen for the red circle. Scalar variability, in which the standard deviation of the estimated intervals increases linearly with their mean, is a verified feature associated with temporal processing (Rakitin

et al., 1998; Brannon et al., 2008). Thus, an exponential function was adopted to select durations to match the difficulty of discrimination between each pair of adjacent durations. The four durations were: 100 ( $100 \times 2^0$ ), 200 ( $100 \times 2^1$ ), 400 ( $100 \times 2^2$ ), and 800 ( $100 \times 2^3$ ) ms. The refresh rate of the computer monitor was 85 Hz, and the computer screen was placed approximately 75 cm from the participant during the task.

## Procedure

The temporal version of the n-back task was used in this study. A 1-back task was defined as low load (LL), and a 2-back task was defined as high load (HL). The order of the two memory load conditions was counterbalanced across subjects. There were four blocks for each memory load condition, and 25 trials for each duration in each block.

The trial sequence was identical for the 1-back and 2-back tasks (**Figure 1**). Temporal jitter between stimuli was used (Luck, 2005) to reduce the distortion that results from overlapping neural activity between previous and subsequent stimuli. Randomized temporal jitter was controlled by E-prime 1.1 (Psychology Software Tools, Inc.). During each trial, a red circle was presented for a randomly selected duration (100, 200, 400, or 800 ms). After a random delay of 400–800 ms, a question mark was presented in the center of the screen until a response was made, or for a maximum of 2000 ms. Participants were informed that they had to respond within 2000 ms. Trials were presented with a random inter-trial interval of 800–1600 ms.

Participants performed a duration comparison task in which they were required to remember the presentation duration of the red circle at two levels of difficulty (LL and HL). In the LL condition, participants indicated whether the duration of the current red circle was the same as that of the previous red circle (**Figure 1A**). In the HL condition, participants indicated whether the duration of the current red circle was the same as that of the red circle presented two presentations previously (**Figure 1B**). The percentages of matched and unmatched trials were both 50% in both the 1-back and 2-back tasks. When the question mark was presented, participants were instructed to press “1” if the memorized durations of the two red circles were the same and “2” if the memorized durations were different. Half of the participants responded with their left hand (pressing “1” with their middle finger and “2” with their index finger), and the other half responded with their right hand (pressing “1” with their index finger and “2” with their middle finger).

## Electrophysiological Recording

Continuous electroencephalography (EEG) was acquired from Ag/AgCl electrodes mounted in an elastic cap (Brain Products GmbH, Gilching, Germany). Sixty-four electrodes were positioned according to the extended 10–20 system. Additional electrodes were placed on the mastoids. Horizontal electrooculograms (EOGs) were acquired using bipolar electrodes positioned at the external ocular canthi, and vertical EOGs were recorded from electrodes placed

above and below the left eye. The EEG and EOG were digitized at 500 Hz with an amplifier bandpass of 0.01–100 Hz, including a 50-Hz notch filter, and stored for offline analysis. All electrode impedances were maintained below 5 k $\Omega$ .

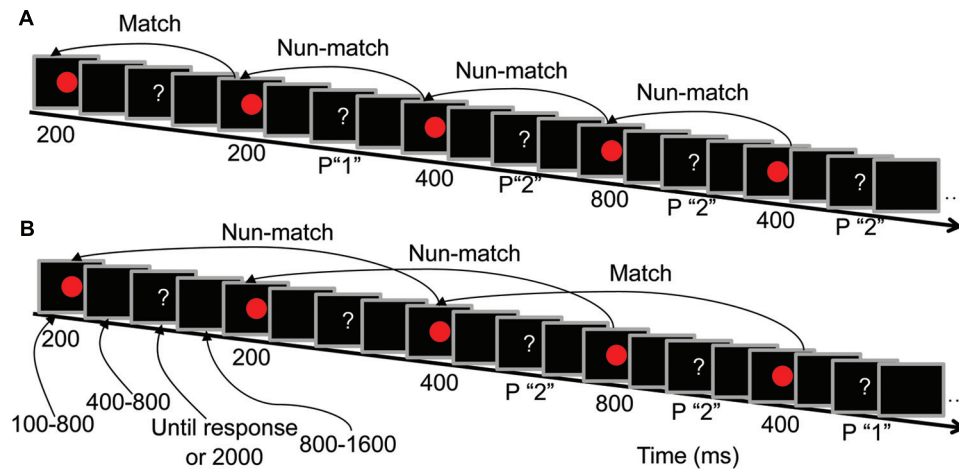
## EEG Analysis

EEGLAB (Delorme and Makeig, 2004) and MATLAB (The MathWorks, Natick, MA, USA) were used for offline EEG data processing. Continuous EEG data were re-referenced to the average of the right and left mastoids, and digitally low-pass filtered at 40 Hz. EEG epochs were segmented in 3-s time windows (pre-stimulus 1 s and post-stimulus 2 s, 0 was onset of stimulus) and baseline-subtracted in the time domain from –1000 to 0 ms. Baseline correction in the time domain effectively subtracts the direct current with no impact on frequency components (Addante et al., 2011). Trials with EOG artifacts (mean EOG voltage exceeding  $\pm 80 \mu\text{V}$ ) and those contaminated with artifacts due to amplifier clipping or peak-to-peak deflection exceeding  $\pm 80 \mu\text{V}$  were excluded. Remaining EOG artifacts were visually identified and removed using independent component analysis according to scalp maps and activity profiles; independent components related to eye movements had a large EOG channel contribution and a frontal scalp distribution (Jung et al., 2000a,b).

Segmented and artifact-free data were used for power spectral analysis. Time-frequency EEG power data were obtained using Hanning-windowed sinusoidal wavelets of three cycles at 3 Hz, rising linearly to approximately 20 cycles at 40 Hz (Gevins et al., 1997; Makeig et al., 2004). The present study focused on ongoing EEG power rather than event-related changes in the power spectrum (Gevins et al., 1997, 1998; Gevins and Smith, 2000; Lei and Roetting, 2011). Thus, the pre-stimulus baseline was not subtracted from ongoing EEG power (Addante et al., 2011; **Figure 4**).

Following previous studies (Hsieh et al., 2011; Chen et al., 2015), electrodes were grouped into nine different clusters: left-frontal (AF7, F7, F5), middle-frontal (F1, Fz, F2), right-frontal (AF8, F8, F6), left-central (C3, C5, T7), middle-central (C1, Cz, C2), right-central (C4, C6, T8), left-posterior (P5, P7, PO7), middle-posterior (O1, O2, Oz), and right-posterior (P6, P8, PO8).

Theta band (4–7 Hz), alpha band (7.5–12 Hz), and beta band (13–34 Hz) powers were analyzed separately. These oscillatory bands were defined by the conventional International Federation of Clinical Neurophysiology (IFCN) guidelines (Nuwer et al., 1999). As shown in **Figure 4**, a posterior alpha decrease and a temporal region-distributed beta increase were observed with increasing WM load from –400 to 1400 ms. Three-way repeated-measures analyses of variance (ANOVAs) were conducted on mean theta, alpha, and beta power in the –400 to 1400 ms time intervals with factors memory load (LL and HL), duration (100, 200, 400, and 800 ms) and region (nine electrode clusters). A Greenhouse–Geisser correction was used to correct for any violations of sphericity (Greenhouse and Geisser, 1959).



**FIGURE 1 | Trial sequences and the duration of each screen presentation. (A)** Low load condition (LL, 1-back task) procedures. **(B)** High load condition (HL, 2-back task) procedures. P "1" indicates that pressing "1" is the correct response, and P "2" indicates that pressing "2" is the correct response.

## RESULTS

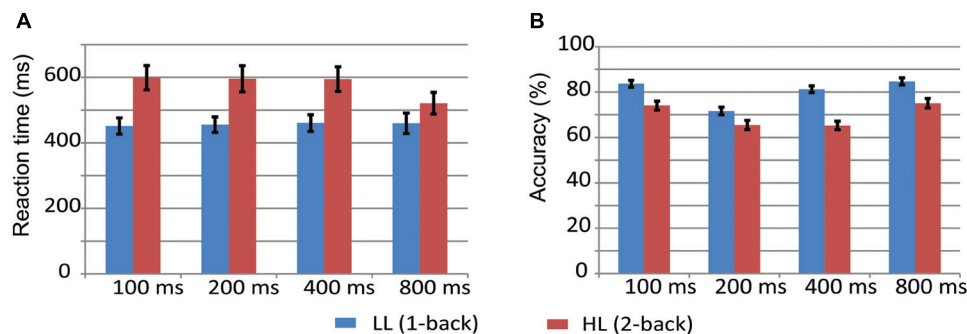
### Behavioral Data

**Figure 2** displays the mean values and standard error of accuracy and reaction time (RT) for 100-, 200-, 400-, and 800-ms durations in the LL and HL conditions. A two-way repeated measures ANOVA on RT with memory load and duration as within-participant factors revealed significant main effects of memory load [ $F(1,17) = 21.966, p < 0.001, \eta_p^2 = 0.564$ ] and duration [ $F(3,51) = 7.086, p < 0.01, \eta_p^2 = 0.294$ ], and a memory load  $\times$  duration interaction [ $F(1.971,33.511) = 9.567, p < 0.01, \eta_p^2 = 0.360$ ]. Simple effects analyses on the memory load  $\times$  duration interaction revealed that RT was longer in the HL condition than the LL condition for all durations [100 ms:  $F(1,17) = 31.570, p < 0.001, \eta_p^2 = 0.650$ ; 200 ms:  $F(1,17) = 22.136, p < 0.001, \eta_p^2 = 0.566$ ; 400 ms:  $F(1,17) = 24.098, p < 0.001, \eta_p^2 = 0.586$ ; 800 ms:  $F(1,17) = 4.466, p = 0.05, \eta_p^2 = 0.208$ ].

A repeated-measures ANOVA on accuracy revealed significant main effects of memory load [ $F(1,17) = 51.169,$

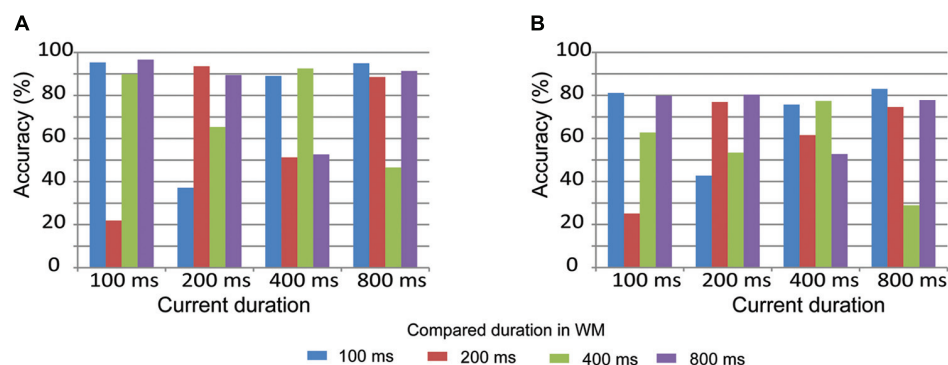
$p < 0.001, \eta_p^2 = 0.751$ ] and duration [ $F(3,51) = 33.503, p < 0.001, \eta_p^2 = 0.663$ ], and a significant memory load  $\times$  duration interaction [ $F(2.239,38.04) = 11.484, p < 0.001, \eta_p^2 = 0.403$ ]. Simple effects analyses on the memory load  $\times$  duration interaction revealed that accuracy was lower in the HL condition than the LL condition for all durations [100 ms:  $F(1,17) = 34.713, p < 0.001, \eta_p^2 = 0.671$ ; 200 ms:  $F(1,17) = 13.523, p < 0.01, \eta_p^2 = 0.443$ ; 400 ms:  $F(1,17) = 114.192, p < 0.001, \eta_p^2 = 0.870$ ; 800 ms:  $F(1,17) = 18.280, p < 0.01, \eta_p^2 = 0.518$ ].

Pairwise comparisons of duration accuracy are displayed in **Figure 3** to further understand how participants compare the current duration with the duration stored in WM. Similar results were observed on both the 1-back and 2-back tasks. Accuracy was low when the current duration was adjacent to the compared duration. For example, accuracy was low when a current duration of 200 ms was compared to adjacent durations of 400 or 600 ms. In contrast, accuracy was high when the current duration was equal to or not adjacent to the compared duration. For example, accuracy was high when a current duration of 200 ms was compared to an equal duration of 200 ms or a non-adjacent



**FIGURE 2 | Reaction time and accuracy in the LL (A, 1-back) and HL (B, 2-back) conditions.**





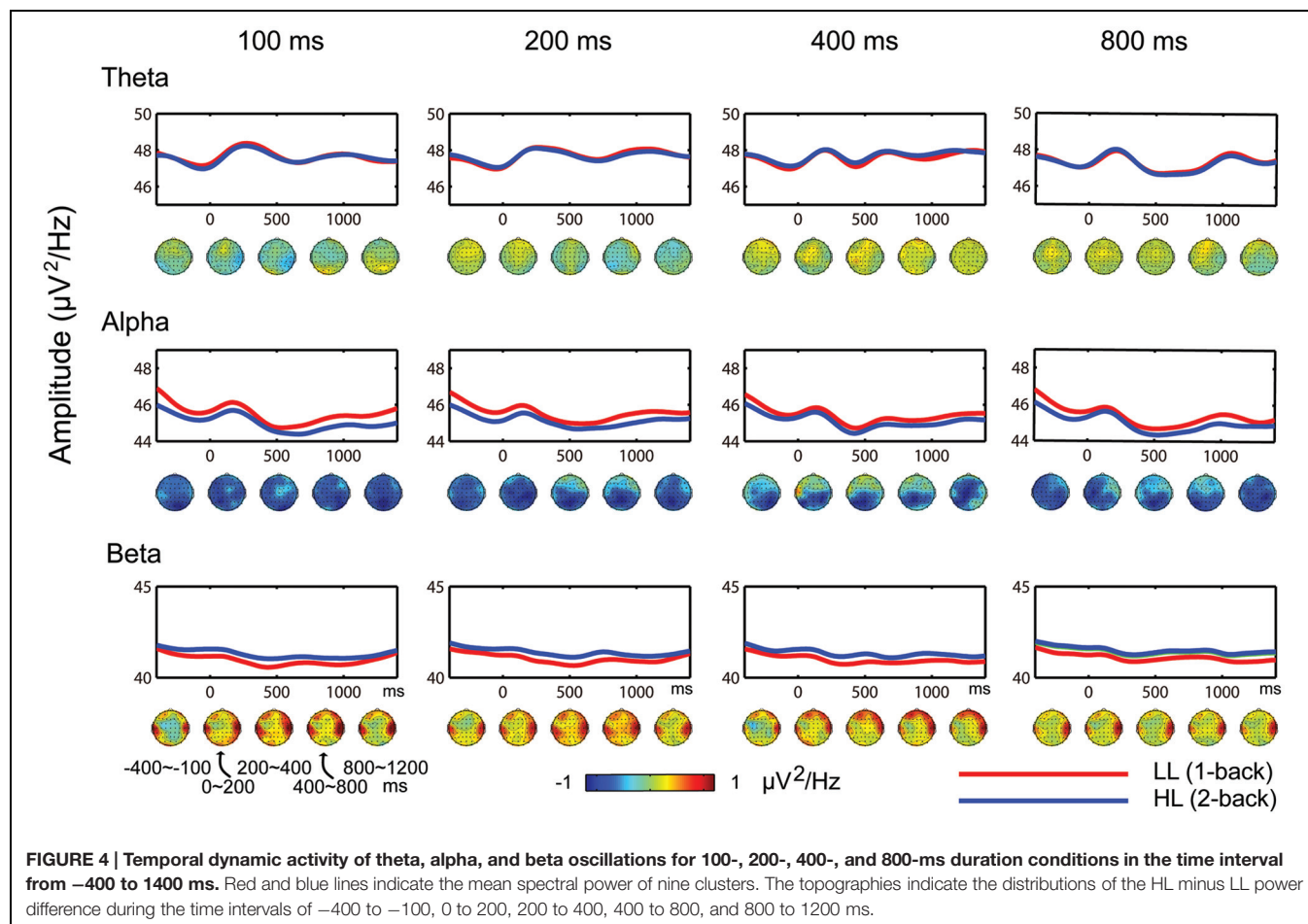
**FIGURE 3 |** Pairwise comparisons of duration accuracy in the LL (A, 1-back task) and HL (B, 2-back task) conditions.

duration of 800 ms. These results suggest that durations were effectively maintained in WM in both the 1-back and 2-back tasks.

## EEG Data

Similar results were obtained from the 100-, 200-, 400-, and 800-ms durations (Figure 4). Theta band oscillations (4–7 Hz) were similar between the HL and LL conditions.

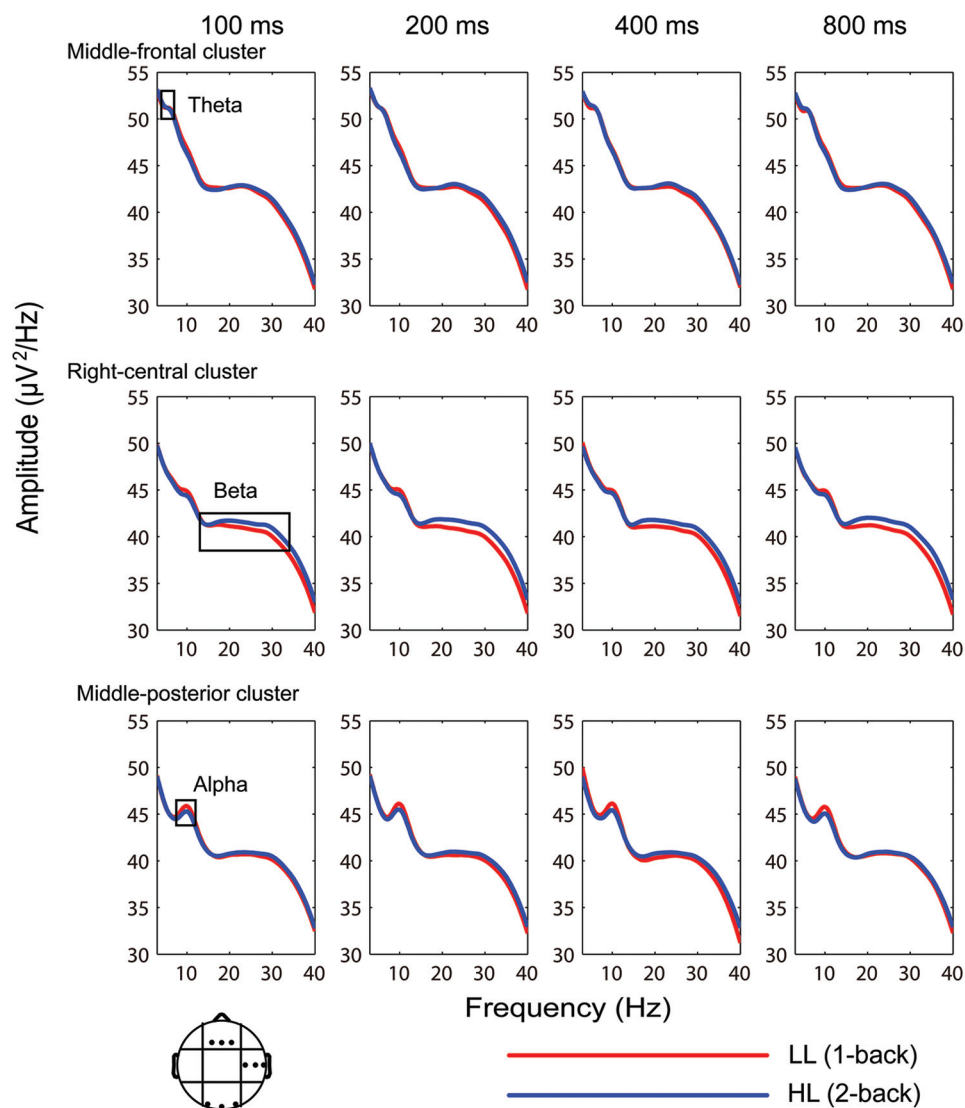
Alpha band power (7.5–12 Hz) over the posterior region from –400 to 1400 ms was lower in the HL condition than the LL condition. Beta band power (13–35 Hz) over the temporal region from –400 to 1400 ms was higher in the HL condition than the LL condition. These results are consistent with the higher WM load that participants are under in the 2-back task, even during the inter-trial interval.



**FIGURE 4 |** Temporal dynamic activity of theta, alpha, and beta oscillations for 100-, 200-, 400-, and 800-ms duration conditions in the time interval from –400 to 1400 ms. Red and blue lines indicate the mean spectral power of nine clusters. The topographies indicate the distributions of the HL minus LL power difference during the time intervals of –400 to –100, 0 to 200, 200 to 400, 400 to 800, and 800 to 1200 ms.

Electroencephalography spectral power was averaged over the time interval from  $-400$  to  $1400$  ms (**Figure 5**). ANOVA conducted on theta band power ( $4\text{--}7$  Hz) revealed significant main effects of duration [ $F(2.506, 42.608) = 20.262$ ,  $p > 0.001$ ,  $\eta_p^2 = 0.544$ ] and region [ $F(3.648, 62.013) = 108.840$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.865$ ]. Theta band amplitude was significantly lower in the 800-ms condition compared with the 100, 200, and 400-ms conditions ( $p$ -values  $< 0.001$ ); the differences between the 100, 200, and 400-ms conditions were not significant ( $p$ -values  $> 0.05$ ). Theta power was highest over the middle-frontal cluster ( $51.056 \pm 0.441 \mu\text{V}^2/\text{Hz}$ ). Main effects of memory load and interactions of memory load  $\times$  duration, memory load  $\times$  region, duration  $\times$  region, and memory load  $\times$  duration  $\times$  region were not significant ( $p$ -values  $> 0.05$ ).

Analysis of variance conducted on alpha band power revealed significant effects of memory load [ $F(1, 17) = 12.945$ ,  $p < 0.01$ ,  $\eta_p^2 = 0.432$ ], duration [ $F(2.054, 34.912) = 4.830$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.221$ ], region [ $F(2.525, 42.918) = 16.270$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.489$ ], and duration  $\times$  region interaction [ $F(6.773, 115.145) = 3.045$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.152$ ]. Alpha band power was higher in the LL condition ( $45.247 \pm 0.759 \mu\text{V}^2/\text{Hz}$ ) than the HL condition ( $44.810 \pm 0.704 \mu\text{V}^2/\text{Hz}$ ). Simple effects analyses on the duration  $\times$  region interaction revealed a significant effect of duration over the left-frontal [ $F(3, 15) = 3.644$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.422$ ], middle-central [ $F(3, 15) = 3.880$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.437$ ], left-posterior [ $F(3, 15) = 3.977$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.443$ ], and middle-posterior [ $F(3, 15) = 10.526$ ,  $p < 0.01$ ,



**FIGURE 5 |** Average EEG spectral power ( $-400$  to  $1400$  ms) for the 100-, 200-, 400-, and 800-ms durations under the low (LL) and high (HL) memory load conditions in the middle-frontal, right-central, and middle-posterior clusters.

$\eta_p^2 = 0.678$ ] clusters such that the alpha power band amplitude was significantly lower in the 800-ms condition than the 100, 200, and 400-ms conditions. Interactions of memory load  $\times$  duration, memory load  $\times$  region, and memory load  $\times$  duration  $\times$  region were not significant ( $p$ -values  $> 0.05$ ).

Analysis of variance conducted on beta band power revealed significant effects of memory load [ $F(1,17) = 6.439$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.275$ ], region [ $F(2.728,46.378) = 15.112$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.471$ ], and memory load  $\times$  region interaction [ $F(3.586,60.969) = 3.064$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.153$ ]. Simple effects analyses on the memory load  $\times$  region interaction revealed that beta band power was significantly lower in the LL condition than the HL condition over the right-frontal [ $F(1,17) = 4.760$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.219$ ], left-central [ $F(1,17) = 7.890$ ,  $p < 0.05$ ,  $\eta_p^2 = 0.317$ ], and right-central [ $F(1,17) = 13.262$ ,  $p < 0.01$ ,  $\eta_p^2 = 0.438$ ] clusters. The main effect of duration and memory load  $\times$  duration, duration  $\times$  region, and memory load  $\times$  duration  $\times$  region interactions were not significant ( $p$ -values  $> 0.05$ ).

Given that similar results were obtained across all four duration conditions (Figures 4 and 5), the oscillation power was averaged across durations to plot the topographies of the oscillations. Theta band was highest over the frontal region, alpha band was highest over the frontal, central, and parietal regions, and beta band was highest over the frontal region in both the LL and HL conditions. LL subtracted from HL revealed an alpha decrease distributed over the posterior region and a beta increase distributed over the temporal region (Figure 6).

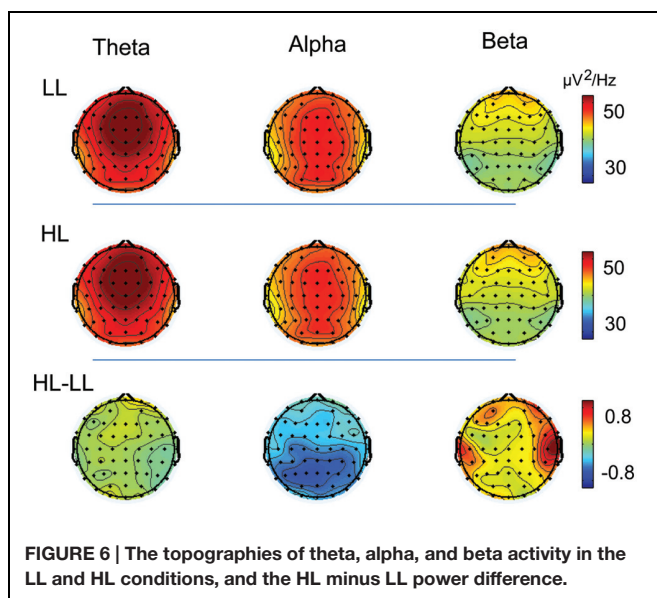
## DISCUSSION

Accuracy decreases and RT increases with increasing WM load on spatial and verbal versions of the  $n$ -back task (Gevins et al., 1997; McEvoy et al., 1998). The present study found that accuracy

was decreased and RT was increased in the HL condition (2-back task) compared with the LL condition (1-back task) for the 100-, 200-, 400-, and 800-ms duration conditions (Figure 2), which suggests that memory load was effectively manipulated. We found a significant memory load  $\times$  duration interaction on RT. This significance was driven by a smaller difference in RT between the 1-back and 2-back task in the 800-ms condition [mean difference (MD): 61.06 ms] compared with the 100-ms (MD: 147.54 ms), 200-ms (MD: 140.07 ms), and 400-ms (MD: 133.85 ms) conditions. Similarly, a significant memory load  $\times$  duration interaction on accuracy is due to a larger difference in accuracy between the 1-back and 2-back tasks in the 400-ms condition (MD: 16.0%) compared with the 100-ms (MD: 9.6%), 200-ms (MD: 6.2%), and 800-ms (MD: 9.6%) conditions. These interactions did not influence the effective manipulation of memory load, and therefore, they will not be further discussed.

Time-frequency analysis was conducted on EEG data to identify the temporal dynamic activity of oscillations (Figure 4). Decreases in alpha band and increases in beta band were observed with increasing temporal WM load from  $-400$  to  $1400$  ms. This result suggests that the WM load is higher in the 2-back task than the 1-back task even during the inter-trial interval. For this reason, the present study focused on ongoing EEG power rather than event-related changes in the power spectrum. If the pre-stimulus baseline is subtracted from ongoing EEG power, then neural activity related to WM load would be removed. Furthermore, alpha band decreases and beta band increases emerged during the time interval from  $-400$  to  $0$  ms (Figure 4), a phase during which temporal encoding (i.e., timing) does not exist. Thus, this result indicates that alpha band decreases and beta band increases are due to increased WM load rather than temporal encoding. In addition, the present study found that theta and alpha band amplitudes were lower in the 800-ms condition than the 100-, 200-, and 400-ms conditions. This result may represent neural oscillatory correlates of temporal encoding, and will not be further discussed.

Consistent with previous studies on WM (Gevins et al., 1997, 1998; Gevins and Smith, 2000; Jensen and Tesche, 2002; Onton et al., 2005; Lei and Roetting, 2011), a pronounced theta power was distributed over the frontal midline in both the LL and HL conditions (Figure 6). This theta activity emanates from the anterior cingulate cortex (Onton et al., 2005; Womelsdorf et al., 2010). The present study found that theta power was not modulated by increasing temporal memory load. To determine whether this result was due to the time-frequency analysis method, time-frequency EEG power data were obtained using Hanning-windowed sinusoidal wavelets of three cycles at 3 Hz (Makeig et al., 2004). This analysis method was previously adopted to extract frontal midline theta during a Sternberg WM task (Onton et al., 2005). We performed a supplementary analysis in which each set of EEG data (5-s epoch) was subjected to Fast-Fourier Transform (FFT) analysis (Chen et al., 2008). No distinct difference in theta band power between the LL and HL conditions was observed. This result suggests that the lack of an effect of temporal WM load on theta band power is not due to the time-frequency analysis method.



Previous studies have shown that theta band reflects the organization of sequentially ordered items in WM (Hsieh et al., 2011; Roberts et al., 2013; Roux and Uhlhaas, 2014). The number of temporal order relationships among items in WM increases as WM load increases, which in turn increases the amplitude of theta band power (Hsieh et al., 2011). This finding was not confirmed in the present study. In previous studies, letters, digits, locations, or visual objects were held in WM, and the visual representation of each item was different (Roland and Gulyas, 1994). In the present study, one duration was stored in WM for the 1-back task, and two durations were stored in WM for the 2-back task. However, the same red circle was presented in each trial, and thus the visual representation of each item was identical in the LL and HL conditions. Our results suggest that the amplitude of the theta band increases as a function of the number of temporal order relationships only when different visual representations are stored in WM. This hypothesis should be tested further in future studies.

Consistent with previous n-back studies (Gevins et al., 1997, 1998; Gevins and Smith, 2000; Lei and Roetting, 2011), the present study found that alpha power decreased with increased memory load. This result is consistent with the finding that increases in alpha oscillation amplitudes reflect increases in cortical inhibition, and decreases in alpha band reflect task-relevant cortical activity (Pfurtscheller, 2001; Klimesch et al., 2007). Functional neuroimaging studies revealed that areas involved in WM (prefrontal and parietal cortex) vary as a function of memory load, with greater activation for higher load levels (Cohen et al., 1997; Owen et al., 2005). Thus, decreases in alpha band in posterior sites reflect increases in cortical activity with increased memory load.

The present study supports the role of beta band oscillations in maintenance rather than inhibition. Given that alpha and beta oscillations are proposed to reflect inhibition of interfering visual memories (Waldhauser et al., 2012), decreased beta band would be expected to be observed with an increased WM load (Zanto and Gazzaley, 2009). However, others have proposed that beta oscillations are related to the maintenance of item information, such that beta band power would increase with increased temporal WM load (Deiber et al., 2007). The beta band increase in the present study supports the maintenance hypothesis. This result is consistent with previous studies that beta increase is associated with maintaining an existing steady state in motor control (Gilbertson et al., 2005; Pogosyan et al., 2009). The present study found that the increased beta was largest over temporal region, which is in agreement with a neuroimaging study that cortico-striatal circuits and superior temporal lobe engage in maintenance of duration in WM (Harrington et al., 2010). Previous studies revealed that phase synchrony in beta oscillations plays an important role for connectivity and communication between/within cortico-striatal circuits and auditory cortex (Fujioka et al., 2012), which may explain how beta oscillations maintain information in neural networks.

Our research will inspire future studies on temporal information processing. First, as the first step, we showed that n-back task is suitable for studying maintenance and

manipulation of duration in WM, and revealed functions of alpha and beta bands in maintenance and manipulation of duration in WM. This experimental paradigm can be used to identify several unsolved scientific problems about representation of duration in WM. E.g., Whether auditory and visual duration is represented differently in WM; whether there is any difference in representations between short and long durations. Second, comparing with previous studies, our study revealed a specific neural activity pattern for duration maintenance in WM. We found that temporal region-distributed beta bands reflect maintenance of duration in WM. Deiber et al. (2007) demonstrated the reactivity of the beta oscillations to the verbal WM load, more pronounced in the right parietal region. Differences in topographies of beta bands are consistent with a previous meta-analysis study which revealed subregional and lateralized differences in activation of a frontoparietal network in response to contents of WM (such as locations, letters, sharps; Owen et al., 2005). Our study indicates the specific neural activity pattern for temporal WM load which can be further identified using n-back task combining with magnetoencephalogram (MEG) or functional magnetic resonance imaging (fMRI). Third, our study proposed an open question. It is not clear why isn't theta power modulated by increasing temporal memory load. It indicates that there are certain differences in organization of durations and other types of information (such as locations, letters). By solving this open question, it is helpful to understand how temporal durations are organized in WM.

To summarize, the present study applied an n-back paradigm to explore neural oscillatory correlates of maintenance and manipulation of duration in WM. We found that frontal midline theta activity was not modulated by increased duration memory load, whereas alpha power was decreased over the posterior region and beta power was increased over the temporal region in the HL compared with the LL condition. The relationship between theta band and the organization of duration in WM needs to be further investigated. Our results are consistent with previous studies in which posterior alpha band was shown to reflect the inhibition of task-irrelevant information. This study also revealed an important role of temporal region-distributed beta in the active maintenance of duration in WM.

## AUTHOR CONTRIBUTIONS

YC and XH developed the experimental concept and the design. YC collected data. YC performed the data analysis and interpretation under the supervision of XH. YC and XH wrote the manuscript. All authors approved the final version of the manuscript for submission.

## FUNDING

This study was supported by a grant from the National Natural Science Foundation of China (31200855, 31300845), the Key Research Institute of Humanities and Social Science in Chongqing (10SKB23), the Doctoral Foundation of Southwest University (SWU110037).



## REFERENCES

- Addante, R. J., Watrous, A. J., Yonelinas, A. P., Ekstrom, A. D., and Ranganath, C. (2011). Prestimulus theta activity predicts correct source memory retrieval. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10702–10707. doi: 10.1073/pnas.1014528108
- Allan, L. G. (1998). The influence of the scalar timing model on human timing research. *Behav. Process.* 44, 101–117. doi: 10.1016/S0376-6357(98)00043-6
- Baddeley, A. (1992). Working memory. *Science* 255, 556–559. doi: 10.1126/science.1736359
- Baddeley, A. (2010). Working memory. *Curr. Biol.* 20, R136–R140. doi: 10.1016/j.cub.2009.12.014
- Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422
- Bledowski, C., Kaiser, J., and Rahm, B. (2010). Basic operations in working memory: contributions from functional imaging studies. *Behav. Brain Res.* 214, 172–179. doi: 10.1016/j.bbr.2010.05.041
- Bocková, M., Chládek, J., Jurák, P., Haláček, J., and Rektor, I. (2007). Executive functions processed in the frontal and lateral temporal cortices: intracerebral study. *Clin. Neurophysiol.* 118, 2625–2636. doi: 10.1016/j.clinph.2007.07.025
- Brannon, E. M., Libertus, M. E., Meck, W. H., and Woldorff, M. G. (2008). Electrophysiological measures of time processing in infant and adult brains: weber's law holds. *J. Cogn. Neurosci.* 20, 193–203. doi: 10.1162/jocn.2008.20016
- Chen, A. C., Feng, W., Zhao, H., Yin, Y., and Wang, P. (2008). EEG default mode network in the human brain: spectral regional field powers. *Neuroimage* 41, 561–574. doi: 10.1016/j.neuroimage.2007.12.064
- Chen, Y. G., Chen, X., Kuang, C. W., and Huang, X. T. (2015). Neural oscillatory correlates of duration maintenance in working memory. *Neuroscience* 290, 389–397. doi: 10.1016/j.neuroscience.2015.01.036
- Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., et al. (1997). Temporal dynamics of brain activation during a working memory task. *Nature* 386, 604–608. doi: 10.1038/386604a0
- Coull, J. T., Nazarian, B., and Vidal, F. (2008). Timing, storage, and comparison of stimulus duration engage discrete anatomical components of a perceptual timing network. *J. Cogn. Neurosci.* 20, 2185–2197. doi: 10.1162/jocn.2008.20153
- Deiber, M. P., Missonnier, P., Bertrand, O., Gold, G., Fazio-Costa, L., Ibanez, V., et al. (2007). Distinction between perceptual and attentional processing in working memory tasks: a study of phase-locked and induced oscillatory brain dynamics. *J. Cogn. Neurosci.* 19, 158–172. doi: 10.1162/jocn.2007.19.1.158
- 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. doi: 10.1016/j.jneumeth.2003.10.009
- Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic  $\beta$  oscillations. *J. Neurosci.* 32, 1791–1802. doi: 10.1523/JNEUROSCI.4107-11.2012
- Gevens, A., and Smith, M. E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cereb. Cortex* 10, 829–839. doi: 10.1093/cercor/10.9.829
- Gevens, A., Smith, M. E., Leong, H., McEvoy, L., Whitfield, S., Du, R., et al. (1998). Monitoring working memory load during computer-based tasks with EEG pattern recognition methods. *Hum. Factors* 40, 79–91. doi: 10.1518/001872098779480578
- Gevens, A., Smith, M. E., McEvoy, L., and Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. *Cereb. Cortex* 7, 374–385. doi: 10.1093/cercor/7.4.374
- Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychol. Rev.* 84, 279–325. doi: 10.1037/0033-295X.84.3.279
- Gibbon, J., Church, R. M., and Meck, W. (1984). "Scalar timing in memory," in *Annals of the New York Academy of Sciences: Timing and Time Perception*, eds J. Gibbon and L. Allan (New York, NY: New York Academy of Sciences), 52–77.
- Gilbertson, T., Lalo, E., Doyle, L., Di Lazzaro, V., Cioni, B., and Brown, P. (2005). Existing motor state is favored at the expense of new movement during 13-35 Hz oscillatory synchrony in the human corticospinal system. *J. Neurosci.* 25, 7771–7779. doi: 10.1523/JNEUROSCI.1762-05.2005
- Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112. doi: 10.1007/BF02289823
- Gruber, O., Kleinschmidt, A., Binkowski, F., Steinmetz, H., and Von Cramon, D. Y. (2000). Cerebral correlates of working memory for temporal information. *Neuroreport* 11, 1689–1693. doi: 10.1097/00001756-200006050-00019
- Harrington, D. L., Zimelman, J. L., Hinton, S. C., and Rao, S. M. (2010). Neural modulation of temporal encoding, maintenance, and decision processes. *Cereb. Cortex* 20, 1274–1285. doi: 10.1093/cercor/bhp194
- Hsieh, L. T., Ekstrom, A. D., and Ranganath, C. (2011). Neural oscillations associated with item and temporal order maintenance in working memory. *J. Neurosci.* 31, 10803–10810. doi: 10.1523/JNEUROSCI.0828-11.2011
- Imperioli, C., Farina, B., Brunetti, R., Gnoni, V., Testani, E., Quintiliani, M. I., et al. (2013). Modifications of EEG power spectra in mesial temporal lobe during n-back tasks of increasing difficulty. A sLORETA study. *Front. Hum. Neurosci.* 7:109. doi: 10.3389/Fnhum.2013.00109
- Jensen, O., and Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *Eur. J. Neurosci.* 15, 1395–1399. doi: 10.1046/j.1460-9568.2002.01975.x
- Jokisch, D., and Jensen, O. (2007). Modulation of gamma and alpha activity during a working memory task engaging the dorsal or ventral stream. *J. Neurosci.* 27, 3244–3251. doi: 10.1523/JNEUROSCI.5399-06.2007
- Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., McKeown, M. J., Iragui, V., et al. (2000a). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37, 163–178. doi: 10.1111/1469-8986.3720163
- Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., and Sejnowski, T. J. (2000b). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin. Neurophysiol.* 111, 1745–1758. doi: 10.1016/S1388-2457(00)00386-2
- Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* 53, 63–88. doi: 10.1016/j.brainresrev.2006.06.003
- Krause, C. M., Pesonen, M., and Hamalainen, H. (2010). Brain oscillatory 4-30 Hz electroencephalogram responses in adolescents during a visual memory task. *Neuroreport* 21, 767–771. doi: 10.1097/WNR.0b013e32833bfcbb
- Lei, S. G., and Roetting, M. (2011). Influence of task combination on EEG spectrum modulation for driver workload estimation. *Hum. Factors* 53, 168–179. doi: 10.1177/0018720811400601
- Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. London: The MIT Press.
- Makeig, S., Delorme, A., Westerfield, M., Jung, T. P., Townsend, J., Courchesne, E., et al. (2004). Electroencephalographic brain dynamics following manually responded visual targets. *PLoS Biol.* 2:e176. doi: 10.1371/journal.pbio.0020176
- Manza, P., Hau, C. L., and Leung, H. C. (2014). Alpha power gates relevant information during working memory updating. *J. Neurosci.* 34, 5998–6002. doi: 10.1523/JNEUROSCI.4641-13.2014
- McEvoy, L. K., Smith, M. E., and Gevins, A. (1998). Dynamic cortical networks of verbal and spatial working memory: effects of memory load and task practice. *Cereb. Cortex* 8, 563–574. doi: 10.1093/cercor/8.7.563
- Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cogn. Brain Res.* 3, 227–242. doi: 10.1016/0926-6410(96)00009-2
- Meck, W. H. (2006). Temporal memory in mature and aged rats is sensitive to choline acetyltransferase inhibition. *Brain Res.* 1108, 168–175. doi: 10.1016/j.brainres.2006.06.047
- Meck, W. H., and Church, R. M. (1987). Cholinergic modulation of the content of temporal memory. *Behav. Neurosci.* 101, 457–464. doi: 10.1037//0735-7044.101.4.457
- Meegan, D. V., Purc-Stephenson, R., Honsberger, M. J., and Topan, M. (2004). Task analysis complements neuroimaging: an example from working memory research. *Neuroimage* 21, 1026–1036. doi: 10.1016/j.neuroimage.2003.10.011
- Nuwer, M. R., Comi, G., Emerson, R., Fuglsang-Frederiksen, A., Guerit, J. M., Hinrichs, H., et al. (1999). IFCN standards for digital recording of clinical EEG. The international federation of clinical neurophysiology. *Electroencephalogr. Clin. Neurophysiol. Suppl.* 52, 11–14.
- Onton, J., Delorme, A., and Makeig, S. (2005). Frontal midline EEG dynamics during working memory. *Neuroimage* 27, 341–356. doi: 10.1016/j.neuroimage.2005.04.014
- Owen, A. M., Mcmillan, K. M., Laird, A. R., and Bullmore, E. (2005). N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Hum. Brain Mapp.* 25, 46–59. doi: 10.1002/hbm.20131

- Palomaki, J., Kivikangas, M., Alafuzoff, A., Hakala, T., and Krause, C. M. (2012). Brain oscillatory 4-35 Hz EEG responses during an n-back task with complex visual stimuli. *Neurosci. Lett.* 516, 141–145. doi: 10.1016/j.neulet.2012.03.076
- Pesonen, M., Hamalainen, H., and Krause, C. M. (2007). Brain oscillatory 4-30 Hz responses during a visual n-back memory task with varying memory load. *Brain Res.* 1138, 171–177. doi: 10.1016/j.brainres.2006.12.076
- Pfurtscheller, G. (2001). Functional brain imaging based on ERD/ERS. *Vision Res.* 41, 1257–1260. doi: 10.1016/S0042-6989(00)00235-2
- Pogosyan, A., Gaynor, L. D., Eusebio, A., and Brown, P. (2009). Boosting cortical activity at Beta-band frequencies slows movement in humans. *Curr. Biol.* 19, 1637–1641. doi: 10.1016/j.cub.2009.07.074
- Rakitin, B. C., Gibbon, J., Penney, T. B., Malapani, C., Hinton, S. C., and Meck, W. H. (1998). Scalar expectancy theory and peak-interval timing in humans. *J. Exp. Psychol. Anim. Behav. Process.* 24, 15–33. doi: 10.1037/0097-7403.24.1.15
- Roberts, B. M., Hsieh, L. T., and Ranganath, C. (2013). Oscillatory activity during maintenance of spatial and temporal information in working memory. *Neuropsychologia* 51, 349–357. doi: 10.1016/j.neuropsychologia.2012.10.009
- Roland, P. E., and Gulyas, B. (1994). Visual imagery and visual representation. *Trends Neurosci.* 17, 281–287. doi: 10.1016/0166-2236(94)90057-4
- Roux, F., and Uhlhaas, P. J. (2014). Working memory and neural oscillations: alpha-gamma versus theta-gamma codes for distinct WM information? *Trends Cogn. Sci.* 18, 16–25. doi: 10.1016/j.tics.2013.10.010
- Sakurai, Y., Takahashi, S., and Inoue, M. (2004). Stimulus duration in working memory is represented by neuronal activity in the monkey prefrontal cortex. *Eur. J. Neurosci.* 20, 1069–1080. doi: 10.1111/j.1460-9568.2004.03525.x
- Sauseng, P., Griesmayr, B., Freunberger, R., and Klimesch, W. (2009). Control mechanisms in working memory: a possible function of EEG theta oscillations. *Neurosci. Biobehav. Rev.* 34, 1015–1022. doi: 10.1016/j.neubiorev.2009.12.006
- Sauseng, P., Klimesch, W., Schabus, M., and Doppelmayr, M. (2005). Frontoparietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *Int. J. Psychophysiol.* 57, 97–103. doi: 10.1016/j.ijpsycho.2005.03.018
- Schmiedt, C., Brand, A., Hildebrandt, H., and Basar-Eroglu, C. (2005). Event-related theta oscillations during working memory tasks in patients with schizophrenia and healthy controls. *Brain Res. Cogn. Brain Res.* 25, 936–947. doi: 10.1016/j.cogbrainres.2005.09.015
- Tuladhar, A. M., Ter Huurne, N., Schoffelen, J. M., Maris, E., Oostenveld, R., and Jensen, O. (2007). Parieto-occipital sources account for the increase in alpha activity with working memory load. *Hum. Brain Mapp.* 28, 785–792. doi: 10.1002/hbm.20306
- Waldhauser, G. T., Johansson, M., and Hanslmayr, S. (2012). Alpha/beta oscillations indicate inhibition of interfering visual memories. *J. Neurosci.* 32, 1953–1961. doi: 10.1523/JNEUROSCI.4201-11.2012
- Womelsdorf, T., Johnston, K., Vinck, M., and Everling, S. (2010). Theta-Activity in anterior cingulate cortex predicts task rules and their adjustments following errors. *Proc. Natl. Acad. Sci. U.S.A.* 107, 5248–5253. doi: 10.1073/pnas.0906194107
- World Medical Association (2013). “WMA declaration of Helsinki: ethical principles for medical research involving human subjects,” in *Proceedings of the 64th WMA General Assembly*, Fortaleza. Available at: <http://www.wma.net/en/30publications/10policies/b3/index.html>
- Zanto, T. P., and Gazzaley, A. (2009). Neural suppression of irrelevant information underlies optimal working memory performance. *J. Neurosci.* 29, 3059–3066. doi: 10.1523/JNEUROSCI.4621-08.2009

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Chen and Huang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Corticostratial Field Potentials Are Modulated at Delta and Theta Frequencies during Interval-Timing Task in Rodents

Eric B. Emmons<sup>1</sup>, Rafael N. Ruggiero<sup>1,2</sup>, Ryan M. Kelley<sup>1</sup>, Krystal L. Parker<sup>1</sup> and Nandakumar S. Narayanan<sup>1,3\*</sup>

<sup>1</sup> Department of Neurology, Carver College of Medicine, The University of Iowa, Iowa City, IA, USA, <sup>2</sup> Department of Neuroscience and Behavioral Sciences, University of São Paulo, São Paulo, Brazil, <sup>3</sup> Aging Mind and Brain Initiative, Carver College of Medicine, The University of Iowa, Iowa City, IA, USA

## OPEN ACCESS

### Edited by:

Hugo Merchant,  
Universidad Nacional Autónoma  
de México, Mexico

### Reviewed by:

Bon-Mi Gu,  
Duke University, USA  
Wilbert Zarco,  
The Rockefeller University, USA

### \*Correspondence:

Nandakumar S. Narayanan  
nandakumar-narayanan@uiowa.edu

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 10 November 2015

**Accepted:** 15 March 2016

**Published:** 05 April 2016

### Citation:

Emmons EB, Ruggiero RN,  
Kelley RM, Parker KL  
and Narayanan NS (2016)  
Corticostratial Field Potentials  
Are Modulated at Delta and Theta  
Frequencies during Interval-Timing  
Task in Rodents.  
Front. Psychol. 7:459.  
doi: 10.3389/fpsyg.2016.00459

Organizing movements in time is a critical and highly conserved feature of mammalian behavior. Temporal control of action requires corticostratial networks. We investigate these networks in rodents using a two-interval timing task while recording LFPs in medial frontal cortex (MFC) or dorsomedial striatum. Consistent with prior work, we found cue-triggered delta (1–4 Hz) and theta activity (4–8 Hz) primarily in rodent MFC. We observed delta activity across temporal intervals in MFC and dorsomedial striatum. Rewarded responses were associated with increased delta activity in MFC. Activity in theta bands in MFC and delta bands in the striatum was linked with the timing of responses. These data suggest both delta and theta activity in frontostriatal networks are modulated during interval timing and that activity in these bands may be involved in the temporal control of action.

**Keywords:** prefrontal cortex, striatum, dorsomedial striatum, Parkinson's disease, medial frontal cortex, local field potential, temporal control, interval timing

## INTRODUCTION

The cortex and striatum are critical for the temporal control of action in mammals (Buhusi and Meck, 2005). These regions are dysfunctional in neuropsychiatric disorders such as schizophrenia and PD, resulting in impaired temporal processing and other cognitive deficits (Malapani et al., 1998; Matell et al., 2003; Ward et al., 2012; Parker et al., 2015). The underlying mechanisms of temporal control by corticostratial systems remain unclear. A better understanding of these circuits could provide insight into both mammalian behavior and human disease.

Temporal control of action can be studied using an interval-timing task. This task requires subjects to estimate an interval of several seconds by making a motor response. Interval timing requires both working memory for temporal rules and attention to the passage of time. Goal-directed timing behavior also shares resources with other executive processes (Brown et al., 2013; Parker et al., 2013). In both humans and rodents, prefrontal areas and dorsal striatum are required for temporal processing (Meck and Benson, 2002; Meck, 2006; Coull et al., 2011). In rodents, inactivation of medial frontal cortex (MFC) impairs interval timing (Uylings et al., 2003; Narayanan et al., 2012; Kim et al., 2013). MFC projects to dorsal and medial regions of the rodent

**Abbreviations:** ERP, event-related potential; LFP, local field potential; PD, Parkinson's disease.

striatum which are also required for temporal control of action, unlike the ventral striatum (Matell and Meck, 2004; Meck, 2006; Kurti and Matell, 2011).

In medial frontal regions of humans and rodents, low-frequency activity is associated with cognitive control and organizing goal-directed activity in time (Cavanagh et al., 2012; Narayanan et al., 2013; Cavanagh and Frank, 2014). Activity around 4 Hz coordinates task information in prefrontal, midbrain, and hippocampal areas (Fujisawa and Buzsáki, 2011). However, it is unclear if activity in this band extends to the basal ganglia. Strikingly, cue-triggered activity in delta and theta bands in MFC is highly conserved in humans and rodents during timing tasks (Narayanan et al., 2013; Parker et al., 2015). In frontal cortex, activity in these bands is coherent with neurons that may encode the accumulation of temporal information (Parker et al., 2014). In both humans and rodents, this cue-triggered delta and theta activity depends on dopamine via medial frontal D1 dopamine receptors (Parker et al., 2015). In the striatum, the striatal beat frequency model proposes that oscillations in the activity of individual neurons may act as a mechanism for the representation of time (Matell and Meck, 2004; Oprisan and Buhusi, 2011). These and many other findings suggest that low-frequency activity may be an important component of temporal processing (Gu et al., 2015).

To further explore the role of low-frequency activity in temporal processing, we recorded LFPs from rodent MFC and the dorsomedial striatum during performance of an interval-timing task with two intervals in rodents. Because our prior work has found delta (1–4 Hz) and theta (4–8 Hz) activity associated with temporal processing, we restricted our analyses to these bands in the present manuscript (Narayanan et al., 2013; Cavanagh and Frank, 2014; Parker et al., 2014, 2015; Laubach et al., 2015). We tested the hypothesis that delta/theta activity is related to temporal processing in corticostratial circuits. We found cue-triggered activity in delta and theta bands in MFC. Delta activity was found in MFC and striatum across temporal intervals, and was observed around rewarded responses in MFC. Frontostriatal delta/theta activity was related to when animals responded in time during the interval. These data indicate that delta/theta activity in corticostratial circuits is involved in the temporal control of action.

## MATERIALS AND METHODS

### Subjects

Eight Long-Evans rats (age 2 months; 200–225 g) were trained to perform an interval-timing task using standard operant procedures. Animals were motivated by regulated access to water, while food was available *ad libitum*. Rats consumed 5–6 ml of water/100 g body weight each day. 5–10 ml were consumed during the behavioral session and any additional water needed was provided 1–3 h after each behavioral session in the home cage. Rats were singly housed and kept on a 12-h light/dark cycle; all experiments took place during the light cycle. Rats were kept at ~90% of their free-access body weight during these experiments, and received 1 day of free access to water per week. All procedures

were approved by the Animal Care and Use Committee at the University of Iowa.

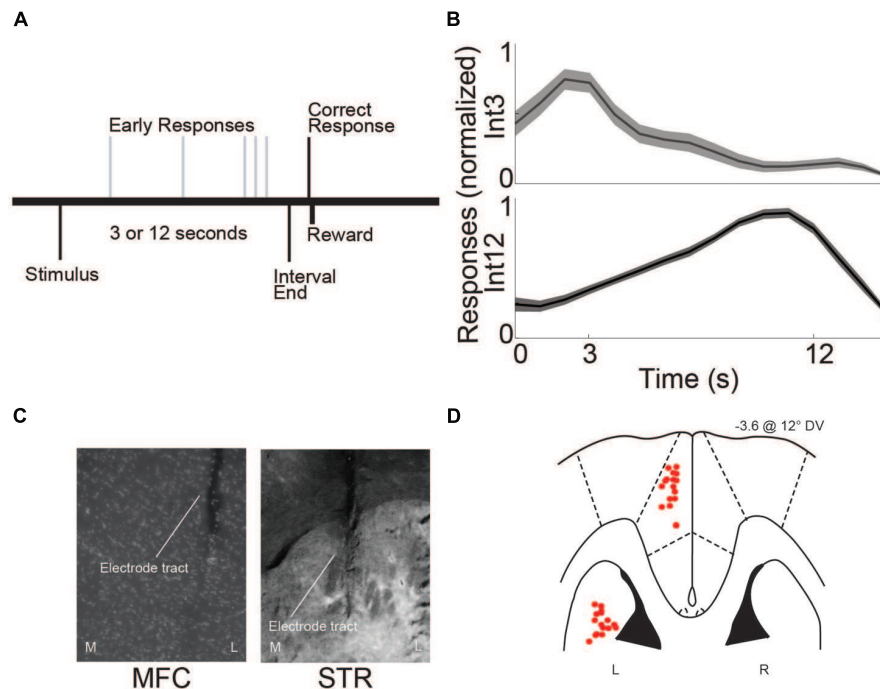
### Interval-Timing Task

Rats were trained on the interval-timing task with a standard operant approach described in detail previously (Narayanan et al., 2012; Parker et al., 2013). First, animals went through fixed-ratio training to make operant lever presses to receive water reward. Next, animals were trained in a 12 s fixed-interval timing task where rewards were delivered for responses made following a 12 s interval (**Figure 1A**). Rewarded presses were signaled by a click and an ‘off’ houselight. Each rewarded trial was immediately followed by a 6, 8, 10, or 12 s pseudorandom intertrial interval which concluded with an ‘on’ houselight signaling the beginning of the next trial. Responses occurring before 12 s were not reinforced. The houselight stayed on from trial onset until the onset of the intertrial interval. Training sessions were 60 min long. Importantly, rodents were allowed to make multiple responses per trial. Average response times were used to determine central tendency of response time per trial. The timing of each response was used to generate time-response histograms. To compare across animals, time-response histograms were normalized to the highest response rate during the interval. After animals learned the 12 s interval—indicated by a peak in their time-response histograms around 12 s—a second delay of 3 s was added. This 3 s interval was signaled with an additional light on the right side of the lever. Operant chambers (MedAssociates, St Albans, VT, USA) were equipped with a lever, a drinking tube, and a speaker driven to produce an 8 kHz tone at 72 dB. Behavioral arenas were housed in sound-attenuating chambers (MedAssociates). Water rewards were delivered via a pump (MedAssociates) connected to a standard metal drinking tube (AnCare) via Tygon tubing.

### Surgical and Perfusion Procedures

Rats trained in the interval-timing task were implanted with a microwire array in MFC or dorsomedial striatum according to procedures described previously (Narayanan and Laubach, 2006). Briefly, animals were anesthetized using Ketamine (100 mg/kg) and Xylazine (10 mg/kg). A surgical level of anesthesia was maintained with hourly (or as needed) Ketamine supplements (10 mg/kg). Under aseptic surgical conditions, the scalp was retracted and the skull was leveled between bregma and lambda. A craniotomy was drilled over the area above MFC or dorsomedial striatum and four holes were drilled for skull screws. A microelectrode array consisting of 50  $\mu\text{m}$  stainless steel wires (250  $\mu\text{m}$  between wires and rows; impedance measured *in vitro* at ~400 k $\Omega$ ; Plexon: Dallas, TX) configured in 4  $\times$  4 ( $n = 4$ ) or 2  $\times$  8 ( $n = 4$ ) was implanted in eight animals in either MFC (coordinates from bregma: AP +3.2, ML  $\pm$  1.2, DV  $-3.6$  @ 12° in the lateral plane) or in dorsomedial striatum (coordinates from bregma: AP +0.0, ML  $\pm$  4.2, DV  $-3.6$  @ 12° in the lateral plane). The electrode ground wire was wrapped around the skull screws. Electrode arrays were inserted while recording neuronal activity to verify implantation in layer II/III of MFC or in the most dorsal portion of dorsomedial striatum. The craniotomy was sealed with cyanoacrylate (‘SloZap’, Pacer Technologies,





**FIGURE 1 | Interval-timing task and histological validation. (A)** Rats estimated a 3 or 12 s interval by pressing a lever. The first response after the end of the interval was rewarded with water. Multiple responses per trial were permitted. **(B)** Average response-timing curves from all eight animals included in the study. Responses were normalized within both 3 s (Int3) and 12 s (Int12) trials. **(C)** Representative photomicrograph of electrode tracts in the left hemisphere of medial frontal cortex (MFC) and dorsomedial striatum (STR) from a coronal view. M, Medial, L, Lateral. **(D)** Horizontal image of histological reconstruction from the four animals implanted with multielectrode arrays either in the MFC or STR. L, Left, R, Right. Red circles correspond to the locations of electrodes used in the study.

Rancho Cucamonga, CA, USA) accelerated by 'ZipKicker' (Pacer Technologies) and methyl methacrylate (i.e., dental cement; AM Systems, Port Angeles, WA, USA). Following implantation, animals recovered for 1 week before being reacclimated to behavioral and recording procedures.

Following experiments, rats were anesthetized, sacrificed by injections of 100 mg/kg sodium pentobarbital, and transcardially perfused with 4% formalin. Brains were post-fixed in a solution of 4% formalin and 20% sucrose before being sectioned on a freezing microtome. Brain slices were mounted on gelatin-subbed slides and stained for cell bodies using DAPI. Histological reconstruction was completed using postmortem analysis of electrode placements by confocal microscopy or stereology microscopy in each animal. These data were used to determine electrode location within MFC or dorsomedial striatum (Figures 1C,D).

## Neurophysiological Recordings

Neuronal ensemble recordings in MFC or dorsomedial striatum were made using a multi-electrode recording system (Plexon, Dallas, TX, USA). LFPs were recorded using wide-band boards with bandpass filters between 0.07 and 8000 Hz. Analysis of neuronal activity and quantitative analysis of basic firing properties were carried out using NeuroExplorer (Nex Technologies, Littleton, MA, USA) and with custom routines for MATLAB. Microwire electrode arrays were comprised of 16

electrodes. In each animal, one electrode without single units was reserved for local referencing and filtering out of noise, yielding 15 electrodes per rat. LFPs were recorded from four low-noise electrodes in each rodent. We recorded LFPs using wide-band boards with analog filters between 0.7 and 100 Hz.

## Time-Frequency Analyses

In line with our previous work, all analyses were restricted to delta and theta bands (Narayanan et al., 2013; Cavanagh and Frank, 2014; Parker et al., 2014, 2015; Laubach et al., 2015). Time-frequency calculations were computed using custom-written MATLAB routines (Cavanagh et al., 2009). Time-frequency measures were computed by taking the inverse FFT of the convolution of a fast Fourier transformed (FFT) LFP power spectrum and a set of complex Morlet wavelets (defined as

a Gaussian-windowed complex sine wave:  $e^{i2\pi tf} e^{-\frac{t^2}{2\sigma^2}}$  where  $t$  is time,  $f$  is frequency [increasing from 1 to 50 Hz in 50 logarithmically spaced steps], and  $\sigma$  is scaling, defined as cycles/( $2\pi f$ ), with four cycle wavelets) (Narayanan et al., 2013; Parker et al., 2014, 2015; Laubach et al., 2015). We varied the number of cycles and other parameters to balance time-frequency resolution for the bands we were interested in here (delta/theta bands) and the time windows used for analysis ( $\sim 1$  s). Wavelet transformation results in estimates of instantaneous power which were subsequently normalized to a decibel (dB) scale ( $10 \cdot \log_{10}[\text{power}(t)/\text{power}(\text{baseline})]$ ),

allowing a direct comparison of effects across frequency bands. Hypothesis-driven statistical significance was computed via a paired *t*-test in the delta (1–4 Hz) or theta (4–8 Hz) frequency bands by calculating the average power change in a period of interest vs. baseline across all subjects. We defined the baseline period as –500 to –300 ms prior to stimulus presentation. For cue-evoked analyses, 0–500 ms post-stimulus onset was compared to baseline. For whole-trial analyses, 0.5 (to exclude the immediate post-stimulus period) to 3 s post-stimulus onset was compared to baseline for 3 s trials (Int3 trials) and 0.5–12 s post-stimulus onset was compared to baseline for 12 s trials (Int12 trials). For response analyses, mean activity from –500 to 0 ms pre-response in Int3 and Int12 trials was compared to the mean baseline. Rewarded presses were the first press after interval end that resulted in reward (3 s after the cue for Int3; 12 s after the cue for Int12). Unrewarded presses occurred prior to the end of the interval (0–2.9 s for Int3; 0–11.9 s for Int12). To match variance with rewarded trials we randomly subsampled the number of unrewarded presses so that comparisons between rewarded and unrewarded trials had the same number of trials in each category. Error bars were computed from variance across subjects and represent the standard error of the mean.

## Linear Models

To investigate the relationship between the timing of responses and frontostriatal field potentials, we used linear regression (*fitlm.m* in MATLAB) where delta or theta activity calculated from –500–0 ms prior to the response was regressed against the timing of the response. To reduce the role of cue-related activity, this analysis was calculated from 500 ms to 3 s for Int3 trials and 500 ms–12 s for Int12 trials. Delta and theta power were derived according to methods above. Slope was calculated as the change in delta/theta power ( $\Delta$ dB) over the change in the timing of responses (in seconds). Significant linear fits were derived from analysis of variance.

## RESULTS

### Interval-Timing Behavior

The eight rats used in this study were trained on the 3 and 12 s interval-timing task described above (Int3 and Int12, respectively; **Figure 1A**). The mean response time for Int3 was  $4.8 \pm 0.28$  s and the mean response time for Int12 was  $11.7 \pm 0.43$  s (**Figure 1B**). Mean response times were significantly different on Int3 vs. Int12 trials [ $t_{(7)} = 12.5$ ,  $p < 0.001$ ]. The variability of interval-timing behavior was similar to that seen in previous studies (Kim et al., 2013; Parker et al., 2014, 2015; Xu et al., 2014).

### Cue-Triggered Delta/Theta Activity in Medial Frontal Cortex and Dorsomedial Striatum

To test the idea that delta and theta bands are modulated during interval timing, we recorded field potentials from MFC

or dorsomedial striatum of rats trained to perform an interval-timing task (**Figures 1C,D**). In MFC, a large cue-triggered ERP was found (**Figure 2A**). The average latency to the positive peak on Int3 trials was  $126 \pm 6.4$  ms, followed by a negative peak at  $208 \pm 8.3$  ms. On Int12 trials the average latency to the positive peak was  $122 \pm 5.6$  ms, followed by a negative peak at  $192 \pm 8.4$  ms. Time-frequency analysis demonstrated strong delta and theta activity 0–0.5 s after cue onset (**Figures 2B,C**). Direct comparison of MFC activity revealed significant cue-related modulation of delta and theta bands relative to baseline in both Int3 trials [delta:  $t_{(15)} = 4.1$ ,  $p < 0.01$ ; theta:  $t_{(15)} = 2.7$ ,  $p < 0.05$ ] and Int12 [delta:  $t_{(15)} = 5.2$ ,  $p < 0.01$ ; theta:  $t_{(15)} = 3.1$ ,  $p < 0.01$ ; **Figure 2D**]. Notably, cue-related activity was similar on Int3 and Int12 trials (**Figure 2D**). These data indicate that delta and theta modulations early in the trial are cue-triggered and do not significantly differ based on interval length.

In dorsomedial striatum, a less distinct pattern was observed. On Int3 trials, the average latency to the positive peak was  $130 \pm 4.6$  ms, followed by a less prominent negative peak at  $192 \pm 9.8$  ms. On Int12 trials, the average latency to the positive peak was  $130 \pm 4.2$  ms, followed by a less distinct negative potential at  $178 \pm 10.9$  ms. Low-frequency modulation by the cue was visible on both trial types (**Figures 2E–G**). However, there was only a significant increase from baseline in the striatal delta band on Int3 trials [delta:  $t_{(15)} = 4.7$ ,  $p < 0.01$ ; **Figure 2H**]. Taken together, these data demonstrate that cue-related delta and theta activity is primarily modulated in MFC.

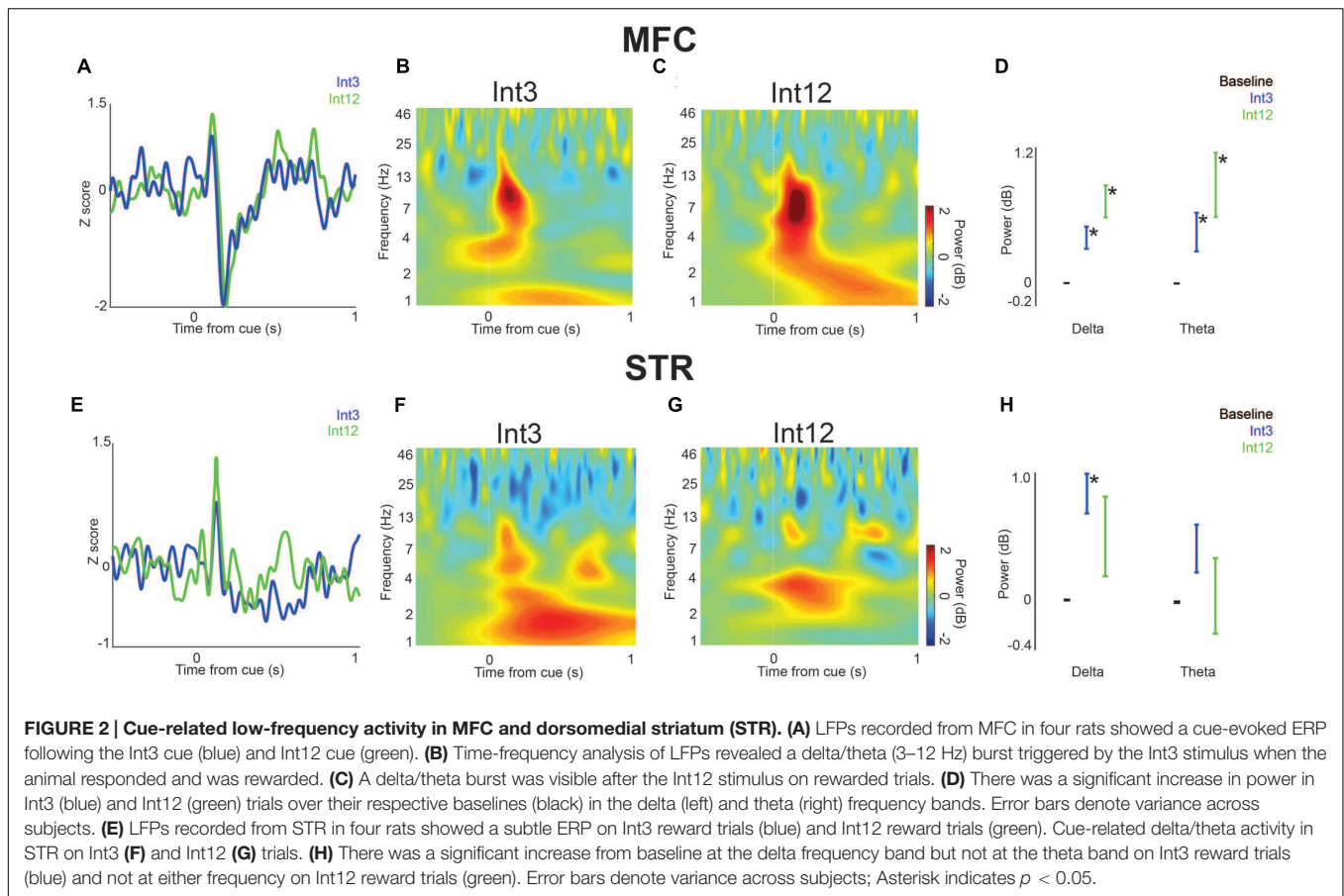
### Interval-Related Delta/Theta Activity in Medial Frontal Cortex and Dorsomedial Striatum

Next, we examined field potentials over the duration of the trial in MFC and dorsomedial striatum. To compare Int3 and Int12 trials, we averaged activity in delta and theta bands across the interval. In MFC, we found that average delta and theta activity bands across the interval were significantly higher than baseline for Int12 trials [0.5–12 s following cue; delta:  $t_{(15)} = 2.4$ ,  $p < 0.05$ ; theta:  $t_{(15)} = 2.4$ ,  $p < 0.05$ ; **Figures 3A–C**]. No significant difference from baseline was seen on Int3 trials. These data indicate that medial frontal delta activity is engaged on longer intervals.

In the dorsomedial striatum, only delta activity was significantly higher than baseline for Int3 and Int12 conditions [Int3—delta:  $t_{(15)} = 3.2$ ,  $p < 0.01$ ; Int12—delta:  $t_{(15)} = 2.5$ ,  $p < 0.05$ ; **Figures 3D–F**]. These data suggest that delta activity is modulated across temporal intervals in frontostriatal circuits.

### Medial Frontal Delta Activity Is Related to Rewarded Responses

Next, we analyzed field potentials around lever presses. Trials on which the animal pressed the lever after the 12 s interval were rewarded. We found marked press-related potentials in both MFC and dorsomedial striatum (**Figure 4A**). The average latency to the positive peak in MFC on rewarded trials was  $-8 \pm 16.5$  ms, followed by a negative peak at  $206 \pm 11.6$  ms. The average latency to the positive peak on unrewarded responses was  $-82 \pm 8.5$  ms,



**FIGURE 2 | Cue-related low-frequency activity in MFC and dorsomedial striatum (STR).** (A) LFPs recorded from MFC in four rats showed a cue-evoked ERP following the Int3 cue (blue) and Int12 cue (green). (B) Time-frequency analysis of LFPs revealed a delta/theta (3–12 Hz) burst triggered by the Int3 stimulus when the animal responded and was rewarded. (C) A delta/theta burst was visible after the Int12 stimulus on rewarded trials. (D) There was a significant increase in power in Int3 (blue) and Int12 (green) trials over their respective baselines (black) in the delta (left) and theta (right) frequency bands. Error bars denote variance across subjects. (E) LFPs recorded from STR in four rats showed a subtle ERP on Int3 reward trials (blue) and Int12 reward trials (green). Cue-related delta/theta activity in STR on Int3 (F) and Int12 (G) trials. (H) There was a significant increase from baseline at the delta frequency band but not at the theta band on Int3 reward trials (blue) and not at either frequency on Int12 reward trials (green). Error bars denote variance across subjects; Asterisk indicates  $p < 0.05$ .

followed by a negative peak at  $6 \pm 10.9$  ms. In MFC, only delta activity was significantly higher on rewarded responses both compared to baseline [ $t_{(15)} = 3.0$ ,  $p < 0.05$ ] and compared to unrewarded presses [ $t_{(15)} = 2.6$ ,  $p < 0.05$ ; **Figures 4B–D**].

A similar press-related potential was found in dorsomedial striatum (**Figure 4E**). In dorsomedial striatum, the average latency to the positive peak on rewarded responses was  $4 \pm 11.8$  ms, followed by a negative peak at  $190 \pm 18.3$  ms. The average latency to the positive peak on unrewarded responses was  $-44 \pm 10.4$  ms, followed by a negative peak at  $172 \pm 18.0$  ms. In contrast to MFC, striatal delta power was significantly higher than baseline on both rewarded and unrewarded responses [rewarded:  $t_{(15)} = 4.3$ ,  $p < 0.01$ ; unrewarded:  $t_{(15)} = 2.6$ ,  $p < 0.05$ ; **Figures 4B–H**]. There was not a significant difference between rewarded and unrewarded responses in either the delta or theta bands. Thus, in MFC delta activity was associated with rewarded presses, while in dorsomedial striatum delta activity was associated with all lever presses. These data provide insight into delta and theta activity throughout corticostriatal circuits during interval timing.

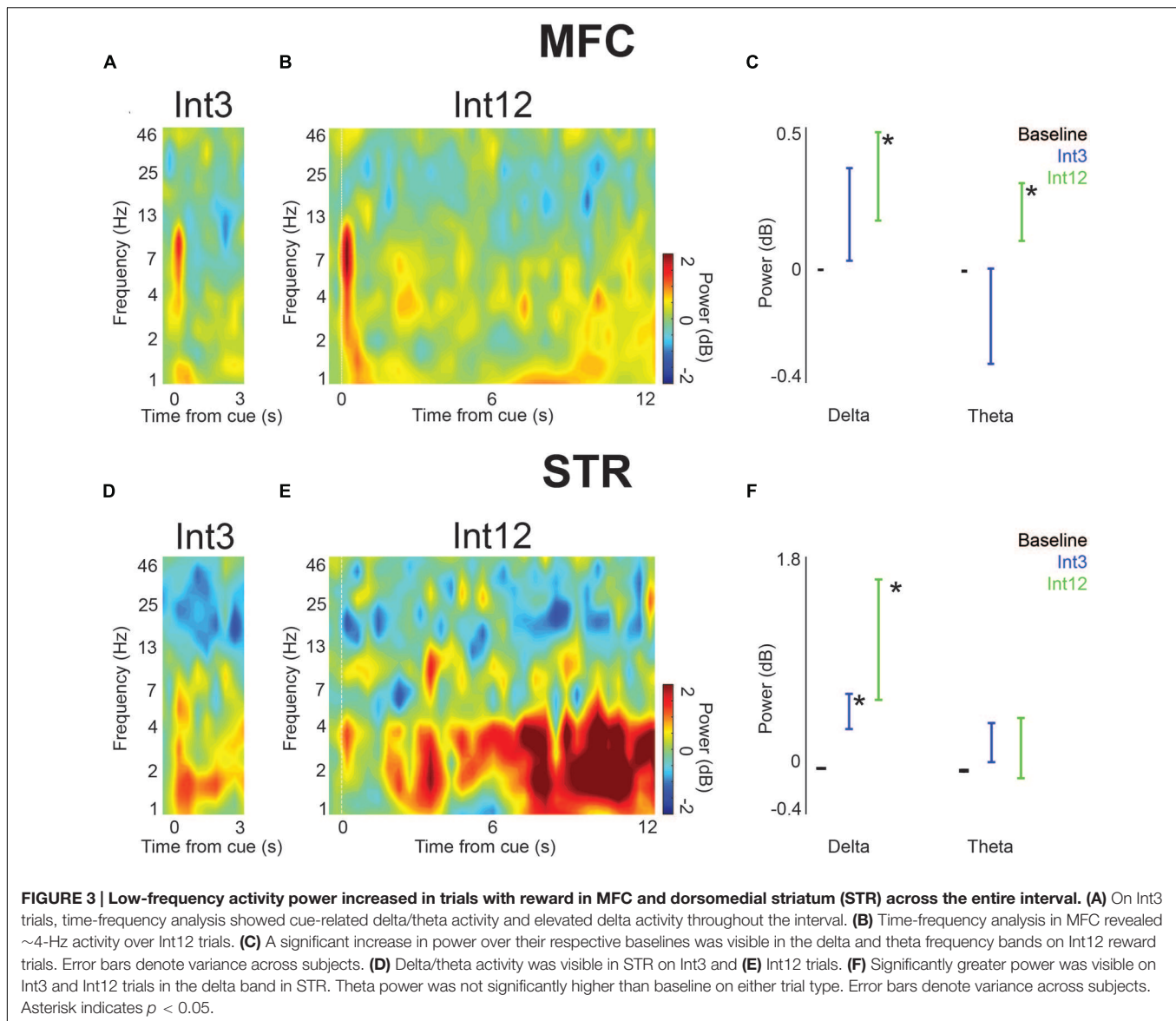
## Delta/Theta Activity and Temporal Control of Responding

To examine how delta/theta activity in the MFC and striatum predicted when animals responded, we used linear models

of frontostriatal field potential activity vs. response time. We examined delta and theta activity  $-500$ – $0$  ms prior to lever press. Significant linear fits are indicated in **Table 1** as changes in delta or theta power in dB per second of response time. In MFC, theta activity immediately prior to lever press predicted when animals responded for both Int3 and Int12 trials (Int3:  $p < 0.03$ ; Int12:  $p < 0.02$ ). By contrast, in dorsomedial striatum delta activity immediately prior to lever press predicted when animals responded (Int3:  $p < 0.0001$ ; Int12:  $p < 10^{-8}$ ). These data indicate that response-related theta activity in MFC and delta activity in the striatum depends on when animals press the lever during the interval, and indicate that these bands are involved in the temporal control of action in frontostriatal circuits.

## DISCUSSION

Here we studied rodent frontostriatal circuits using LFPs during an interval-timing task. Because our previous work implicates low-frequency activity in delta and theta ranges in the temporal control of action, we focused on these bands in this study. We report four main findings. First, we observed cue-triggered modulations in delta and theta activity primarily in MFC. Second, we found delta activity in MFC and dorsomedial striatum across the temporal interval. Thirdly, we observed increased delta activity in MFC prior to rewarded responses, while striatal



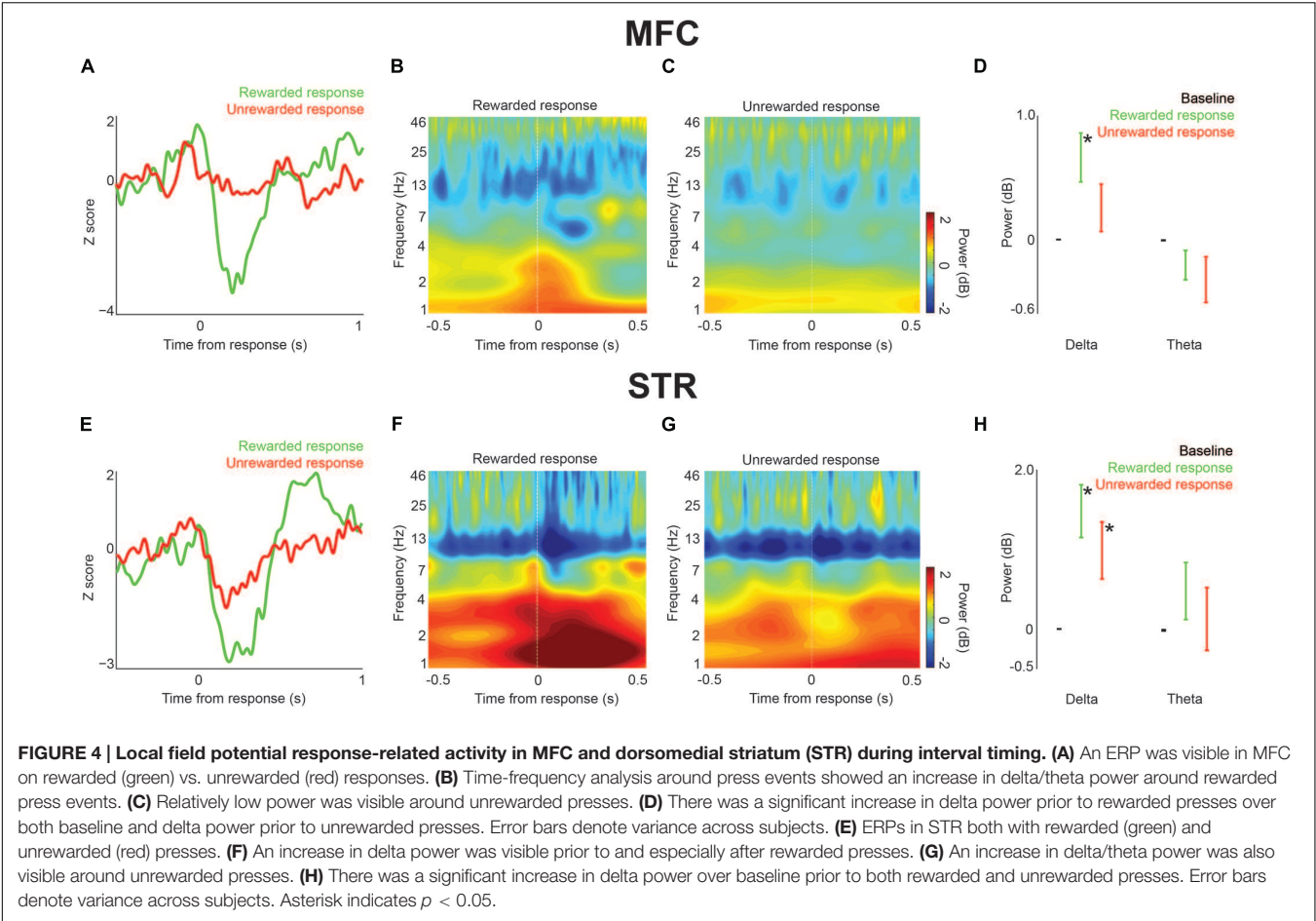
delta modulation was observed prior to all responses. Finally, theta activity in MFC and delta activity in the striatum was related to when animals responded during the interval. These data contribute to an understanding of low-frequency activity in corticostriatal circuits that is highly conserved across humans and rodents (Cavanagh et al., 2012; Narayanan et al., 2013; Parker et al., 2015). This similarity could help approach human EEG as well as human intracortical recordings from patients undergoing epilepsy or deep-brain stimulation surgeries (Brown and Williams, 2005; Emeric et al., 2008; Kingyon et al., 2015).

The low-frequency activity observed in MFC after the instructional cue is broadly consistent with past research. Frontal theta and delta bands during elementary cognitive tasks are similar between humans and rodents (Narayanan et al., 2013; Parker et al., 2015; Warren et al., 2015). To our knowledge these are the first field potential data from the dorsomedial striatum in

rodents during a timing task. Delta and theta bands have been associated with errors, conflict, working memory, and attention (Curtis and D'Esposito, 2003; Emeric et al., 2008; Liebe et al., 2012; Totah et al., 2013; Cavanagh and Frank, 2014; Chen et al., 2014; Parker et al., 2014; Laubach et al., 2015). Activity in this range may provide a means of synchronizing frontal activity with other brain regions (Fujisawa and Buzsáki, 2011). The pronounced burst of low-frequency activity in MFC following the cue is similar to that seen in our previous work (Parker et al., 2014, 2015). This activity was not unique to either of the interval lengths—it is likely related to the salience of the cue and communicates the need for cognitive control (Cavanagh and Frank, 2014).

Low-frequency activity was observed throughout the duration of interval-timing tasks in MFC and dorsomedial striatum. Delta and theta activity was significantly increased in MFC on





longer-interval trials and was increased on both trial types in dorsomedial striatum. This result suggests that sustained low-frequency activity in the MFC is more engaged on intervals of longer, more demanding duration. Moreover, low-frequency activity in both areas was significantly related to when animals made a response in time. That is, activity in these bands was different if the animal pressed the lever early or late in the interval, indicating that pre-response delta/theta activity can be influenced by temporal preparation of responding (Table 1).

**TABLE 1 | Relationship of response-related delta/theta power compared to the timing of responses by interval length (Int3–3 s intervals; Int12–12 s intervals): Output of the linear regression model calculated for the change in power (ΔdB) over time (s).**

		DELTA		THETA	
		Slope (ΔdB/s)	p-value	Slope (ΔdB/s)	p-value
Int3	MFC	0.20	0.21	−0.38	0.03
	STR	1.77	0.0001	−0.11	0.76
Int12	MFC	0.01	0.50	0.05	0.02
	STR	0.36	10 <sup>−8</sup>	0.17	10 <sup>−6</sup>

Significant linear fits were found for theta activity in the medial frontal cortex (MFC) and delta activity in the striatum (STR) in relation to the timing of response (ΔdB/s).

Strong reward-related delta activity was observed around responses. Delta activity was increased in dorsomedial striatum on all responses, regardless of reward. Delta activity has been reported from rodent cortex and striatum, and has been associated with motor action, reward processing, and temporal expectation (Stefanics et al., 2010; Cavanagh et al., 2012; Laubach et al., 2015). We observed different relationships between delta activity and interval-timing behavior in the MFC and dorsomedial striatum. One possibility is that medial frontal delta activity reflects reward anticipation during interval timing (Cavanagh et al., 2012; Narayanan et al., 2013; Parker et al., 2014, 2015).

Low frequencies in MFC may represent temporal processing while field potentials in dorsomedial striatum may also reflect the motor output of this processing. Many lines of evidence suggest that the striatum is critical for interval timing (Matell et al., 2003; Matell and Meck, 2004; Meck, 2006; Coull et al., 2011; Merchant et al., 2013). Notably, spiking activity in striatal ensembles robustly encodes temporal processing (Matell et al., 2003; Mello et al., 2015). In contrast, LFP may reflect input to the striatum from a variety of sources (Wall et al., 2013)—MFC being but one of them—making temporal signals relatively difficult to isolate at the level of field potentials. High-frequency gamma and beta activity in the primate striatum have been linked to interval

timing, particularly in terms of coherence and entrainment of neural populations (Bartolo et al., 2014). It remains to be seen how striatal field potentials couple with neuronal activity in other brain areas such as MFC.

This study is limited by several factors. Rodent LFP recordings are not a perfect analog to EEG in human subjects, though progress has been made recently in comparing these two systems (Narayanan et al., 2013; Parker et al., 2015; Warren et al., 2015). Due to the scope of this study, we constrained our analyses to delta and theta activity as we have found these bands to be reliably modulated in prior human and rodent work during timing tasks (Narayanan et al., 2013; Parker et al., 2014, 2015; Laubach et al., 2015). Although striatal delta power was distinct on unrewarded vs. rewarded lever presses and correlated with response time, rewarded presses generally occur when the response rate is high and could be affected by movement. By contrast, in MFC, theta power had a more complex relationship with movement on Int3 and Int12 trials and could not be directly accounted for by movement-related activity. Finally, we did not examine sensory aspects of frontostriatal LFPs. Future work will look at other frequency bands, neuronal spike data, and at the interactions between spikes and LFPs. Because

recordings in MFC and dorsomedial striatum were done in separate groups of animals, we are unable to make conclusions about the simultaneous activity of corticostratial ensembles. In subsequent studies we hope to address these issues by more directly comparing rodent and human data, exploring changes in LFP activity during learning of temporal rules, and looking at the simultaneous activity of neuronal ensembles in both of these structures.

## AUTHOR CONTRIBUTIONS

EE, KP, and NN designed these experiments; EE and KP collected data for these experiments, RR, EE, and NN analyzed data, and EE, RR, RK, and NN wrote the paper.

## FUNDING

This work was supported by National Institutes of Neurological Disorders and Stroke Grants R01 NS089470 and São Paulo Research Foundation (FAPESP) 2014/22817-1.

## REFERENCES

- Bartolo, R., Prado, L., and Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *J. Neurosci.* 34, 3910–3923. doi: 10.1523/JNEUROSCI.2679-13.2014
- Brown, P., and Williams, D. (2005). Basal ganglia local field potential activity: character and functional significance in the human. *Clin. Neurophysiol.* 116, 2510–2519. doi: 10.1016/j.clinph.2005.05.009
- Brown, S. W., Collier, S. A., and Night, J. C. (2013). Timing and executive resources: dual-task interference patterns between temporal production and shifting, updating, and inhibition tasks. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 947–963. doi: 10.1037/a0030484
- Buhusi, C. V., and Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6, 755–765. doi: 10.1038/nrn1764
- Cavanagh, J. F., Cohen, M. X., and Allen, J. J. B. (2009). Prelude to and resolution of an error: EEG phase synchrony reveals cognitive control dynamics during action monitoring. *J. Neurosci. Off. J. Soc. Neurosci.* 29, 98–105. doi: 10.1523/JNEUROSCI.4137-08.2009
- Cavanagh, J. F., and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* 18, 414–421. doi: 10.1016/j.tics.2014.04.012
- Cavanagh, J. F., Zambrano-Vazquez, L., and Allen, J. J. B. (2012). Theta lingua franca: a common mid-frontal substrate for action monitoring processes. *Psychophysiology* 49, 220–238. doi: 10.1111/j.1469-8986.2011.01293.x
- Chen, H., Wang, Y., Yang, L., Hu, C., Ke, X., Fan, Z., et al. (2014). Predictive nature of prefrontal theta oscillation on the performance of trace conditioned eyeblink responses in guinea pigs. *Behav. Brain Res.* 265, 121–131. doi: 10.1016/j.bbr.2014.02.020
- Coull, J. T., Cheng, R.-K., and Meck, W. H. (2011). Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* 36, 3–25. doi: 10.1038/npp.2010.113
- Curtis, C. E., and D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends Cogn. Sci.* 7, 415–423. doi: 10.1016/S1364-6613(03)00197-9
- Emeric, E. E., Brown, J. W., Leslie, M., Pouget, P., Stuphorn, V., and Schall, J. D. (2008). performance monitoring local field potentials in the medial frontal cortex of primates: anterior cingulate cortex. *J. Neurophysiol.* 99, 759–772. doi: 10.1152/jn.00896.2006
- Fujisawa, S., and Buzsáki, G. (2011). A 4 Hz oscillation adaptively synchronizes prefrontal, VTA, and hippocampal activities. *Neuron* 72, 153–165. doi: 10.1016/j.neuron.2011.08.018
- Gu, B.-M., van Rijn, H., and Meck, W. H. (2015). Oscillatory multiplexing of neural population codes for interval timing and working memory. *Neurosci. Biobehav. Rev.* 48, 160–185. doi: 10.1016/j.neubiorev.2014.10.008
- Kim, J., Ghim, J.-W., Lee, J. H., and Jung, M. W. (2013). Neural correlates of interval timing in rodent prefrontal cortex. *J. Neurosci. Off. J. Soc. Neurosci.* 33, 13834–13847. doi: 10.1523/JNEUROSCI.1443-13.2013
- Kingyon, J., Behroozmand, R., Kelley, R., Oya, H., Kawasaki, H., Narayanan, N. S., et al. (2015). High-gamma band fronto-temporal coherence as a measure of functional connectivity in speech motor control. *Neuroscience* 305, 15–25. doi: 10.1016/j.neuroscience.2015.07.069
- Kurti, A. N., and Matell, M. S. (2011). Nucleus accumbens dopamine modulates response rate but not response timing in an interval timing task. *Behav. Neurosci.* 125, 215–225. doi: 10.1037/a0022892
- Laubach, M., Caetano, M. S., and Narayanan, N. S. (2015). Mistakes were made: neural mechanisms for the adaptive control of action initiation by the medial prefrontal cortex. *J. Physiol. Paris* 109, 104–107. doi: 10.1016/j.jphysparis.2014.12.001
- Liebe, S., Hoerzer, G. M., Logothetis, N. K., and Rainer, G. (2012). Theta coupling between V4 and prefrontal cortex predicts visual short-term memory performance. *Nat. Neurosci.* 15, S1–S2. doi: 10.1038/nn.3038
- Malapani, C., Rakitin, B., Levy, R., Meck, W. H., Deweer, B., Dubois, B., et al. (1998). Coupled temporal memories in parkinson's disease: a dopamine-related dysfunction. *J. Cogn. Neurosci.* 10, 316–331. doi: 10.1162/089892998562762
- Matell, M. S., and Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Brain Res. Cogn. Brain Res.* 21, 139–170. doi: 10.1016/j.cogbrainres.2004.06.012
- Matell, M. S., Meck, W. H., and Nicolelis, M. A. L. (2003). Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behav. Neurosci.* 117, 760–773. doi: 10.1037/0735-7044.117.4.760
- Meck, W. H. (2006). Neuroanatomical localization of an internal clock: a functional link between mesolimbic, nigrostriatal, and mesocortical dopaminergic systems. *Brain Res.* 1109, 93–107. doi: 10.1016/j.brainres.2006.06.031
- Meck, W. H., and Benson, A. M. (2002). Dissecting the brain's internal clock: how frontal-striatal circuitry keeps time and shifts attention. *Brain Cogn.* 48, 195–211. doi: 10.1006/brcg.2001.1313

- Mello, G. B. M., Soares, S., and Paton, J. J. (2015). A scalable population code for time in the striatum. *Curr. Biol.* 25, 1113–1122. doi: 10.1016/j.cub.2015.02.036
- 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. doi: 10.1146/annurev-neuro-062012-170349
- Narayanan, N. S., Cavanagh, J. F., Frank, M. J., and Laubach, M. (2013). Common medial frontal mechanisms of adaptive control in humans and rodents. *Nat. Neurosci.* 16, 1888–1897. doi: 10.1038/nn.3549
- Narayanan, N. S., Land, B. B., Solder, J. E., Deisseroth, K., and DiLeone, R. J. (2012). Prefrontal D1 dopamine signaling is required for temporal control. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20726–20731. doi: 10.1073/pnas.1211258109
- Narayanan, N. S., and Laubach, M. (2006). Top-down control of motor cortex ensembles by dorsomedial prefrontal cortex. *Neuron* 52, 921–931. doi: 10.1016/j.neuron.2006.10.021
- Oprisan, S. A., and Buhusi, C. V. (2011). Modeling pharmacological clock and memory patterns of interval timing in a striatal beat-frequency model with realistic, noisy neurons. *Front. Integr. Neurosci.* 5:52. doi: 10.3389/fnint.2011.00052
- Parker, K. L., Alberico, S. L., Miller, A. D., and Narayanan, N. S. (2013). Prefrontal D1 dopamine signaling is necessary for temporal expectation during reaction time performance. *Neuroscience* 255, 246–254. doi: 10.1016/j.neuroscience.2013.09.057
- Parker, K. L., Chen, K.-H., Kingyon, J. R., Cavanagh, J. F., and Narayanan, N. S. (2014). D1-dependent 4 Hz oscillations and ramping activity in rodent medial frontal cortex during interval timing. *J. Neurosci.* 34, 16774–16783. doi: 10.1523/JNEUROSCI.2772-14.2014
- Parker, K. L., Chen, K.-H., Kingyon, J. R., Cavanagh, J. F., and Narayanan, N. S. (2015). Medial frontal 4-Hz activity in humans and rodents is attenuated in PD patients and in rodents with cortical dopamine depletion. *J. Neurophysiol.* 114, 1310–1320. doi: 10.1152/jn.00412.2015
- Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., and Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. *J. Neurosci. Off. J. Soc. Neurosci.* 30, 13578–13585. doi: 10.1523/JNEUROSCI.0703-10.2010
- Totah, N. K. B., Jackson, M. E., and Moghaddam, B. (2013). Preparatory attention relies on dynamic interactions between prefrontal cortex and anterior cingulate cortex. *Cereb. Cortex* 23, 729–738. doi: 10.1093/cercor/bhs057
- Uylings, H. B. M., Groenewegen, H. J., and Kolb, B. (2003). Do rats have a prefrontal cortex? *Behav. Brain Res.* 146, 3–17. doi: 10.1016/j.bbr.2003.09.028
- Wall, N. R., De La Parra, M., Callaway, E. M., and Kreitzer, A. C. (2013). Differential Innervation of Direct- and Indirect-Pathway Striatal Projection Neurons. *Neuron* 79, 347–360. doi: 10.1016/j.neuron.2013.05.014
- Ward, R. D., Kellendonk, C., Kandel, E. R., and Balsam, P. D. (2012). Timing as a window on cognition in schizophrenia. *Neuropharmacology* 62, 1175–1181. doi: 10.1016/j.neuropharm.2011.04.014
- Warren, C. M., Hyman, J. M., Seamans, J. K., and Holroyd, C. B. (2015). Feedback-related negativity observed in rodent anterior cingulate cortex. *J. Physiol. Paris* 109, 87–94. doi: 10.1016/j.jphysparis.2014.08.008
- Xu, M., Zhang, S., Dan, Y., and Poo, M. (2014). Representation of interval timing by temporally scalable firing patterns in rat prefrontal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 111, 480–485. doi: 10.1073/pnas.1321314111

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Emmons, Ruggiero, Kelley, Parker and Narayanan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Rubberband Effect in Temporal Control of Mismatch Negativity

Lingyan Wang<sup>1,2</sup>, Xiaoxiong Lin<sup>1</sup>, Bin Zhou<sup>3</sup>, Ernst Pöppel<sup>1,3,4</sup> and Yan Bao<sup>1,4\*</sup>

<sup>1</sup> School of Psychological and Cognitive Sciences, Key Laboratory of Machine Perception (Ministry of Education) and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China, <sup>2</sup> Departments of Neurosurgery and Neuroscience, Baylor College of Medicine, Houston, TX, USA, <sup>3</sup> Institute of Psychology, Chinese Academy of Sciences, Beijing, China, <sup>4</sup> Institute of Medical Psychology and Human Science Center, Ludwig-Maximilians-University, Munich, Germany

Mismatch negativity (MMN) is a difference event-related potential (ERP) wave reflecting the brain's automatic reaction to deviant sensory stimuli, and it has been proven to be a useful tool in research on cognitive functions or clinical disorders. In most MMN studies, amplitude, peak latency, or the integral of the responses, in rare cases also the slopes of the responses, have been employed as parameters of the ERP responses for quantitative analyses. However, little is known about correlations between these parameters. To better understand the relations between different ERP parameters, we extracted and correlated several different parameters characterizing the MMN waves. We found an unexpected correlation which gives new insight into the temporal control of MMN: response amplitudes are positively correlated with downside slopes, whereas barely correlated with upside slopes. This result suggests an efficient feedback mechanism for the MMN to return to the baseline within a predefined time window, contradicting an exponential decay function as one might expect. As a metaphor we suggest a rubberband effect for the MMN responses, i.e., the larger the distance of the response from neural equilibrium, the stronger the return force to equilibrium.

**Keywords:** event related potential, mismatch negativity, oddball paradigm, time window, correlation

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College, USA

### Reviewed by:

Paula Maarit Virtala,  
University of Helsinki, Finland  
Tao Liu,  
Zhejiang University, China

### \*Correspondence:

Yan Bao  
baoyan@pku.edu.cn

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 31 March 2016

**Accepted:** 12 August 2016

**Published:** 31 August 2016

### Citation:

Wang L, Lin X, Zhou B, Pöppel E  
and Bao Y (2016) Rubberband Effect  
in Temporal Control of Mismatch  
Negativity. *Front. Psychol.* 7:1299.  
doi: 10.3389/fpsyg.2016.01299

## INTRODUCTION

Mismatch negativity (MMN) is a negative event-related potential (ERP) component when subtracting brain responses to standard stimuli from those to rare stimuli, usually peaking at 150 to 250 ms after deviant onset (Näätänen et al., 2007). MMN was discovered by Näätänen et al. (1978) using an auditory oddball paradigm and equivalent responses have been observed in other sensory modalities such as in vision (Pazo-Alvarez et al., 2003) or in olfaction (Krauel et al., 1999). In a typical auditory oddball paradigm, an infrequent deviant sound is occasionally presented within a sequence of frequent standard sounds, and the participants either actively detect the deviant or ignore the entire sequence while focusing on signals of another modality. The elicitation of MMN in the latter case indicates that it can be observed independent of attention, and such independence is more pronounced in sleeping infants (Cheour et al., 2002; Martynova et al., 2003). MMN apparently discloses a neural mechanism to detect novel stimuli, and it is even elicited in complex situations when abstract rules are violated (Saarinen et al., 1992; Schröger et al., 2007).

In research on MMN, the amplitude of the response, usually quantified by the most negative value within the conventional MMN time window, has been proven to be a useful parameter (Sinkkonen and Tervaniemi, 2000; Paavilainen, 2013). Typically, the MMN amplitude gets larger



when the oddball stimuli become more different, i.e., when the magnitude of oddball's deviation from standard stimuli increases (Näätänen, 1992; Jaramillo et al., 2000; Pakarinen et al., 2007; but see Horváth et al., 2008). For example, Pakarinen et al. (2007) varied the frequency distance between the standard and deviant sounds and found that larger relative to smaller deviation (e.g., 523 and 609 Hz relative to 523 and 546 Hz) led to higher MMN amplitude measured at the Fz electrode. Similar relationships between the magnitude of deviation and the MMN amplitude can be demonstrated on other deviant dimensions as well, such as stimulus intensity (Pakarinen et al., 2007), duration (Jaramillo et al., 2000; Pakarinen et al., 2007), and perceived location (Pakarinen et al., 2007). Moreover, when the number of standard stimuli before a deviant increases, MMN amplitudes also get enhanced (Haenschel et al., 2005).

Beside amplitude, peak latency measured by the time from the deviant onset to the MMN peak is also widely used as an indicator in MMN research (Sinkkonen and Tervaniemi, 2000; Paavilainen, 2013). It has been observed that MMN peak latency gets shorter when stimuli deviation increases (Amenedo and Escera, 2000; Pakarinen et al., 2007). Both MMN amplitude and peak latency are good predictors of behavioral performances. While the accuracy of detecting deviants among series of standards is paralleled by the MMN amplitude (Lang et al., 1990, 1995; Jaramillo et al., 2000; Novitski et al., 2004; Pakarinen et al., 2007), the MMN peak latency in some cases predicts the speed of behavioral responses, i.e., the shorter the latency, the faster the reaction (Novitski et al., 2004; Pakarinen et al., 2007). Furthermore, both amplitude and latency have been well established as biomarkers in clinical cases of psychiatric disorders such as schizophrenia (Kargel et al., 2014) or autism (Roberts et al., 2011). These observations indicate that the MMN amplitude and latency are ecologically relevant indicators, which may reflect the neural operations of sensory novelty detection. Thus, an in-depth analysis of these indicators is necessary.

In previous ERP studies, amplitude or averaged amplitude in a short window are most frequently chosen for statistical analyses (Näätänen et al., 1978). In some cases temporal parameters like onset, offset or peak latency are also analyzed with the onset and offset latency being estimated by the time at the most positive value immediately before and after the MMN peak, respectively (Baldeweg et al., 1999). Sometimes the slopes of the waves indexed by the fluxion of the extrapolated line between the onset and the peak of MMN or the integral of the half MMN wave area are explored (Korostenskaja et al., 2003).

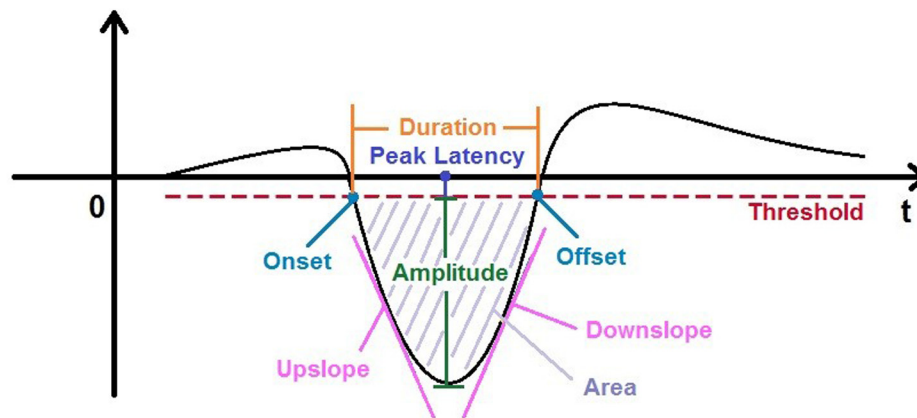
Although various parameters have been used in previous ERP studies, it remains unclear in detail how the different parameters are related to each other. Previous studies provide only limited evidence which shows no clear relationship between the MMN amplitude and different latencies (Lang et al., 1995). With respect to other ERP components rather than MMN, Polich et al. (1997) found negative correlations between the P3 amplitude and its latency, i.e., the larger the P3 amplitude, the shorter the latency, in a task requiring participants to discriminate an infrequent target from frequent standards; this correlation was somewhat biased to the right frontal electrodes, suggesting that

the generation of P3 may involve initially the attentional control system of the right frontal cortex. In another study (Intriligator and Polich, 1994) it was found that the EEG power in lower frequency bands correlated positively with the P3 amplitude, suggesting a possible mediating role of attention resources. These observations indicate that exploring correlations between different parameters is a useful method to look into the neural mechanisms of ERP components.

The above-described research has examined mostly the relationship between the amplitude and latencies; other potentially indicative parameters such as upside and downside slopes are largely ignored. As suggested in some previous studies (Korostenskaja et al., 2003), the upside slope reflects the speed of the rise of MMN wave and indexes the speed of neuronal arousal processes associated with MMN responses. On this basis, it is important to extend the ERP analysis from focusing only on amplitude and latencies to more parameters such as upside and downside slopes as well as other temporal parameters.

In the study presented here, we aimed to systematically examine nine parameters in the MMN waveform, i.e., onset, offset, peak amplitude, averaged peak amplitude (shortened as "ampavg"), duration, area, peak latency, upside and downside slopes (see **Figure 1**), trying to find out whether there exist certain correlations between different parameters of MMN. The data of this analysis were taken from a previous study (Wang et al., 2015), in which MMN amplitudes elicited by frequency deviants of sinusoidal tones in a passive oddball paradigm were compared for four inter-stimulus interval (ISI) conditions (1.5, 3, 4.5, and 6 s). The original results demonstrated significantly larger MMNs over central-frontal scalp areas for shorter ISIs up to 3 s as compared to longer ones, suggesting that the temporal modulation of 3 s provides a basic process of sequential segmentation which can be operated pre-attentively or pre-semantically (Pöppel, 1997, 2009). This finding also posed another question for the current analysis whether the suspected correlations could be present in all or a subset of ISI conditions. Considering MMN as a negative ERP component being present at various ISIs, we focused our analysis on the potential correlations between different MMN parameters, which might capture the substantial underlying neural processes of novelty detection. We anticipated that the suspected correlations should be independent of ISIs since their generation is due to generic operations of the neural system dealing with the departure from the equilibrium.

Among all the parameters we examined, latency, amplitude (including averaged amplitude) and slope (both upside slope and downside slope) were of special interest. Since previous study did not show any clear relationship between MMN latency and amplitude (Lang et al., 1995), we anticipated no correlation between MMN peak latency and amplitude or averaged amplitude as well. Since MMN slopes are waveform-related parameters while latency is a temporal one, we did not anticipate any correlation between MMN peak latency and slopes (both upside and downside slopes) as well. Regarding the relationship between MMN amplitude and slopes, which was of our major interest in the present study, we anticipated a different picture. As suggested previously (Korostenskaja et al.,



**FIGURE 1 |** Nine parameters characterizing the simplified mismatch negativity (MMN) wave.

2003), the upside slope indicates the speed of the rising MMN wave. Thus, the faster the rising speed, the larger the MMN amplitude. We hence hypothesized a positive correlation between MMN amplitude and upside slope. However, once the MMN wave reached its peak, neural attenuation processes were expected independent of the peak amplitude reflecting an exponential decay. Thus, we hypothesized no correlation between MMN amplitude and downside slope.

Taken together, our present study took a new perspective to gain better understanding of the neural processing underlying auditory change detection. We investigated systematically the relationships between different MMN parameters with certain predictions on the relationships between major parameters of our concern.

## MATERIALS AND METHODS

### Participants

Twenty four right-handed Peking University students participated in this study; they received considerable financial reward afterward according to local standards. All participants passed an auditory test to guarantee normal hearing. They reported no neurological or psychiatric problems and had normal or corrected-to-normal vision. All the participants were informed that their brain activities would be recorded during the experiment, but they were naïve with respect to the real purpose of the study. The study was approved by the departmental ethics committee of Peking University, and all participants signed an informed consent before the experiment. In the data analysis, two participants were excluded from further analysis because of severe EEG artifacts; thus, 22 participants (11 females) remained for the final analyses. The mean age of the participants was 24.6 years (range 18–28 years).

### Stimuli and Procedure

Data were taken from a previous study, thus, the materials and methods, except the ERP data analysis, were identical to the study of Wang et al. (2015). A passive auditory oddball paradigm was

employed in which a sinusoidal tone of 1000 Hz served as the standard stimulus and another sinusoidal tone of 1500 Hz served as the deviant stimulus. The occurrence probability of the deviant was 20%. Tones were of 100 ms duration and had an intensity of 65–75 dB when measured by a decibel meter at the ear locations. Video clips of a silent documentary movie about the Indian River Ganges served as task relevant stimuli to which participants were asked to pay attention during the experiment.

Auditory stimuli were presented with a constant ISI within each of eight blocks. Altogether four different ISIs (1.5, 3, 4.5, and 6 s) were used. They were assigned to the first four blocks in a Latin square order and to the remaining four blocks in a reversed order. Each block contained 150 auditory stimuli with 120 standard tones and 30 deviant tones. Deviant tones were separated from each other at least by two standard tones to avoid a potential decrease of the oddball effect. Furthermore, in each block the first five stimuli were standard tones, in order to set a baseline for the participants. During the experiment each participant sat in a comfortable armchair in a dimly lit and electrically shielded room. Participants were told to continuously watch the video clips while ignoring the auditory stimuli presented from a speaker 40 cm behind them. They were asked to pay full attention to the subtitles and video images, and were told that they would be tested during the breaks. About every 15 min participants were encouraged to take a break and they were quizzed by the experimenters on the movie contents. All participants memorized the movie content remarkably well. During the recordings, participants were asked to restrain from frequent eye blinks and head movements to avoid EEG artifacts. All experiments were done from 9 to 12 am to avoid potential circadian fluctuations (Pöppel and Bao, 2014; Zhou et al., 2014; Bao et al., 2015). On average, one recording lasted for about 90 min. After the experiment, participants were debriefed about the purpose of the experiment.

### Electrophysiological Recordings

Electroencephalographic (EEG) data were recorded with a 64-channel NeuroScan 4.3 system. The electrodes positions were

chosen according to the extended 10–20 system. The average of the bilateral mastoids served as the online reference and the forehead served as ground. Vertical eye movements were monitored with bipolar electrodes above and below the left eye; horizontal eye movements were monitored with electrodes placed at the bilateral temples. The recording sampling rate was 500 Hz, high-pass filtered at 0.05 Hz and low-pass filtered at 100 Hz. In every recording, the impedance of each electrode was below 5 k $\Omega$ .

## EEG Data Preprocessing

Fieldtrip Toolbox (Oostenveld et al., 2011) was used for offline pre-processing of EEG raw data. Raw data were first band-pass filtered between 1 and 100 Hz and then an independent component analysis (ICA, Belouchrani et al., 1993) was utilized to remove eye and muscle artifacts. We defined each epoch from 400 ms before to 800 ms after the onset of each stimulus. The ICA-processed data were corrected to the baseline of –200 ms to 0 ms and then low-pass filtered at 25 Hz. Deviant and standard epochs for each condition were separately averaged to obtain the waveforms. For each participant, standard waves were subtracted from deviant waves to obtain the MMN.

## ERP Parameters Extraction

In a first step, altogether nine categories were extracted from the MMN waves, i.e., amplitude, averaged amplitude, onset latency, offset latency, duration, area, peak latency, upside slope, and downside slope (named as upslope and downslope thereafter). Descriptions of each parameter and their defining criteria are listed in **Table 1** and shown in **Figure 1**. It should be mentioned that a threshold (Th) value was calculated in each MMN wave by the formula  $Th = Amp_{base} - Std_{base}$ , where  $Amp_{base}$  is the mean amplitude of the baseline epoch (from –200 ms to 0 ms) and the  $Std_{base}$  is the corresponding standard deviation. MMN waves without any amplitude more negative than Th were considered as “fake-MMNs” and discarded from further analysis.

## Correlation Analysis

Test of normality (Kolmogorov–Smirnov test) of parameters showed that some of them violated the normal distribution ( $p > 0.05$ ), and therefore Spearman Correlations ( $n = 22$ )

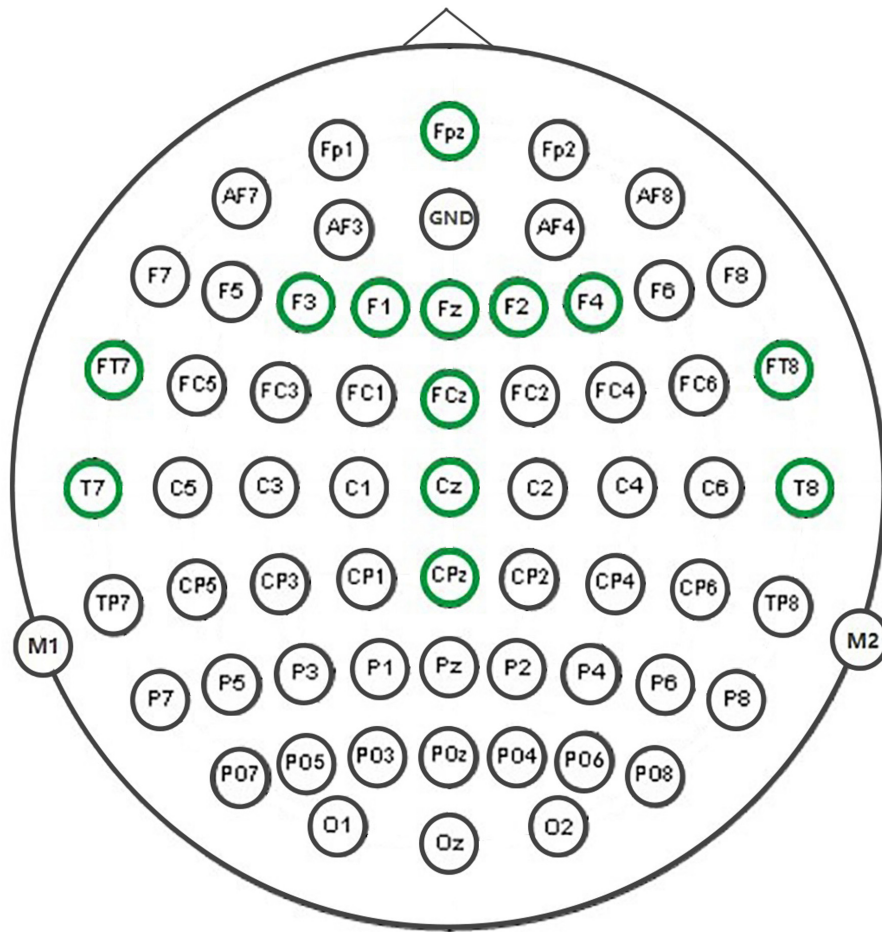
coefficients were computed between all pairs of parameters for each of MMN waves, i.e., the waves for the four ISI conditions. Combinations of two out of nine parameters altogether result in  $C_9^2 = 36$  pairs. Consequently, we have  $36 \times 4 = 144$  correlation coefficients and corresponding  $p$ -values at each electrode. Thirteen electrodes were chosen for this correlation analysis, namely FPZ, FZ, FCZ, CZ, CPZ, F1, F3, F2, F4, FT7, FT8, T7, and T8, shown in **Figure 2**. These electrodes mainly are located over centro-frontal and temporal areas, and are often chosen as target electrodes for auditory MMN analysis (for a review see Näätänen et al., 2007). To this end, we first counted the number of significant and marginally significant correlations across electrodes for each ISI condition and each paired parameter. This approach does not simply test whether or not there is significant correlation among parameters, rather it calculates approximate indices for the relative reliability of the relationship between parameters based on the statistical tests, thus, reducing though not eliminating the potential impact of Type I error in drawing conclusions. More generally speaking, if one pair of parameters showed significant correlations at most of selected electrodes while another pair barely showed any, we would have decent confidence that the former (vs. the latter) revealed a relatively reliable relationship between the parameters. In this sense, although the temporal electrodes are close to the mastoids and exhibited smaller MMNs, the inclusion of them would not significantly alter the results and conclusions described below. Furthermore, ERP signals from adjacent electrodes might correlate with each other and our simple counting approach might conflate the strength of between-parameters relationship in some situations. Therefore, we conducted a further region of interest (ROI) analysis to first obtain averaged MMN waves within each of predefined regions and then calculate the parameters from the averaged MMNs for later correlation analysis. Altogether four regions were defined, i.e., frontal (FZ, FPZ, F1, F2, F3, and F4), central (FCZ, CZ, and CPZ), left temporal (T7 and FT7), and right temporal (T8 and FT8) regions.

## RESULTS

On the basis of our analytical procedure we obtained a large number of correlation coefficients. To draw clear patterns from

**TABLE 1 | Parameters and their measuring criteria.**

Parameter	Measuring criteria
Amplitude	Amplitude (in microvolts) of the most negative deflection (peak) between 100 and 300 ms after onset; for correlation analysis, its absolute value is used
Ampavg	Averaged value of $\pm 20$ ms around the peak; for correlation analysis, its absolute value is used
Onset	Time at the first point the MMN wave crossed the threshold
Offset	Time at the second point the MMN wave crossed the threshold
Peak latency	Time at the peak amplitude
Area	Accumulated area between onset and offset times
Duration	Offset minus onset time
Upslope	Largest derivative of the ascending side of the MMN wave; for correlation analysis, its absolute value is used
Downslope	Largest derivative of the descending side of the MMN wave; for correlation analysis, its absolute value is used



**FIGURE 2 |** Thirteen marked electrodes were used for the correlational analysis.

these correlations, we summed significant ( $p < 0.05$ ; two-tailed and thereafter) or marginally significant ( $0.05 \leq p < 0.1$ ) correlations out of the 13 electrodes (**Figure 2**) for each condition (**Table 2**). For example, if seven electrodes out of 13 show significant or marginally significant correlations between two parameters, the corresponding number in **Table 2** would be 7. By counting the number of significant correlations across electrodes, we could obtain a first impression of the relative strength of the relationship between any two parameters.

As shown in **Table 2**, large amounts of correlations were found for several expected correlational pairs across electrodes. The first group of these pairs is due to their similarity in nature, and the typical example is the Ampavg-Amplitude which showed average  $r > 0.9$  at all 13 electrodes. The second group is due to their mathematical relationship, examples including Amplitude-Area, Ampavg-Area, Onset-Duration, Offset-Duration, Onset-Area, and Offset-Area; within these pairs, the value of one parameter is dependent on the other parameter. Third, the large amounts of correlations for Amplitude-Duration, Ampavg-Duration, Peak

Latency-Onset, and Peak Latency-Offset are expected in a commonsense way. For the other combinations, a clear trend was that temporal parameters (onset, offset, peak latency) and shape parameters (amplitude, ampavg, upslope and downslope) were usually not correlated as judged from the numbers across electrodes, consistent with the findings from Lang et al. (1995).

Surprisingly, we found an unexpected asymmetry in correlational relationships between amplitude and upslope/downslope, which contradict our hypotheses. While rather high positive correlations in Amplitude-Downslope were observed (on average in more than 8 among 13 electrodes), the correlational relationships between Amplitude and Upslope barely exist (on average less than 1). A similar asymmetry was observed between Ampavg-Downslope (average number of correlated electrodes  $> 6$ ) and Ampavg-Upslope (average number = 0). These correlations correspond to the low correlations between upslope and downslope (average number  $< 2$ ). Furthermore, no systematic differences between the four ISI conditions were suspected from the counts in **Table 2**. To visually illustrate the results, in **Figure 3** for all electrodes



**TABLE 2 | Counts of correlations across electrodes.**

Pairs	1.5 s	3 s	4.5 s	6 s	Pairs	1.5 s	3 s	4.5 s	6 s
Ampavg-Amplitude	13 (13)	13 (13)	13 (13)	13 (13)	Upslope-Downslope	0	1 (1)	5 (2)	2 (2)
Ampavg-Upslope	1 (0)	1 (0)	0	1 (0)	Upslope-Duration	1 (1)	0	1 (0)	0
Ampavg-Peak latency	0	1 (0)	2 (0)	0	Upslope-Area	0	2 (0)	0	0
Ampavg-Onset	12 (9)	7 (4)	4 (1)	5 (3)	Peak latency-Onset	11 (10)	13 (13)	13 (13)	13 (13)
Ampavg-Offset	4 (2)	3 (2)	11 (7)	2 (2)	Peak latency-Offset	13 (13)	13 (13)	13 (13)	13 (13)
Ampavg-Downslope	8 (7)	7 (5)	7 (5)	8 (8)	Peak latency-Downslope	0	0	1 (0)	1 (0)
Ampavg-Duration	13 (13)	13 (13)	13 (13)	13 (12)	Peak latency-Duration	0	1 (0)	1 (0)	0
Ampavg-Area	13 (13)	13 (13)	13 (13)	13 (13)	Peak latency-Area	0	0	0	0
Amplitude-Upslope	0	1 (0)	1 (0)	0	Onset-Offset	5 (4)	10 (8)	9 (6)	13 (12)
Amplitude-Peak latency	0	0	0	0	Onset-Downslope	1 (1)	4 (2)	0	4 (1)
Amplitude-Onset	12 (10)	8 (5)	5 (3)	7 (5)	Onset-Duration	13 (13)	8 (6)	8 (8)	4 (3)
Amplitude-Offset	3 (2)	4 (2)	6 (6)	3 (3)	Onset-Area	13 (12)	8 (7)	7 (5)	6 (3)
Amplitude-Downslope	9 (9)	10 (9)	10 (7)	10 (10)	Offset-Downslope	0	1 (0)	0	0
Amplitude-Duration	13 (13)	13 (13)	13 (13)	13 (13)	Offset-Duration	9 (8)	11 (8)	12 (11)	11 (7)
Amplitude-Area	13 (13)	13 (13)	13 (13)	13 (13)	Offset-Area	8 (6)	9 (5)	12 (12)	6 (5)
Upslope-Peak latency	1 (0)	1 (0)	2 (2)	1 (1)	Downslope-Duration	1 (1)	0	2 (1)	3 (0)
Upslope-Onset	0	1 (0)	0	1 (1)	Downslope-Area	5 (3)	6 (3)	7 (4)	8 (6)
Upslope-Offset	1 (1)	2 (0)	2 (1)	1 (0)	Duration-Area	13 (13)	13 (13)	13 (13)	13 (13)

Unparenthesized values,  $p < 0.1$ ; parenthesized values,  $p < 0.05$ .

and all ISIs the distribution of the two slopes (upslope and downslope) and the corresponding amplitudes are presented; a clear difference is seen between the Amplitude-Upslope and the Amplitude-Downslope relations.

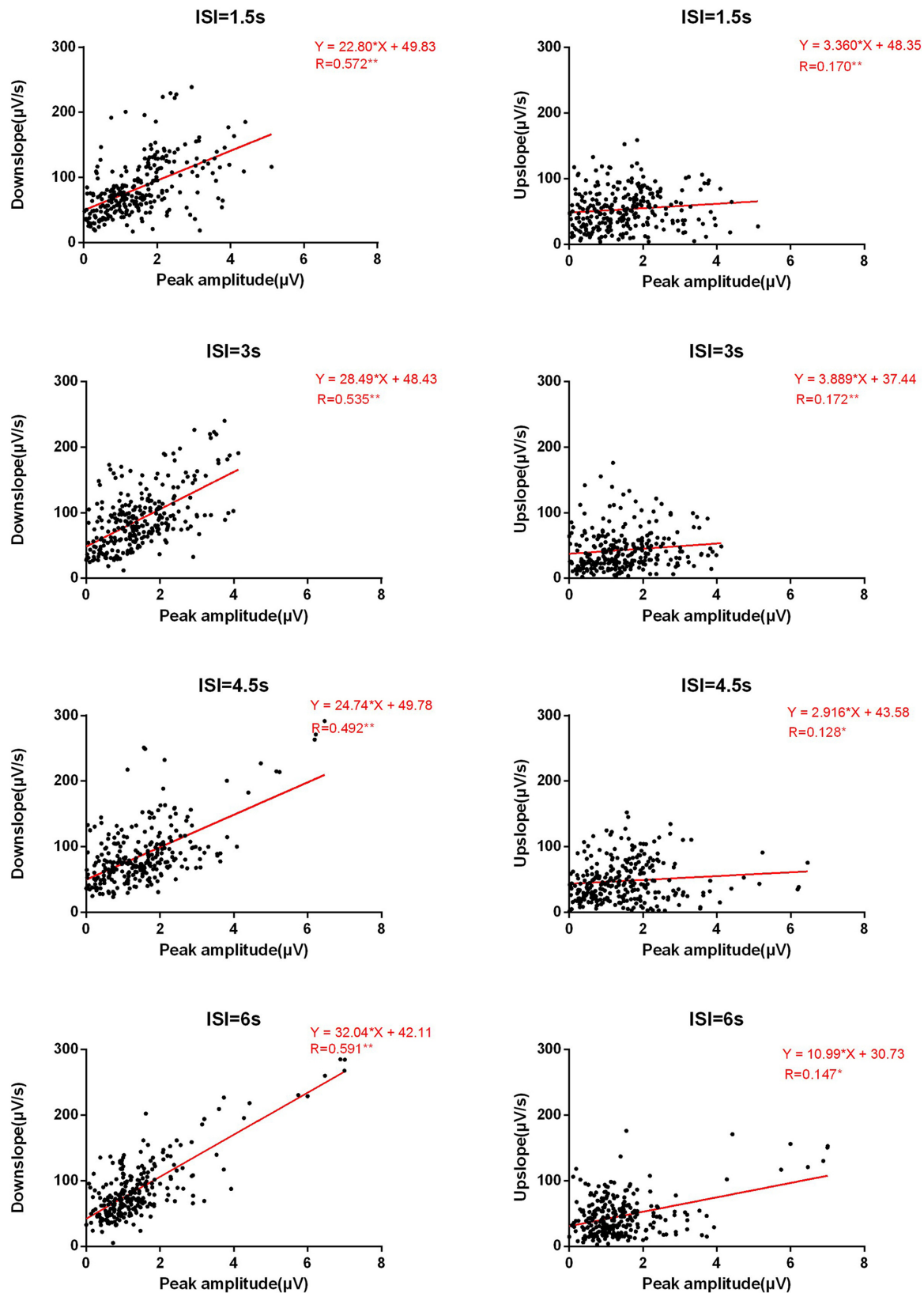
Considering that ERP signals at nearby electrodes might correlate with each other, in a further step, we clustered the selected electrodes into four relatively homogenous regions and performed ROI analyses (see Materials and Methods). On the basis of the results in **Table 2**, we focused our ROI analyses on the correlations between amplitudes and slopes. The results are shown in **Table 3**. Consistent with the results of the counting analysis, **Table 3** shows significant correlations between Amplitude and Downslope but none between Amplitude and Upslope; the correlations between Ampavg with Downslope and Upslope showed equivalent results. Interestingly, the asymmetrical effects between Downslope and Upslope were more pronounced over the frontal and central scalp relative to the bilateral temporal areas.

## DISCUSSION

Our study explored the potential correlations between parameters defining the time course and the shape of MMN wave. The MMNs were elicited using a conventional auditory oddball paradigm and responses in a group of 13 electrodes were analyzed. Consistent with our prediction, the MMN peak latency was not correlated with the amplitude (including ampavg), and no correlation was found for peak latency and MMN slope (upslope/downslope) either. In fact, our results showed that most combinations between temporal parameters (onset, offset, and peak latency) and wave shape parameters (amplitude, ampavg, upslope and downslope) were not correlated with each other, indicating that the MMN wave

shape characteristics are irrelevant with respect to when the MMN is elicited.

Most importantly, we found an unexpected asymmetry regarding the relationship between MMN amplitude and the two types of slopes: the MMN downslope was positively correlated with the MMN amplitude, while the upslope was not correlated with the amplitude. This observation showed an opposite pattern to our hypothetical expectation and revealed an important new phenomenon: the downside decreasing speed of the MMN wave increases with the amplitude. To the best of our knowledge, it is the first time to observe such an asymmetric correlation between MMN amplitude and upslope vs. downslope. The observation that the MMN amplitude is positively correlated with the downslope and not with the upslope is very surprising, since an exponential decay as represented in many passively decreasing biological processes would be expected, i.e., once the MMN wave reaches its peak, the same neural attenuation processes should occur independent of the peak amplitude. The unexpected observation disproves our hypothesis and in our view suggests a different type of neural attenuation. Once the neural response reaches its peak, our brain uses a negative feedback mechanism to actively draw the neural activity back to the optimal operation level which can be referred to as the baseline; the more the deviation from the baseline, the stronger the return force to the baseline within a certain time (see **Figure 4**). This implies that apparently our neural system has a kind of “distance” information available (i.e., how far it is away from the baseline) at the time when the peak amplitude is reached, and can further use this information to actively control the returning process; otherwise, one could not explain that the returning speed is positively correlated with the distance from the baseline. This further suggests an anticipatory temporal control mechanism for the system to be prepared as fast as possible for the next novelty detection.

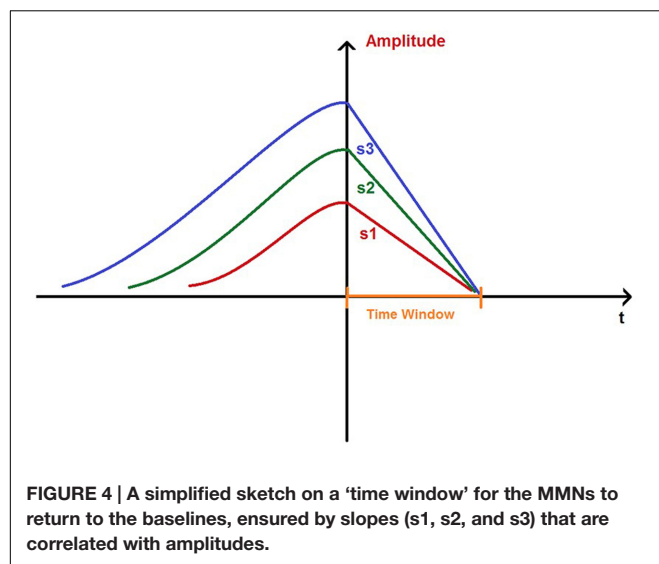


**FIGURE 3 | Basic linear fitting of amplitudes and slopes.** The Y-axis represents the downslope (left panels) or the upslope (right panels) for four ISI conditions in absolute values. The X-axis represents the absolute values of peak amplitude. Each figure has the data points from all electrodes and participants.

**TABLE 3 | Correlations between amplitude and slopes across different regions.**

Areas	Correlation Pairs	1.5 s	3 s	4.5 s	6 s
Frontal	Amplitude-Upslope	0.1553	0.3258	−0.0932	0.1218
	Amplitude-Downslope	0.6318**	0.6533**	0.7064**	0.5504*
Central	Amplitude-Upslope	0.1273	0.2532	0.1485	0.1182
	Amplitude-downslope	0.6805**	0.2519	0.5675**	0.6701**
Left temporal	Amplitude-Upslope	0.3209	0.4180	−0.0597	−0.0120
	Amplitude-downslope	0.6017**	0.5129*	0.1779	0.0647
Right temporal	Amplitude-Upslope	−0.0536	0.0526	−0.3729	0.1818
	Amplitude-downslope	0.2208	−0.0180	0.3383	0.1195

The table presents *R*-values of the correlations. \* $p < 0.05$ , \*\* $p < 0.01$ .



Since the decreasing speed of MMN component is dependent on its deviation from the baseline, we suggest the name 'rubberband effect' as a metaphor to describe this mechanism. It is important to point out, that the effects were clearly observable in all ISIs which in a previous study (Wang et al., 2015) showed different MMN amplitudes. Thus, it seems that the association between the MMN amplitude and downslope might reflect an intrinsic neural mechanism dealing with the disequilibrium, which might be independent of the temporal segmentation (Pöppel, 1997, 2009) that modulates the MMN amplitude.

However, one has to be cautious about the generalization of this independence of the temporal segmentation. The dominant effect derived from the pattern of correlation coefficients (see **Table 3**) is observed in frontal midline areas extending back to central areas. Most importantly, however, we observed a clear hemispheric asymmetry. The right hemisphere did not show any involvement for a fast return to baseline; thus, there was no rubberband effect. However, in the left hemisphere the rubberband effect was observed for 1.5 and 3 s, but not for the longer ISIs. This observation suggests that in addition to the spatial distribution of the rubberband effect, there is also a temporal modulation of the effect. We speculate that the left

hemisphere involvement of the effect and its temporal limitation to approximately 3 s could be related to the 3 s platform observed in verbal behavior; spontaneous speech is segmented in successive 3 s time windows (Vollrath et al., 1992). Thus, the brain is prepared for the next utterance within regular time intervals.

Contrary to our hypothesis as well, the MMN upslope was not correlated with the amplitude. This observation indicates that the generating (up) and diminishing (down) processes of MMN wave are implemented by different neural mechanisms. As indicated by some investigators, the rising phase may reflect the neuronal activation (Korostenskaja et al., 2003) or summation process (Zhou et al., 2010) during the encoding stage of information processing. The present study shows that this activation or summation process seems to be an integration process, which does not anticipate the position of the MMN peak. In other words, the MMN peak can be reached at any time during the summation process, thus resulting in no correlation between the upslope and the amplitude. The proposition that the MMN generating process does not anticipate the position of the peak is further supported by the observation of no correlation between the MMN amplitude and the peak latency. Thus, different from the rubberband-like active operation in the returning phase of the MMN, a passive information accumulation in the MMN generating process is indicated.

Besides our major surprising observation of an asymmetry of neural processes before and after the peak of the amplitude, we also observed that the MMN peak latency was not correlated with the MMN amplitude and the slopes, which is consistent with our hypotheses and also in line with observations by Lang et al. (1995). However, it should be noted that a previous study by Polich et al. (1997) did observe correlational relationships between amplitude and latency, but with the P3 ERP component. Since P3 is closely related to the operation of attention system (Polich, 2007) and the MMN in our case is observed with a passive oddball paradigm, our conclusion regarding the relationship between amplitude and latency should be limited to auditory frequency MMN in a passive oddball paradigm only, since the involvement of attention or the change of the oddball probability could possibly change the correlational relationships.

Finally, the current study also pointed out a simple yet effective data analyzing method by using correlational relationship as a way to explore psychological or neuronal mechanism in ERP signals. In previous studies, there have been discussions on advantages and disadvantages of single measurements like peak amplitude, peak latency or mean amplitude (Luck, 2014). Sometimes it may be difficult which parameter to use, as some might show hypothesized results while others do not. The method we suggest enables us to explore all parameters from a different perspective, and statistically the correlational analysis has a higher tolerance with respect to data variation, and it remains robust with different calculating methods. These advantages might prove to be useful in future studies.

## REFERENCES

- Amenedo, E., and Escera, C. (2000). The accuracy of sound duration representation in the human brain determines the accuracy of behavioral perception. *Eur. J. Neurosci.* 12, 2570–2574. doi: 10.1046/j.1460-9568.2000.00114.x
- Baldeweg, T., Richardson, A., Watkins, S., Foale, C., and Gruzeliér, J. (1999). Impaired auditory frequency discrimination in dyslexia detected with mismatch evoked potentials. *Ann. Neurol.* 45, 495–503.
- Bao, Y., Pöppel, E., Wang, L., Lin, X., Yang, T., Avram, M., et al. (2015). Synchronization as a biological, psychological and social mechanism to create common time: a theoretical frame and a single case study. *Psych. J.* 4, 243–254. doi: 10.1002/pchj.119
- Belouchrani, A., Abed-Meraim, K., Cardoso, J. F., and Moulines, E. (1993). “Second-order blind separation of temporally correlated sources,” in *Proceedings of the International Conference on Digital Signal Processing*, Cyprus, 346–351.
- Cheour, M., Martynova, O., Näätänen, R., Erkkola, R., Sillanpää, M., Kero, P., et al. (2002). Speech sounds learned by sleeping newborns. *Nature* 415, 599–600. doi: 10.1038/415599b
- Haenschel, C., Vernon, D., Dwivedi, P., Gruzeliér, J., and Baldeweg, T. (2005). Event-related brain potential correlates of human auditory sensory memory-trace formation. *J. Neurosci.* 25, 10494–10501. doi: 10.1523/JNEUROSCI.1227-05.2005
- Horváth, J., Czigler, I., Jacobsen, T., Maess, B., Schröger, E., and Winkler, I. (2008). MMN or no MMN: no magnitude of deviance effect on the MMN amplitude. *Psychophysiology* 45, 60–69. doi: 10.1111/j.1469-8986.2007.00599.x
- Intriligator, J., and Polich, J. (1994). On the relationship between background EEG and the P300 event-related potential. *Biol. Psychol.* 37, 207–218. doi: 10.1016/0301-0511(94)90003-5
- Jaramillo, M., Paavilainen, P., and Näätänen, R. (2000). Mismatch negativity and behavioural discrimination in humans as a function of the magnitude of change in sound duration. *Neurosci. Lett.* 290, 101–104. doi: 10.1016/S0304-3940(00)01344-6
- Kargel, C., Sartory, G., Kariofillis, D., Wiltfang, J., and Muller, B. (2014). Mismatch negativity latency and cognitive function in schizophrenia. *PLoS ONE* 9:e84536. doi: 10.1371/journal.pone.0084536
- Korostenskaja, M., Dapsys, K., Maciulis, V., and Ruksenas, O. (2003). Evaluation of new MMN parameters in schizophrenia. *Acta Neurobiol. Exp. (Wars)* 63, 383–388.
- Krauel, K., Schott, P., Sojka, B., Pause, B. M., and Ferstl, R. (1999). Is there a mismatch negativity analogue in the olfactory event-related potential? *J. Psychophysiol.* 13, 49–55. doi: 10.1027//0269-8803.13.1.49
- Lang, A. H., Eerola, O., Korpilahti, P., Holopainen, I., Salo, S., and Aaltonen, O. (1995). Practical issues in the clinical application of mismatch negativity. *Ear Hear.* 16, 118–130. doi: 10.1097/00003446-199502000-00009
- Lang, H. A., Nyrke, T., Ek, M., Aaltonen, O., Raimo, I., and Näätänen, R. (1990). “Pitch discrimination performance and auditory event-related potentials” in *Psychophysiological Brain Research*, Vol. 1, eds C. H. M. Brunia, A. W. K.

## AUTHOR CONTRIBUTIONS

LW, EP, and YB designed the experiment. LW, XL, and BZ collected and analyzed the data. LW, BZ, EP, and YB wrote the manuscript.

## ACKNOWLEDGMENTS

This work was supported by grants from the National Natural Science Foundation of China (No. 31371018, 91120004, 31100735 and J1103602) and the Chinese Academy of Sciences (CAS Visiting Professorships for Senior International Scientists, 2013T1S0029).

- Gaillard, A. Kok, G. Mulder, and M. N. Verbaten (Tilburg: Tilburg University Press), 294–298.
- Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. Cambridge: MIT Press.
- Martynova, O., Kirjavainen, J., and Cheour, M. (2003). Mismatch negativity and late discriminative negativity in sleeping human newborns. *Neurosci. Lett.* 340, 75–78. doi: 10.1016/S0304-3940(02)01401-5
- Näätänen, R. (1992). *Attention and Brain Function*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Näätänen, R., Gaillard, A. W. K., and Mäntysalo, S. (1978). Early selective-attention effect on evoked potential reinterpreted. *Acta Psychol.* 42, 313–329. doi: 10.1016/0001-6918(78)90006-9
- Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clin. Neurophysiol.* 118, 2544–2590. doi: 10.1016/j.clinph.2007.04.026
- Novitski, N., Tervaniemi, M., Huottilainen, M., and Näätänen, R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cogn. Brain Res.* 20, 26–36. doi: 10.1016/j.cogbrainres.2003.12.011
- 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. doi: 10.1155/2011/156869
- Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: a review. *Int. J. Psychophysiol.* 88, 109–123. doi: 10.1016/j.ijpsycho.2013.03.015
- Pakarinen, S., Takegata, R., Rinne, T., Huottilainen, M., and Näätänen, R. (2007). Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). *Clin. Neurophysiol.* 118, 177–185. doi: 10.1016/j.clinph.2006.09.001
- Pazo-Alvarez, P., Cadaveira, F., and Amenedo, E. (2003). MMN in the visual modality: a review. *Biol. Psychol.* 63, 199–236. doi: 10.1016/S0301-0511(03)00049-8
- Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clin. Neurophysiol.* 118, 2128–2148. doi: 10.1016/j.clinph.2007.04.019
- Polich, J., Alexander, J. E., Bauer, L. O., Kuperman, S., Morzorati, S., O'Connor, S. J., et al. (1997). P300 topography of amplitude/latency correlations. *Brain Topogr.* 9, 275–282. doi: 10.1007/BF01464482
- Pöppel, E. (1997). A hierarchical model of temporal perception. *Trends Cogn. Sci.* 1, 56–61. doi: 10.1016/S1364-6613(97)01008-5
- Pöppel, E. (2009). Pre-semantically defined temporal windows for cognitive processing. *Philos. Trans. R. Soc. B* 364, 1887–1896. doi: 10.1098/rstb.2009.0015
- Pöppel, E., and Bao, Y. (2014). “Temporal windows as bridge from objective time to subjective time,” in *Subjective Time: The Philosophy, Psychology, and Neuroscience of Temporality*, eds D. Lloyd and V. Arstila (Cambridge: MIT Press), 241–261.
- Roberts, T. P. L., Cannon, K. M., Tavabi, K., Blaskey, L., Khan, S. Y., Monroe, J. F., et al. (2011). Auditory magnetic mismatch field latency: a



- biomarker for language impairment in autism. *Biol. Psychiatry* 70, 262–269. doi: 10.1016/j.biopsych.2011.01.015
- Saarienen, J., Paavilainen, P., Schöger, E., Tervaniemi, M., and Näätänen, R. (1992). Representation of abstract attributes of auditory stimuli in the human brain. *Neuroreport* 3, 1149–1151. doi: 10.1097/00001756-199212000-00030
- Schröger, E., Bendixen, A., Trujillo-Barreto, N. J., and Roeber, U. (2007). Processing of abstract rule violations in audition. *PLoS ONE* 2:e1131. doi: 10.1371/journal.pone.0001131
- Sinkkonen, J., and Tervaniemi, M. (2000). Towards optimal recording and analysis of the mismatch negativity. *Audiol. Neurotol.* 5, 235–246. doi: 10.1159/000013885
- Vollrath, M., Kazenwadel, J., and Krüger, H.-P. (1992). A universal constant in temporal segmentation of human speech. A reply to Schleidt and Feldhütter (1989). *Naturwissenschaften* 79, 479–480. doi: 10.1007/BF01139205
- Wang, L., Lin, X., Zhou, B., Pöppel, E., and Bao, Y. (2015). Subjective present: a window of temporal integration indexed by mismatch negativity. *Cogn. Process.* 16, 131–135. doi: 10.1007/s10339-015-0687-8
- Zhou, B., Bao, Y., Sander, T., Trahms, L., and Pöppel, E. (2010). Dissociation of summation and peak latencies in visual processing: an MEG study on stimulus eccentricity. *Neurosci. Lett.* 483, 101–104. doi: 10.1016/j.neulet.2010.07.070
- Zhou, B., Pöppel, E., and Bao, Y. (2014). In the jungle of time: the concept of identity as an escape. *Front. Psychol.* 5:844. doi: 10.3389/fpsyg.2014.00844

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Wang, Lin, Zhou, Pöppel and Bao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# Detecting Temporal Change in Dynamic Sounds: On the Role of Stimulus Duration, Speed, and Emotion

Annett Schirmer<sup>1,2,3\*</sup>, Nicolas Escoffier<sup>1,2</sup>, Xiaoqin Cheng<sup>2,4</sup>, Yenju Feng<sup>2,4</sup> and Trevor B. Penney<sup>1,2\*</sup>

<sup>1</sup> Department of Psychology, National University of Singapore, Singapore, Singapore, <sup>2</sup> Life Sciences Institute Programme in Neurobiology and Ageing, National University of Singapore, Singapore, Singapore, <sup>3</sup> Duke-NUS Graduate Medical School, Singapore, Singapore, <sup>4</sup> Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore, Singapore

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College, USA

### Reviewed by:

William Matthews,  
University of Cambridge, UK  
Raquel Cocenas Silva,  
University of São Paulo, Brazil

### \*Correspondence:

Annett Schirmer  
schirmer@nus.edu.sg;  
Trevor B. Penney  
penney@nus.edu.sg

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 29 October 2015

**Accepted:** 24 December 2015

**Published:** 13 January 2016

### Citation:

Schirmer A, Escoffier N, Cheng X,  
Feng Y and Penney TB (2016)  
Detecting Temporal Change  
in Dynamic Sounds: On the Role  
of Stimulus Duration, Speed,  
and Emotion. *Front. Psychol.* 6:2055.  
doi: 10.3389/fpsyg.2015.02055

For dynamic sounds, such as vocal expressions, duration often varies alongside speed. Compared to longer sounds, shorter sounds unfold more quickly. Here, we asked whether listeners implicitly use this confound when representing temporal regularities in their environment. In addition, we explored the role of emotions in this process. Using a mismatch negativity (MMN) paradigm, we asked participants to watch a silent movie while passively listening to a stream of task-irrelevant sounds. In Experiment 1, one surprised and one neutral vocalization were compressed and stretched to create stimuli of 378 and 600 ms duration. Stimuli were presented in four blocks, two of which used surprised and two of which used neutral expressions. In one surprised and one neutral block, short and long stimuli served as standards and deviants, respectively. In the other two blocks, the assignment of standards and deviants was reversed. We observed a climbing MMN-like negativity shortly after deviant onset, which suggests that listeners implicitly track sound speed and detect speed changes. Additionally, this MMN-like effect emerged earlier and was larger for long than short deviants, suggesting greater sensitivity to duration increments or slowing down than to decrements or speeding up. Last, deviance detection was facilitated in surprised relative to neutral blocks, indicating that emotion enhances temporal processing. Experiment 2 was comparable to Experiment 1 with the exception that sounds were spectrally rotated to remove vocal emotional content. This abolished the emotional processing benefit, but preserved the other effects. Together, these results provide insights into listener sensitivity to sound speed and raise the possibility that speed biases duration judgements implicitly in a feed-forward manner. Moreover, this bias may be amplified for duration increments relative to decrements and within an emotional relative to a neutral stimulus context.

**Keywords:** auditory change detection, event-related potentials, vocal affect, sex differences, interval timing, preattentive, prosody

## INTRODUCTION

The human temporal sense depends on the ability to represent external events marking the passage of time. Research has shown that individuals encode such events outside awareness and automatically detect changes in event onset and duration (e.g., Näätänen et al., 1989; Tse and Penney, 2006). We asked whether individuals likewise track the speed with which events unfold and whether emotions benefit such tracking. Specifically, we explored brain responses to unattended neutral and emotional sounds that occasionally accelerated and decelerated, becoming shorter and longer as a consequence.

### Time Perception: On the Role of Stimulus Properties and Context

Time perception, a sixth human sense, critically contributes to meaningful interactions with the environment. Many mental functions depend on time. For example, attention is thought to be governed by temporal rhythms (Jones, 1976; Jones and Boltz, 1989; Escoffier et al., 2015). Learning in the context of classical and operant conditioning is shaped by the delay between two stimuli or between a behavior and its consequence (Gallistel and Gibbon, 2000). Language is temporally sensitive because the comprehension of words and syntactic dependencies relies on durational parameters such as voice-onset-time or intonation (Schirmer, 2004). Additionally, in the context of non-verbal communication, timing tells us how long to hold another's gaze, to laugh at another's joke, and to wait for another's response before responding in turn.

Traditionally, the study of time perception has relied on simple static stimuli such as tones or images for which participants compared stimulus duration to a reference duration in memory. More recently, researchers have explored time perception with dynamic stimuli such as moving objects (Kaneko and Murakami, 2009; Matthews, 2011; Su and Jonikaitis, 2011; Linares and Gorea, 2015), tones (Matthews, 2013), music (Firmino et al., 2009; Droit-Volet et al., 2010; Cocenas-Silva et al., 2011; Darlow et al., 2013), faces (Fayolle and Droit-Volet, 2014), and vocalizations (Schirmer et al., in press). Although ecologically more valid, this approach presents a methodological challenge. With dynamic stimuli, duration always varies in conjunction with one or two other factors: stimulus content and speed. For example, compared to shorter vocalizations, longer vocalizations may contain more ups and downs in pitch (i.e., content varies), and/or the same pitch variation may play out more slowly (i.e., speed varies). At present, our understanding of how these two natural confounds impact time perception is incomplete.

Apart from using more ecologically valid materials, recent timing research has elucidated the role of contextual variables, such as emotion. For example, when asked to time the duration of a sound or image, temporal reproductions or duration judgements are typically longer for emotional as compared to neutral stimuli (e.g., Grommet et al., 2011; for reviews see Droit-Volet and Meck, 2007; Schirmer, 2011). Emotion effects differ depending on whether stimuli are static or dynamic and whether duration manipulations affect stimulus content as opposed to

speed. Static timing stimuli tend to appear longer when they are emotional as compared to neutral (e.g., Grommet et al., 2011). The same is true for dynamic stimuli when duration manipulations are confounded by stimulus content, that is longer stimuli entail more events than do shorter stimuli (Angrilli et al., 1997). However, opposite effects emerge for dynamic stimuli that are confounded by speed, that is, longer stimuli play out more slowly than shorter stimuli (Voyer and Reuangrith, 2015; Schirmer et al., in press). Here, emotionality increases the probability of stimuli being perceived as short.

Although prior research has enhanced our understanding of the human temporal sense in the context of both static and dynamic events, this understanding is still incomplete. Moreover, the focus has been on explicit duration judgements and the representation of stimulus on- and offsets. Thus, we still know little about implicit timing and how individuals represent the temporal course of information between stimulus onset and offset. Additionally, it is unclear how such timing and temporal course representations are modulated by stimulus emotion.

### Mismatch Negativity: An Implicit Measure of Temporal and Emotional Processing

To address the issues outlined here, we adopted a mismatch negativity (MMN) paradigm. In this paradigm, participants pursue a primary activity while an auditory stimulus stream plays in the background. For example, they may watch a silent film while passively listening to a sequence of 440 Hz tones that is occasionally interrupted by a higher pitched deviant (Sams et al., 1985). The electroencephalogram (EEG), recorded throughout this procedure, reveals a negative event related potential (ERP) component peaking around 200 ms following stimulus onset with a fronto-central topography. This component, called the MMN, emerges when the average EEG response to standard stimuli is subtracted from the average EEG response to deviant stimuli. Over the mastoid electrodes, the MMN has positive polarity and this is thought to index generators in primary auditory cortex (for a recent review see Garrido et al., 2009).

The MMN has been used as a marker of implicit temporal processing in several studies (for a review see Ng and Penney, 2014). Stimuli comprised sequences of tones, vowels, syllables or music and stimuli within a sequence differed in duration only. For example, Jacobsen and Schröger (2003) presented several experimental blocks, one of which comprised frequent standard tones of 150 ms and rare deviant tones of 100 ms. Similar to others, these authors found that temporal deviants elicited an MMN peaking around 250 ms following sound onset. In case of shorter deviants, as was true for their block, temporal change occurs at deviant offset. In case of longer deviants, temporal change occurs during the deviant (i.e., at the value of the standard duration). Notably, there is evidence that MMN amplitude differs between these cases (but see Amenedo and Escera, 2000; Peter et al., 2012), although some studies found that long deviants produce a larger MMN than short deviants (e.g., Catts et al., 1995; Jaramillo et al., 1999; Atkinson et al., 2012), whereas other studies found the opposite (e.g., Jaramillo et al., 1999; Takegata et al.,

2008; Colin et al., 2009). Why change direction matters and why its influence varies are still open questions and will be considered further in the discussion of this paper.

Apart from being sensitive to temporal change, the MMN is also sensitive to emotional change. In a first study, Schirmer et al. (2005) presented listeners with the pseudoword “dada” spoken in a neutral and angry voice. In some blocks, neutral sounds served as standards and angry sounds as deviants, whereas in other blocks, angry sounds served as standards and neutral sounds as deviants. A separate experiment used neutral and happy voices. Both emotional sounds elicited an earlier and larger MMN than the neutral sounds, indicating that emotions facilitated change detection. Moreover, sex differences in this effect indicated that women were more sensitive than men to unattended emotional expressions. Subsequent research replicated these results (Schirmer et al., 2007; Schirmer and Escoffier, 2010; Thönnessen et al., 2010; Fan et al., 2013) and pointed to sources in the temporal lobe and insula (Schirmer et al., 2008; Thönnessen et al., 2010).

## The Present Study

Clearly, the MMN can be used to study implicit temporal processing. However, existing MMN studies focused on duration effects, that is whether a deviant expires before or after a standard, rather than on dynamic time course effects that emerge throughout a stimulus. Moreover, the role of emotion, which is known to affect both timing and the MMN, remains to be explored. Here, we sought to address these gaps. We created MMN stimuli by subjecting one surprised and one neutrally spoken “Ah” to a speech manipulation procedure creating a 378 ms and a 600 ms exemplar, of which the former had a fast and the latter a slow speed. We then presented these exemplars in four blocks, two of which comprised surprised and two of which comprised neutral stimuli. Within each emotion condition, one block used the short exemplar as the standard and the long exemplar as the deviant, whereas the other block had a reversed stimulus assignment. Participants passively listened to the four blocks, while watching a silent subtitled movie.

In line with previous research using static stimuli, we expected a duration MMN peaking about 200 ms after deviant offset (short deviant blocks) or the duration value of the standard (long deviant blocks). Additionally, we expected ERP differences between deviants and standards due to the dynamic nature of our stimuli. Specifically, if listeners implicitly process the slowing down and speeding up associated with long and short deviants, respectively, the deviant ERP should become more negative than the standard ERP prior to the stimulus duration mismatch. Moreover, we anticipated an incremental mismatch effect emerging shortly after stimulus onset as information about temporal change continuously accumulated and the temporal disparity between standard and deviant increased. Last, we hypothesized that emotions would facilitate temporal change detection resulting in an earlier and larger mismatch effect for the surprised relative to the neutral blocks, especially for female listeners.

## EXPERIMENT 1

### Methods

This research was approved by the Institutional Review Board of the National University of Singapore.

### Participants

We recruited 35 participants, most of whom were students, through campus advertisements. The data from three participants were discarded due to excessive movement artifacts in the EEG. Half of the remaining participants were male and the other half were female. The average age was 22.2 ( $SD = 2.4$ ). Participants reported normal hearing and normal or corrected to normal vision. They were reimbursed for their time at a rate of S\$10/hour.

### Stimuli

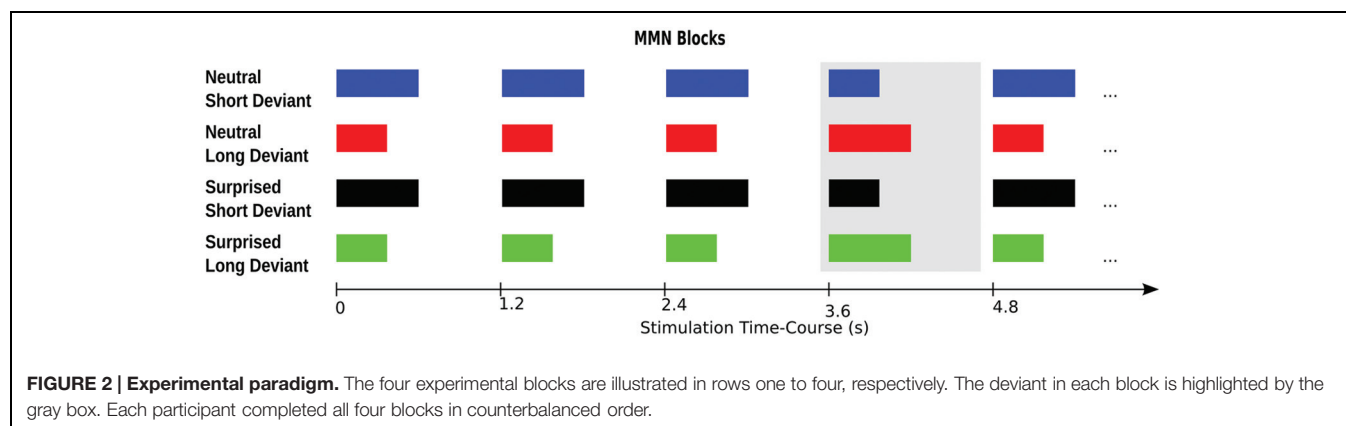
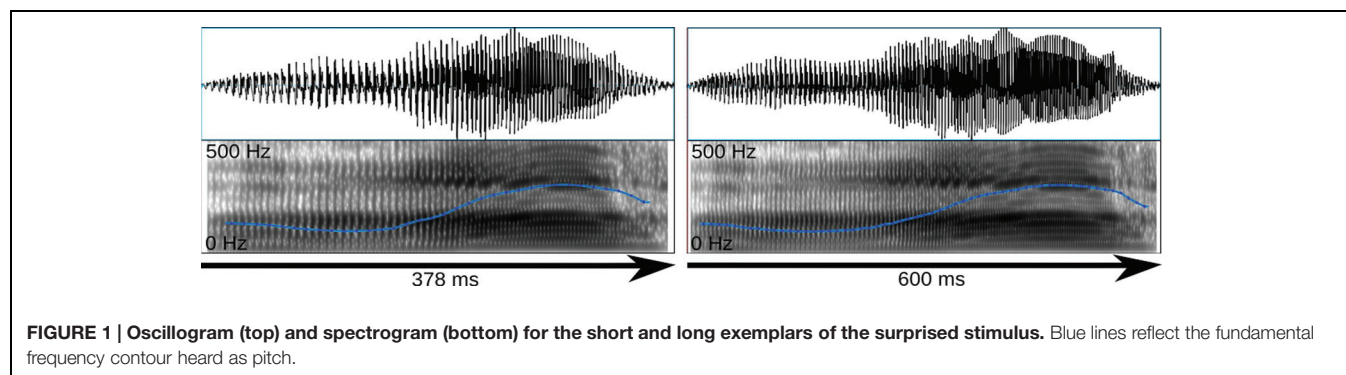
The stimulus material consisted of the interjection “Ah” spoken in a neutral and a surprised voice by a male speaker. The recordings were taken from a larger pool of 231 interjections produced by 33 speakers and including neutral, surprised, sad, angry, happy, disgusted, and fearful expressions. These stimuli were presented to 30 individuals (18 female and 12 male, mean age = 22.1,  $SD = 2.2$ ) who did not participate in the main experiment and who indicated whether a given stimulus expressed neutrality, surprise, sadness, anger, happiness, disgust, fear or another self-determined emotion and rated stimulus arousal on a 5-point scale ranging from 1 (very calm) to 5 (very excited). Recognition probability (i.e., number of raters who correctly identified the emotion divided by the total number of raters) and mean arousal were 0.97 and 2.4 ( $SD = 1.09$ ) for the selected neutral stimulus and 0.7 and 3.13 ( $SD = 0.78$ ) for the selected surprised stimulus. The neutral stimulus had a duration of 501 ms, whereas the surprised stimulus had a duration of 502 ms.

The stimuli were selected based on their rating results and their suitability for the compression/stretching manipulation employed here. Specifically, interjections were manipulated with Celemony Melodyne 2, a commercial sound manipulation software that implements an algorithm for duration change. According to the developer information, this algorithm was designed to enable duration change without altering average pitch and short-term spectral features. Looking at mean pitch and pitch variation we could confirm this claim for a larger stimulus set presented elsewhere (Schirmer et al., in press). However, the algorithm appears to affect harmonics-to-noise ratio (HNR), reducing it in the case of sound compression and increasing it in the case of sound stretching. Nevertheless, we subjected the stimuli described above to the algorithm resulting in a short version of 378 ms and a long version of 600 ms (**Figure 1**). The short and long versions differed in HNR by 2.5 dB only. Sound amplitudes were normalized in MATLAB to the same root-mean-square (RMS) value.

### Paradigm

For an illustration of the paradigm please see **Figure 2**. The sounds were presented in four blocks, each comprising





630 standards and 105 deviants. Two blocks used surprised expressions and the other two blocks used neutral expressions. One surprised and one neutral expression block used the short stimulus as the standard and the long stimulus as the deviant. The other two blocks used the long stimulus as the standard and the short stimulus as the deviant. Deviants were presented pseudorandomly such that they followed a sequence of 3–9 standards. Stimulus onset asynchrony was kept constant at 1.2 s. Block order was counterbalanced across participants using a Latin Square design.

The stimuli were played over ear-insert headphones at a comfortable sound pressure level that was kept constant across participants. While listening to the sounds, participants watched a self-selected, silent, subtitled movie. They were offered nine movies to choose from, which had an average duration of 109 min (range 94–134) and were classified by the British Board of Film as 15, 12a or appropriate for universal audiences. The authors of this research considered the movies to be only mildly arousing as they excluded comedies, dramas, action or extremely violent movies. The experiment lasted about an hour.

### Electrophysiological Recording and Analysis

The EEG was recorded using a 24-channel ANT system at a sampling rate of 256 Hz. Electrodes were placed according to the modified 10–20 system. An online anti-aliasing filter was applied with a cut-off frequency at 0.27 times the sampling rate.

The data were processed in EEGLAB. Continuous recordings were epoched and baseline-corrected using a 200 ms window

prior to stimulus onset and an 800 ms window following stimulus onset. After low- (30 Hz cutoff, 6.75 Hz transition band) and high-pass filtering (0.1 Hz cutoff, 0.2 Hz transition band), the data were scanned visually for non-typical artifacts and subjected to Infomax, an independent components algorithm. Components reflecting eye-blinks and saccades were removed, and the back-projected data scanned once more for residual artifacts.

For statistical analysis, we focused on nine frontal and central channels where the MMN is typically observed (i.e., FP1, F3, C3, Fpz, Fz, Cz, FP2, F4, C4). We included all deviant trials as well as standards immediately preceding a deviant that had survived artifact rejection. Visual inspection of the data failed to reveal a sharp MMN component. Instead, we observed a ramping negativity that developed throughout the stimulus epoch. To adequately capture this negativity, we divided stimulus epochs into eight 100 ms windows and subjected the mean voltages from these windows to an ANOVA with *Duration* (short/long), *Emotion* (neutral/surprised), *Deviance* (standard/deviant), and *Window* (1–8) as repeated measures factors and *Sex* as a between subjects factor. Due to our interest in temporal change detection, we focus this report on main effects and interactions involving *Deviance*.

### Results

Visual inspection of the ERPs revealed a climbing negativity, in some cases preceded by a positive dip, over fronto-central electrodes with polarity inversion over the mastoids. The negativity developed earlier and was larger for long relative to

short deviants and within the surprised relative to the neutral vocal stream (Figures 3 and 4). Effects seemed comparable for female and male listeners.

Statistical analysis confirmed these visual impressions. It revealed a significant interaction between *Deviance* and *Window* [ $F(7,210) = 46.5$ ,  $p < 0.0001$ ] compatible with a growing difference between standards and deviants throughout the course of the stimulus. Additionally, the *Window*  $\times$  *Deviance*  $\times$  *Duration* [ $F(7,210) = 6.9$ ,  $p < 0.0001$ ] and the *Window*  $\times$  *Deviance*  $\times$  *Emotion* [ $F(7,210) = 3.35$ ,  $p < 0.01$ ] interactions were significant. Hence, we pursued these effects for each time bin separately (compare time bin maps illustrated in Figure 4).

The *Deviance*  $\times$  *Duration* interaction was significant for Windows 1, 3, 4, 5, and 6 [ $F_s(1,30) > 7$ ,  $ps < .013$ ]. In the first window, short deviants elicited a more positive ERP than short standards [ $F(1,30) = 13.8$ ,  $p < 0.001$ ], but long deviants and standards did not differ ( $p > 0.1$ ). In the third and fourth windows, long deviants elicited a more negative ERP than long standards [ $F_s(1,30) > 7$ ,  $ps < 0.013$ ], but short deviants and standards did not differ ( $ps > 0.1$ ). In the fifth and sixth windows, the ERPs for both long [ $F_s(1,30) > 41$ ,  $ps < 0.0001$ ] and short duration stimuli [ $F_s(1,30) > 11$ ,  $ps < 0.003$ ] were more negative for deviants than standards.

The *Deviance*  $\times$  *Emotion* interaction was significant for Windows 1, 2, and 3 [ $F_s(1,30) > 6.5$ ,  $ps < 0.016$ ]. In the first and second windows, neutral deviants elicited a more positive ERP than neutral standards [ $F_s(1,30) > 13.7$ ,  $ps < 0.001$ ], whereas surprised deviants and standards did not differ ( $ps > 0.1$ ). In the third window, surprised deviants elicited a more negative ERP than surprised standards [ $F(1,30) = 8.9$ ,  $p < 0.01$ ], whereas neutral deviants and standards did not differ ( $p > 0.1$ ).

In windows 7 and 8, only the *Deviance* main effect reached significance [ $F_s(1,30) = 44$ ,  $p < 0.0001$ ]; irrespective of *Duration* and *Emotion*, deviants elicited a more negative ERP than standards.

## Discussion

Experiment 1 explored whether listeners perceive changes in the acoustic rate of unattended vocalizations and whether this perception is facilitated within an emotional as compared to a neutral auditory context. Our results support both propositions. Although the typical MMN component was absent, we found a climbing negativity for both long and short deviants with MMN-like topography and polarity inversion over the mastoids. Moreover, in the case of long deviants, this negativity emerged at around 200 ms and thus 178 ms before the duration value of short standards. This indicates that listeners were sensitive to a slow-down in stimulus speed. There was also a climbing negativity for short deviants. This negativity developed around 400 ms, thus, roughly coinciding with deviant offset and preceding the long standard offset time by 200 ms. Note that a mismatch response driven entirely by duration should manifest approximately 200 ms after deviant offset. So, although later and smaller than the effect for long deviants, the short deviant effect suggests that listeners perceive an unattended speed-up.

Emotions modulated the emergence of temporal deviant effects in the ERP. The climbing negativity appeared about 100 ms earlier in the surprised relative to the neutral stimulus stream, regardless of deviant duration. Notably, this effect showed irrespective of listener sex suggesting that men and women were equally sensitive to the vocal emotional context.

In the following section, we report a second experiment aimed at pursuing these results further. Specifically, by presenting spectrally rotated versions of the sounds used in Experiment 1, we intended to remove obviously human vocal features from the stimuli, while preserving temporal and spectral stimulus complexity (Blessner, 1972; Scott et al., 2000; Warren et al., 2006; Christmann et al., 2014). The sounds were created by flipping the frequency spectrum around a central frequency resulting in an “alien” sound quality. Thus, we hoped to answer two questions. First, we were interested in determining whether the time course of temporal change detection depends on the presentation of socially relevant vocal expressions as compared with non-vocal sounds. In other words, are differences in sensitivity to a slow-down and speed-up in acoustic rate modulated by whether the sound has human qualities? Second, we wished to determine whether emotion effects in Experiment 1 were due to affective or acoustic stimulus characteristics. Perhaps sound idiosyncrasies rather than their emotional meaning affected the time course of temporal change detection.

## EXPERIMENT 2

### Methods

This research was approved by the Institutional Review Board of the National University of Singapore.

### Participants

We recruited 33 participants. The data from one participant were discarded due to excessive artifacts in the EEG. Half of the remaining participants were male and the other half were female. Their average age was 22.5 ( $SD = 2.3$ ). Participants reported normal hearing and normal or corrected to normal vision. They were reimbursed for their time at a rate of S\$10/hour.

### Stimuli

The interjections from Experiment 1 were low-pass filtered (3.8 kHz) and subjected to a spectral rotation around 2 kHz (Blessner, 1972; Scott et al., 2000). Sound amplitudes were normalized in MATLAB to the same RMS value. The sounds are illustrated in Figure 5.

### Paradigm

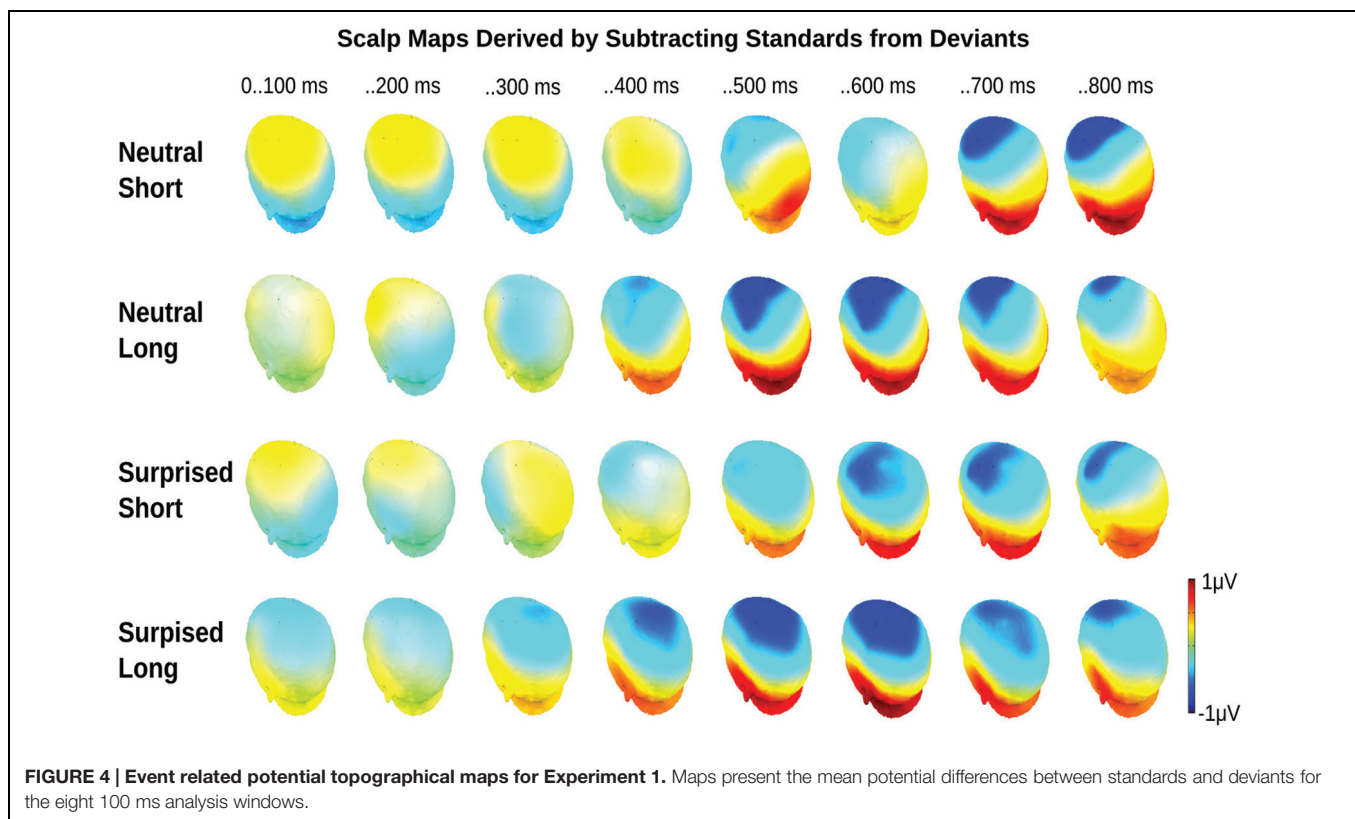
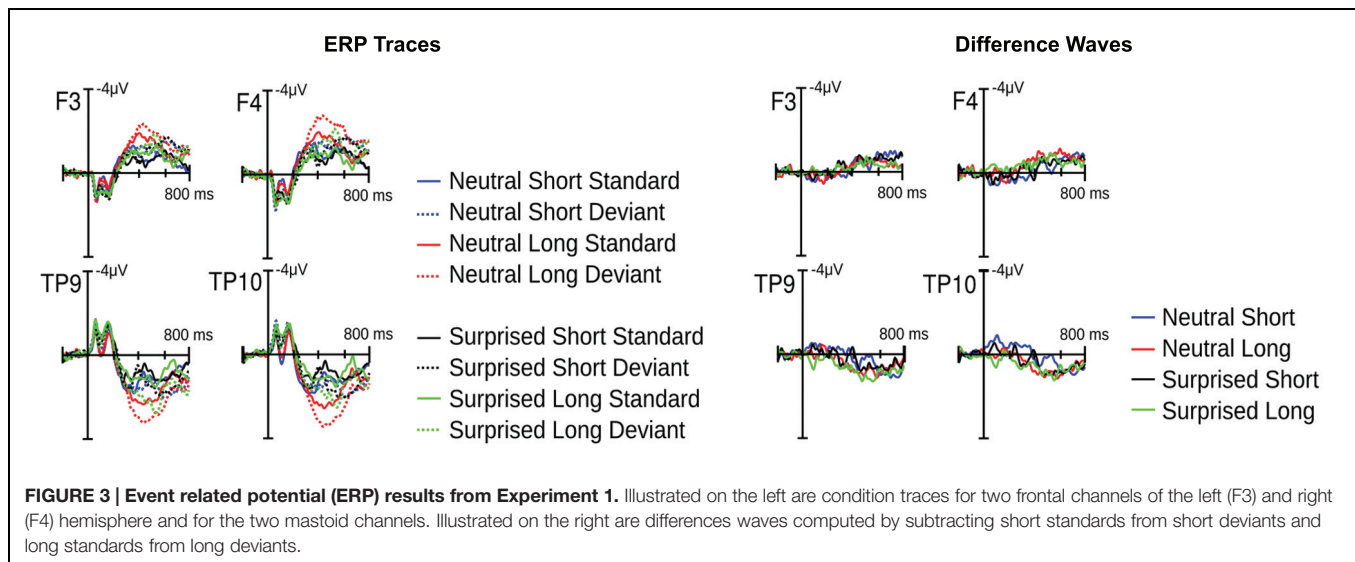
The paradigm was the same as in Experiment 1.

### Electrophysiological Recording and Analysis

Data recording and processing were the same as in Experiment 1.

## Results

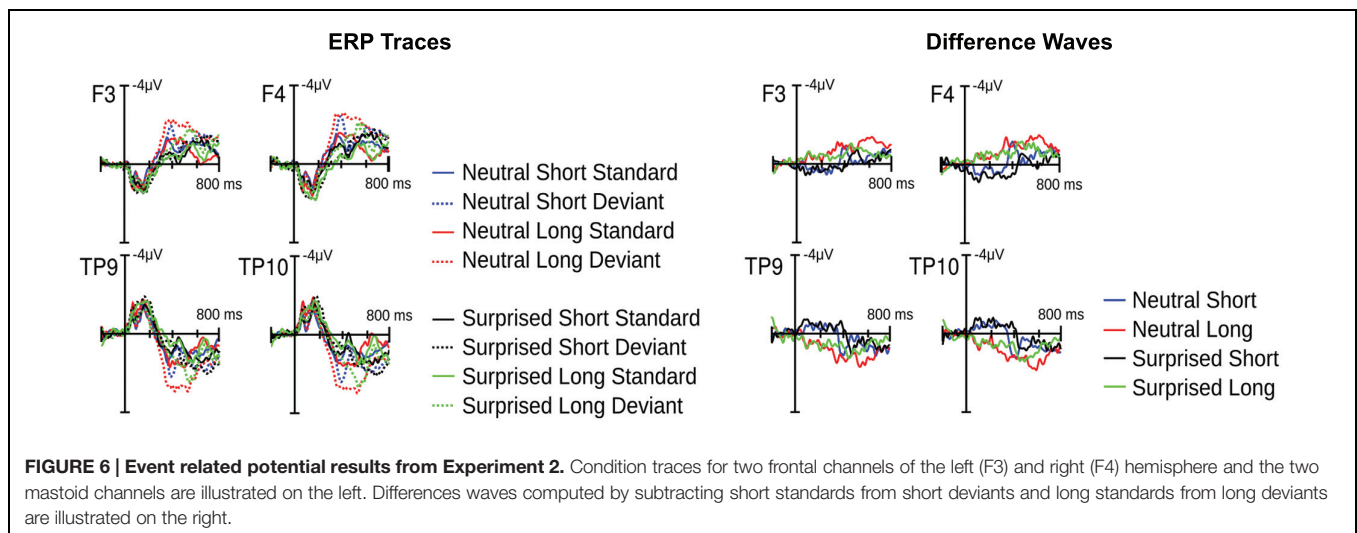
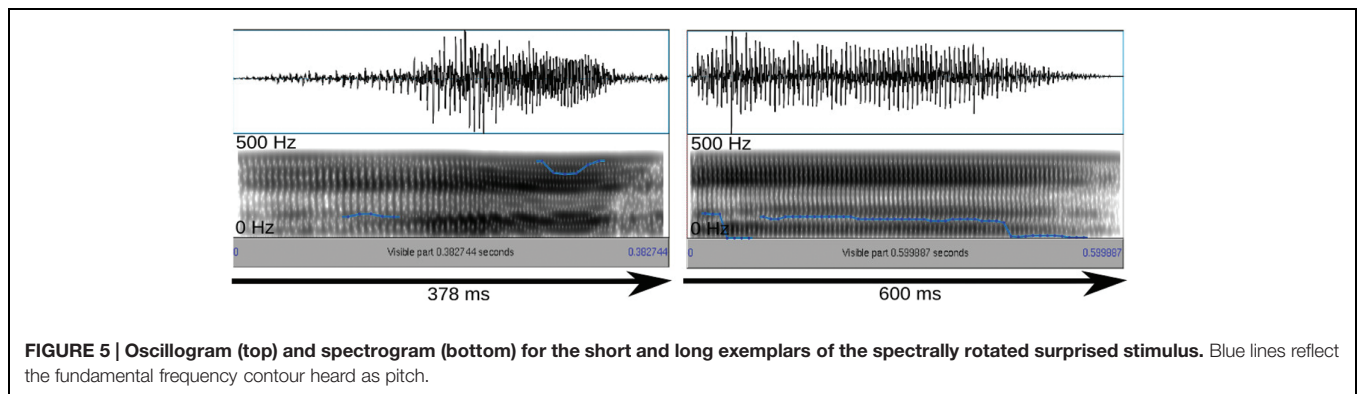
Visual inspection of the ERP suggested similarities with and differences from Experiment 1. There was again a climbing



fronto-central negativity with polarity inversion over the mastoids for deviants relative to standards. Again, in some cases, this negativity was preceded by a positive dip. Moreover, long deviants produced an earlier and larger negativity than short deviants. However, differences between surprised and neutral vocalizations appeared reversed relative to Experiment 1 (Figures 6 and 7).

To probe these visual impressions, we subjected mean voltages from eight 100 ms windows to an ANOVA with

*Duration*, *Emotion*, *Deviance*, and *Window* as repeated measures factors and *Sex* as a between subjects factor. This analysis revealed a significant *Deviance*  $\times$  *Window* interaction [ $F(7,210) = 35.4, p < 0.0001$ ] compatible with a growing difference between standards and deviants throughout the course of the stimulus. Additionally, the *Window*  $\times$  *Deviance*  $\times$  *Duration* [ $F(7,210) = 9.4, p < 0.0001$ ], the *Window*  $\times$  *Deviance*  $\times$  *Emotion* [ $F(7,210) = 5.3, p < 0.0001$ ], and the *Window*  $\times$  *Deviance*  $\times$  *Duration*  $\times$  *Emotion* [ $F(7,210) = 2.1,$



$p < 0.05$ ] interactions were significant. Hence, we pursued these effects for each time bin separately (compare maps presented in Figure 7).

The *Deviance*  $\times$  *Duration*  $\times$  *Emotion* interaction was significant in the first analysis window only [ $F(1,30) = 4.5$ ,  $p < 0.05$ ; other  $ps > 0.1$ ]. A follow-up analysis for long durations revealed an effect of *Deviance* only [ $F(1,30) = 13.5$ ,  $p < 0.001$ ]. Long deviants elicited a more negative ERP than long standards. A follow-up analysis for short durations revealed a *Deviance*  $\times$  *Emotion* interaction [ $F(1,30) = 5.6$ ,  $p < 0.05$ ] indicating that the *Deviance* effect was significant for surprised [ $F(1,30) = 29.9$ ,  $p < 0.0001$ ], but not neutral expressions ( $p > 0.1$ ). Over the first 100 ms following stimulus onset, surprised deviants elicited a more positive ERP than surprised standards.

The *Deviance*  $\times$  *Duration* interaction was significant in windows 2, 3, 4, 5, 6, and 7 [ $Fs(1,30) > 6.7$ ,  $ps < 0.05$ ]. In windows 2, 3, and 4, the ERP was more positive for short deviants than for short standards [ $Fs(1,30) = 36, 24$ ,  $ps < 0.0001$ ;  $F(1,30) = 3.4$ ,  $p = 0.07$ ] and more negative for long deviants than for long standards [ $Fs(1,30) > 25$ ,  $ps < 0.0001$ ]. Subsequently, both short [ $Fs(1,30) > 9$ ,  $ps < 0.01$ ] and long [ $Fs(1,30) > 42$ ,  $ps < 0.0001$ ] deviants elicited more negative potentials than standards. However, the effect for long durations was greater.

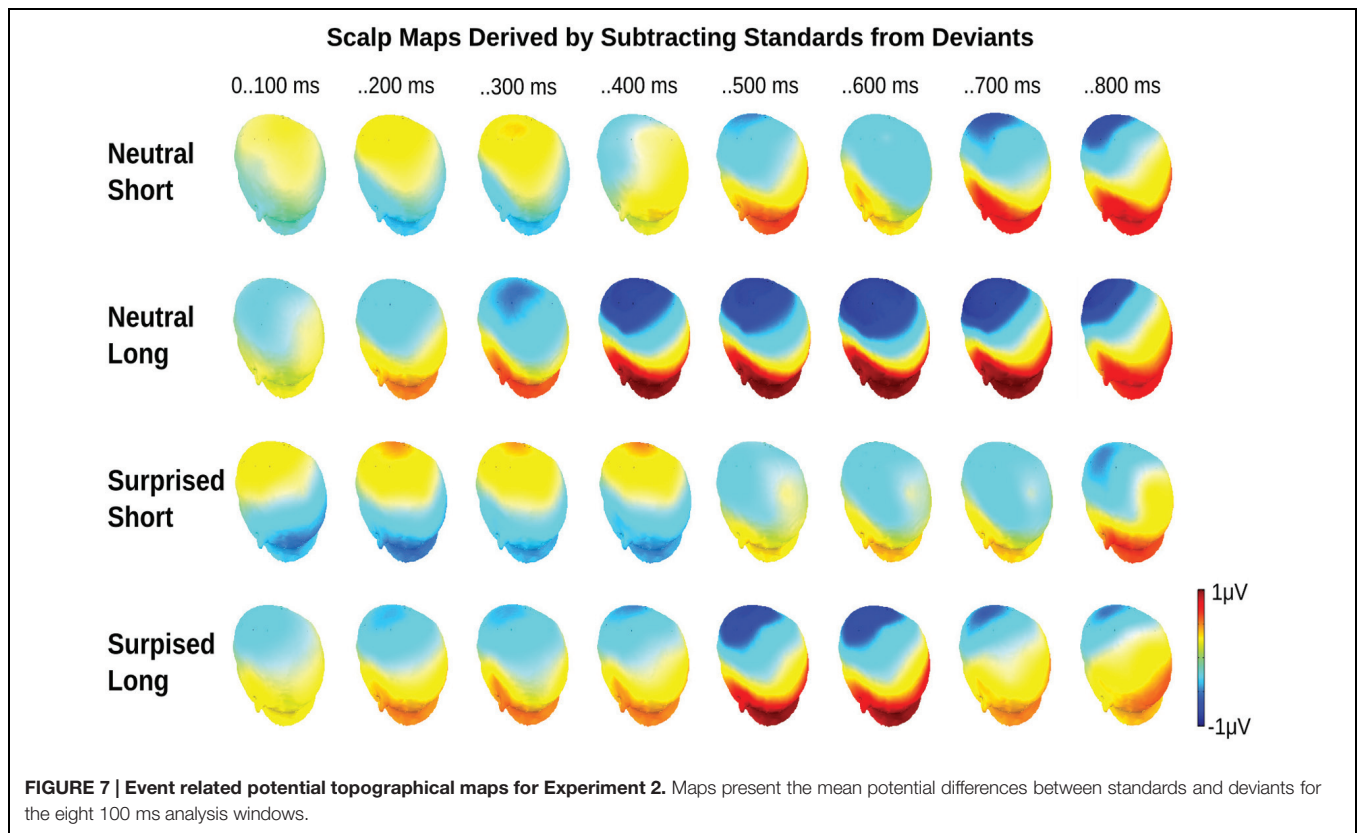
The *Deviance*  $\times$  *Emotion* interaction was significant in windows 4 and 7 [ $Fs(1,30) > 10$ ,  $ps < 0.01$ ]. In window 4, the deviant ERP was more negative than the standard ERP in the neutral [ $F(1,30) = 30.7$ ,  $p < 0.0001$ ], but not the surprised condition ( $p > 0.1$ ). In window 7, the deviant ERP was more negative than the standard ERP in both conditions, but this difference was larger for the neutral [ $F(1,30) = 49.2$ ,  $p < 0.0001$ ] than the surprised [ $F(1,30) = 10.5$ ,  $p < 0.01$ ] condition.

In window 8, only the *Deviance* main effect reached significance [ $F(1,30) = 45.5$ ,  $p < 0.0001$ ]. Irrespective of *Duration* and *Emotion*, deviants elicited a more negative ERP than standards.

## GENERAL DISCUSSION

The present study explored the implicit processing of temporal change within emotional and neutral streams of dynamic stimuli. Using an MMN paradigm, we found evidence that listeners mentally represent task-irrelevant increments and decrements in stimulus speed and that their representations differ as a function of emotion. In the following, we highlight the contributions of these results to the literature on (1) dynamic timing, (2) the asymmetry of temporal





change effects, as well as (3) the role of emotions for time perception.

## Temporal Processing of Dynamic Events

To our knowledge, this is the first demonstration that listeners mentally represent, not only duration, but also the speed with which task-irrelevant events unfold. More negative ERPs for deviants than standards emerged shortly after deviant onset and, thus, before the standard or deviant duration had lapsed. This points to an immediate sensitivity to the rate at which auditory neurons are excited. Furthermore, it suggests that an increasing disparity between standard and deviant time course contributes incrementally to an emerging representation of temporal change. We speculate that this representation then automatically biases an individual's perception of stimulus duration.

Explicit timing research accords with this speculation. Participants asked to judge the duration of dynamic stimuli demonstrate temporal distortions pointing to an influence of stimulus content or speed (Eagleman, 2008; Matthews, 2011; Liverence and Scholl, 2012; Linares and Gorea, 2015). For example, a greater frequency of loops (i.e., change in content) made by a luminance blob was associated with a lengthening of subjective duration (Linares and Gorea, 2015). Furthermore, faster changes in vocal pitch (i.e., change in speed) were associated with a shortening of subjective duration (Schirmer et al., in press). Our results add to this literature by addressing implicit timing and by showing that representations of stimulus speed emerge incrementally throughout a stimulus and could,

hence, contribute to duration perception in a feed-forward manner.

The present auditory change effect differed somewhat from the auditory change effect seen in previous MMN studies (for reviews see Näätänen et al., 2005; Garrido et al., 2009). Previous reports described a negative component with a fronto-central topography, that inverts polarity over the mastoids and peaks about 200 ms following deviant onset. Our effects match these properties with the exception that there was no clearly defined component peak. Instead, we observed a climbing negativity that plateaued later in the epoch. Nevertheless, we suspect this negativity to be an MMN and to index auditory change detection. Differences in time course or amplitude contour probably arise from the nature of our manipulation, which produced an incremental rather than a sudden difference between standards and deviants.

Notably, the MMN effect observed here resembles a negative ERP deflection often reported in the timing literature. This deflection, called the contingent negative variation (CNV), is maximal over fronto-central leads and was shown by some to increase in amplitude with increasing stimulus duration (Macar and Vitton, 1980). As such it was thought to reflect temporal encoding (Macar and Vidal, 2003). More recently, however, the CNV is deemed more likely to be an indicator of response preparation or temporal decision-making (Kononowicz and van Rijn, 2011; Ng et al., 2011; Van Rijn et al., 2011). Given the resemblance of CNV and the present temporal change effect one may wonder whether these are distinct or overlapping

phenomena. We suggest that they are distinct based on effect topography and eliciting conditions. The present climbing negativity, but not the CNV (Ng and Penney, 2014), has sources in the primary auditory cortex as indicated by ERP inversion over mastoid electrodes. Additionally, one would expect a CNV like effect in an MMN paradigm for both standards and deviants with amplitude differences emerging only after the point of duration mismatch. The effects observed here, however, were present before this point.

## Asymmetry in the Sensitivity to Temporal Change

Existing research indicates that listeners differ in their sensitivity to duration decrements and increments (but see Amenedo and Escera, 2000; Peter et al., 2012). Some studies have found a larger MMN to short than to long deviants (Jaramillo et al., 1999; Colin et al., 2009), whereas others (Catts et al., 1995; Takegata et al., 2008; Atkinson et al., 2012), including the present study, have found the opposite. Three factors were cited to explain this variation. First, asymmetry in the sensitivity to temporal change may depend on stimulus properties (Jaramillo et al., 1999; Takegata et al., 2008). For example, Jaramillo et al. (1999) found a larger MMN to short than to long vowels, but a smaller MMN to short than to long tones pointing to processing differences between vocal and non-vocal sounds. Notably, the present study conflicts with this result. The MMN to both interjections (Experiment 1) and their spectrally rotated counterparts (Experiment 2) was greater in the long relative to the short condition.

A second explanation of the asymmetry in temporal change detection invokes a role of absolute stimulus duration. For stimuli that are less than 200 ms long, an increase in duration is perceived as an increase in sound intensity (Takegata et al., 2008). Thus, a 150 ms deviant may be perceived as louder than a preceding 100 ms standard, whereas a 100 ms deviant may be perceived as softer than a preceding 150 ms standard. At these durations, then, differences in MMN magnitude may result from an asymmetry in the perception of sound intensity rather than duration (Peter et al., 2010). For stimuli that exceed 200 ms, perceived sound intensity does not differ as a function of duration. In the present study, stimuli were 378 and 600 ms long, so an intensity illusion is unlikely to account for the observed MMN effects.

Last, asymmetry in temporal change detection has been linked to the method by which an MMN is generated. Traditionally, researchers subtracted standards from deviants in the same block. More recently, however, approaches controlling for the physical difference between standards and deviants have become popular (Jacobsen and Schröger, 2003). One approach involves an additional experimental block in which deviants are presented in an equiprobable manner together with other stimuli. Another approach involves an additional experimental block in which the role of standards and deviants is reversed. In either case, an MMN is generated by subtracting the physically identical control stimulus from the deviant. A comparison of the traditional with the physical control approach indicated that MMN asymmetries for duration deviants are present in the former, but not the latter

(Peter et al., 2010). Again, this explanation does not account for our results as we used the physical control approach, but nevertheless observed differences in the MMN to short and long deviants.

Together, the literature on MMN asymmetry is inconclusive. Although asymmetry is frequently observed, it seems to depend on a number of stimulus parameters. Moreover, these parameters may include two methodological novelties implemented in the present study. Our stimulus durations were fairly long and we manipulated duration alongside speed. Both of these factors may explain why our results are similar to some studies, but differ from others.

Although we are unable to explain variation in MMN asymmetry, we can offer an explanation for why the MMN was larger for increments here. Specifically, behavioral research on timing suggests that repeated or expected stimuli become subjectively shorter than non-repeated or unexpected stimuli, respectively (Pariyadath and Eagleman, 2007; Matthews et al., 2014; for a possible dissociation between repetition and expectation see Matthews, 2015). Moreover, because temporal lengthening and shortening are typically associated with a speed increase and decrease in a dynamic context, these illusions may extend to the perception of speed. Repeated and/or expected events may appear faster than unrepeated and/or unexpected events. Thus, in the present MMN paradigm, long and short deviants might have appeared more and less different from their standard, respectively, and this asymmetry probably emerged as speed discrepancies accumulated.

Before closing the discussion of MMN asymmetry, we would like to add a caveat that complicates the interpretation of duration effects. Likely, here and elsewhere, duration deviants not only violated the expectation for a particular stimulus duration, but also distorted an overall rhythmic structure established by standards (Jones, 1976; Escoffier et al., 2010). Blocks, although comparable in stimulus onset timing, differed in standard duration and speed possibly creating different emphases or metric points that then modulated stimulus processing. As illustrated in **Figure 2**, the longer standards may have been metrically more important than the short standards, thus, producing a stronger entrainment and more readily accommodating temporal deviants. Such a modulation would be revealed by an interaction between stimulus deviance (standard/deviant) and duration (short/long), whereby the deviance effect would be reversed for the short and long duration conditions. In other words, there would be a block effect whereby short standards and long deviants presented together in one block would differ in comparable ways from long standards and short deviants presented together in another block.

Statistical analysis of the present data revealed patterns suggesting such block effects. In Experiment 1 (vocal), initially, short deviants elicited a more positive potential than short standards, whereas the deviance effect for long stimuli was non-significant. In Experiment 2 (non-vocal), initially, short deviants elicited a more positive potential, while long deviants elicited a more negative potential relative to their respective standards. Thus, rhythmic processing likely occurred alongside duration processing and affected the ERP. The present as well

as previous studies cannot dissociate the two. To achieve this, future research could compare temporally regular with irregular stimulus presentations (McAuley and Fromboluti, 2014).

## Emotions and Temporal Change Detection

A final objective of this study was to determine whether and how emotions influence temporal change detection. As expected, we found an earlier change effect for surprised relative to neutral vocalizations. This latency difference reversed when vocalizations were spectrally rotated and human expressiveness was removed (Warren et al., 2006). Moreover, while the time course of ERP effects was comparable for neutral original and rotated sounds, that of surprised original and rotated sounds differed. Additionally, the amplitude of the MMN-like effect in neutral and surprised conditions was comparable for original vocalizations, whereas the same effect was larger for the neutral than the surprised condition for spectrally rotated sounds. Together, these observations suggest that vocal emotions overwrote the processing differences that were due to non-emotional stimulus properties (e.g., stimulus complexity) and facilitated the neural representation of temporal change.

The present emotion effect concurs with prior research demonstrating emotion effects on the MMN and on performance in temporal judgment tasks. As reviewed in the introduction, emotional deviants following neutral standards elicit an earlier and larger MMN than neutral deviants following emotional standards (e.g., Schirmer et al., 2005). Additionally, stimuli with emotional valence are perceived as longer than same-duration stimuli with neutral valence (e.g., Grommet et al., 2011). Together these and the present results suggest that emotions enhance the salience of a stimulus stream ensuring sufficient processing resources despite being task-irrelevant.

This proposition is in line with recent evidence for a role of attention in the link between emotions and time. Rather than asking participants to time emotional and neutral stimuli, Lui et al. (2011) manipulated the emotionality of distractors presented while participants timed simple visual shapes. Specifically, an initial circle (S1) was followed by a distractor and then a second circle (S2) and participants decided whether S2 was shorter or longer than S1. Participants were more likely to judge S2 to be shorter than S1 if the distractor was emotional as compared to neutral. This suggests that S2 indeed seemed shorter in the emotional as compared to neutral context. Moreover, it implies that emotional stimuli capture and bind processing resources at the cost of other neutral stimuli (for similar approaches and results see Halbertsma and Van Rijn, in press; Lake et al., in press).

Change detection research accords with this. Apart from the present study, there is one previous attempt at comparing the MMN in an emotional and neutral context. Lv et al. (2011) asked participants to indicate whether two faces presented side-by-side were identical. In different blocks, faces had a sad or neutral expression. In the background, participants heard a sequence of

tones that contained rare deviants. Unlike in the present study, the MMN was comparable in the emotional and neutral blocks. However, face discrimination was facilitated by sad relative to neutral expressions. Thus, emotions effectively held attention to the visual material thereby benefiting visual categorization performance rather than the processing of an unrelated stream of neutral sounds.

Although in agreement with existing work, the present study failed to identify sex differences in the emotion effect. Temporal change detection was enhanced for surprised relative to neutral vocal streams in both men and women. Given earlier work (Fan et al., 2013; Schirmer et al., 2005, in press), this result was somewhat unexpected. We suspect that the absence of sex effects here relates to the fact that emotions varied in a blocked rather than an event-related manner. There is some indication that sex differences in vocal emotion sensitivity are a matter of processing time and attention. In an implicit emotional priming paradigm, only women showed priming at a short (200 ms) prime-target interval, whereas both sexes showed priming at a longer (750 ms) prime-target interval (Schirmer et al., 2002). Moreover, when emotions were made task-relevant by asking participants to categorize stimuli based on emotion, sex differences likewise disappeared (Schirmer et al., 2006, 2013). Thus, we speculate that the repetition of vocal expressions within blocks enabled participants to “tune into” a particular emotion condition putting male and female processing on par.

## CONCLUSION

For complex environmental stimuli, stimulus duration is necessarily confounded by content and/or speed. Here, we made a first attempt at assessing combined duration and speed perception without highlighting temporal processing to our participants. In the context of an MMN paradigm, we observed a climbing MMN-like negativity emerging shortly after deviant onset. Thus, we conclude that stimulus speed is tracked continuously and suggest that it biases duration perception well before stimulus offset in a feed-forward manner. In the present study, like some others, mismatch effects showed earlier and were larger to duration increments relative to decrements. Although the factors underpinning this are still unclear, we suspect a role of stimulus predictability. Contrasting repeated standards with singular deviants can be expected to augment and diminish perceived temporal change for long and short deviants, respectively. As was demonstrated before, emotions influenced temporal processing. Change was detected earlier within a stream of surprised relative to neutral vocalizations suggesting that the former recruit more processing resources than the latter.

## AUTHOR CONTRIBUTIONS

AS was involved in study design, data analysis, and manuscript writing. NE was involved in study design, experimental



programming, data acquisition and analysis. He also commented on the manuscript. XC and YF were involved in data acquisition and commented on the manuscript. TP was involved in study design and manuscript writing.

## REFERENCES

- Amenedo, E., and Escera, C. (2000). The accuracy of sound duration representation in the human brain determines the accuracy of behavioural perception. *Eur. J. Neurosci.* 12, 2570–2574. doi: 10.1046/j.1460-9568.2000.00114.x
- Angrilli, A., Cherubini, P., Pavese, A., and Mantredini, S. (1997). The influence of affective factors on time perception. *Percept. Psychophys.* 59, 972–982. doi: 10.3758/BF03205512
- Atkinson, R. J., Michie, P. T., and Schall, U. (2012). Duration mismatch negativity and P3a in first-episode psychosis and individuals at ultra-high risk of psychosis. *Biol. Psychiatry* 71, 98–104. doi: 10.1016/j.biopsych.2011.08.023
- Blesser, B. (1972). Speech perception under conditions of spectral transformation. I. Phonetic characteristics. *J. Speech Hear. Res.* 15, 5–41. doi: 10.1044/jshr.1501.05
- Catts, S. V., Shelley, A. M., Ward, P. B., Liebert, B., McConaghy, N., Andrews, S., et al. (1995). Brain potential evidence for an auditory sensory memory deficit in schizophrenia. *Am. J. Psychiatry* 152, 213–219. doi: 10.1176/ajp.152.2.213
- Christmann, C. A., Berti, S., Steinbrink, C., and Lachmann, T. (2014). Differences in sensory processing of German vowels and physically matched non-speech sounds as revealed by the mismatch negativity (MMN) of the human event-related brain potential (ERP). *Brain Lang.* 136, 8–18. doi: 10.1016/j.bandl.2014.07.004
- Colin, C., Hoonhorst, I., Markessis, E., Radeau, M., de Tourtchaninoff, M., Foucher, A., et al. (2009). Mismatch Negativity (MMN) evoked by sound duration contrasts: an unexpected major effect of deviance direction on amplitudes. *Clin. Neurophysiol.* 120, 51–59. doi: 10.1016/j.clinph.2008.10.002
- Darlow, H. M., Dylman, A. S., Gheorghiu, A. I., and Matthews, W. J. (2013). Do changes in the pace of events affect one-off judgments of duration? *PLoS ONE* 8:e59847. doi: 10.1371/journal.pone.0059847
- Droit-Volet, S., Bigand, E., Ramos, D., and Bueno, J. L. O. (2010). Time flies with music whatever its emotional valence. *Acta Psychol. (Amst.)* 135, 226–232. doi: 10.1016/j.actpsy.2010.07.003
- Droit-Volet, S., and Meck, W. H. (2007). How emotions colour our perception of time. *Trends Cogn. Sci.* 11, 504–513. doi: 10.1016/j.tics.2007.09.008
- Eagleman, D. M. (2008). Human time perception and its illusions. *Curr. Opin. Neurobiol.* 18, 131–136. doi: 10.1016/j.conb.2008.06.002
- Escoffier, N., Herrmann, C. S., and Schirmer, A. (2015). Auditory rhythms entrain visual processes in the human brain: evidence from evoked oscillations and event-related potentials. *Neuroimage* 111, 267–276. doi: 10.1016/j.neuroimage.2015.02.024
- Escoffier, N., Sheng, D. Y. J., and Schirmer, A. (2010). Unattended musical beats enhance visual processing. *Acta Psychol. (Amst.)* 135, 12–16. doi: 10.1016/j.actpsy.2010.04.005
- Fan, Y.-T., Hsu, Y.-Y., and Cheng, Y. (2013). Sex matters: n-back modulates emotional mismatch negativity. *Neuroreport* 24, 457–463. doi: 10.1097/WNR.0b013e32836169b9
- Fayolle, S. L., and Droit-Volet, S. (2014). Time perception and dynamics of facial expressions of emotions. *PLoS ONE* 9:e97944. doi: 10.1371/journal.pone.0097944
- Firmino, É.A., Bueno, J. L. O., and Bigand, E. (2009). Travelling through pitch space speeds up musical time. *Music Percept. Interdiscip. J.* 26, 205–209. doi: 10.1525/mp.2009.26.3.205
- Gallistel, C. R., and Gibbon, J. (2000). Time, rate, and conditioning. *Psychol. Rev.* 107, 289–344. doi: 10.1037/0033-295X.107.2.289
- Garrido, M. I., Kilner, J. M., Stephan, K. E., and Friston, K. J. (2009). The mismatch negativity: a review of underlying mechanisms. *Clin. Neurophysiol.* 120, 453–463. doi: 10.1016/j.clinph.2008.11.029
- Grommet, E. K., Droit-Volet, S., Gil, S., Hemmes, N. S., Baker, A. H., and Brown, B. L. (2011). Time estimation of fear cues in human observers. *Behav. Process.* 86, 88–93. doi: 10.1016/j.beproc.2010.10.003
- Halbertsma, H., and Van Rijn, H. (in press). *An Evaluation of the Effect of Auditory Emotional Stimuli on Interval Timing. Timing Time Percept.*
- Jacobsen, T., and Schröger, E. (2003). Measuring duration mismatch negativity. *Clin. Neurophysiol.* 114, 1133–1143. doi: 10.1016/S1388-2457(03)00043-9
- Jaramillo, M., Paavo, A., and Paavilainen, P. (1999). An event-related potential (ERP) study of duration changes I: neuroReport. *Neuroreport* 10, 3301–3305. doi: 10.1097/00001756-199911080-00010
- Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychol. Rev.* 83, 323–355. doi: 10.1037/0033-295X.83.5.323
- Jones, M. R., and Boltz, M. (1989). Dynamic attending and responses to time. *Psychol. Rev.* 96, 459–491. doi: 10.1037/0033-295X.96.3.459
- Kaneko, S., and Murakami, I. (2009). Perceived duration of visual motion increases with speed. *J. Vis.* 9:14. doi: 10.1167/9.7.14
- Kononowicz, T. W., and van Rijn, H. (2011). Slow potentials in time estimation: the role of temporal accumulation and habituation. *Front. Integr. Neurosci.* 5:48. doi: 10.3389/fnint.2011.00048
- Lake, J., Meck, W. H., and LaBar, K. S. (in press). *Discriminative Fear Learners are Resilient to Temporal Distortions during Threat Anticipation. Timing Time Percept.*
- Linares, D., and Gorea, A. (2015). Temporal frequency of events rather than speed dilates perceived duration of moving objects. *Sci. Rep.* 5:8825. doi: 10.1038/srep08825
- Liverence, B. M., and Scholl, B. J. (2012). Discrete events as units of perceived time. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 549–554. doi: 10.1037/a0027228
- Lui, M. A., Penney, T. B., and Schirmer, A. (2011). Emotion effects on timing: attention versus pacemaker accounts. *PLoS ONE* 6:e21829. doi: 10.1371/journal.pone.0021829
- Lv, J.-Y., Wang, T., Tu, S., Zheng, F., and Qiu, J. (2011). The effect of different negative emotional context on involuntary attention: an ERP study. *Brain Res. Bull.* 86, 106–109. doi: 10.1016/j.brainresbull.2011.06.010
- Macar, F., and Vidal, F. (2003). The CNV peak: an index of decision making and temporal memory. *Psychophysiology* 40, 950–954. doi: 10.1111/1469-8986.00113
- Macar, F., and Vitton, N. (1980). CNV and reaction time task in man: effects of inter-stimulus interval contingencies. *Neuropsychologia* 18, 585–590. doi: 10.1016/0028-3932(80)90161-X
- Matthews, W. J. (2011). How do changes in speed affect the perception of duration? *J. Exp. Psychol. Hum. Percept. Perform.* 37, 1617–1627. doi: 10.1037/a0022193
- Matthews, W. J. (2013). How does sequence structure affect the judgment of time? Exploring a weighted sum of segments model. *Cogn. Psychol.* 66, 259–282. doi: 10.1016/j.cogpsych.2013.01.001
- Matthews, W. J. (2015). Time perception: the surprising effects of surprising stimuli. *J. Exp. Psychol. Gen.* 144, 172–197. doi: 10.1037/xge0000041
- Matthews, W. J., Terhune, D. B., Rijn, H., van Eagleman, D. M., Sommer, M. A., and Meck, W. H. (2014). Subjective duration as a signature of coding efficiency: emerging links among stimulus repetition, predictive coding, and cortical GABA levels. *Timing Time Percept. Rev.* 1, 1–11. doi: 10.1163/24054496-00101005
- McAuley, J. D., and Fromboluti, E. K. (2014). Attentional entrainment and perceived event duration. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369:20130401. doi: 10.1098/rstb.2013.0401
- Näätänen, R., Jacobsen, T., and Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. *Psychophysiology* 42, 25–32. doi: 10.1111/j.1469-8986.2005.00256.x
- Näätänen, R., Paavilainen, P., and Reinikainen, K. (1989). Do event-related potentials to infrequent decrements in duration of auditory stimuli demonstrate a memory trace in man? *Neurosci. Lett.* 107, 347–352. doi: 10.1016/0304-3940(89)90844-6
- Ng, K. K., and Penney, T. B. (2014). Probing interval timing with scalp-recorded electroencephalography (EEG). *Adv. Exp. Med. Biol.* 829, 187–207. doi: 10.1007/978-1-4939-1782-2\_11

## FUNDING

This research was supported by an NUS Faculty Research Grant awarded to AS (R-581-000-152-112).



- Ng, K. K., Tobin, S., and Penney, T. B. (2011). Temporal accumulation and decision processes in the duration bisection task revealed by contingent negative variation. *Front. Integr. Neurosci.* 5:77. doi: 10.3389/fnint.2011.00077
- Cocenas-Silva, R., Bueno, J. L. O., Molin, P., and Bigand, E. (2011). Multidimensional scaling of musical time estimations. *Percept. Mot. Skills* 112, 737–748. doi: 10.2466/11.24.PMS.112.3.737-748
- Pariyadath, V., and Eagleman, D. (2007). The effect of predictability on subjective duration. *PLoS ONE* 2:e1264. doi: 10.1371/journal.pone.0001264
- Peter, V., McArthur, G., and Thompson, W. F. (2010). Effect of deviance direction and calculation method on duration and frequency mismatch negativity (MMN). *Neurosci. Lett.* 482, 71–75. doi: 10.1016/j.neulet.2010.07.010
- Peter, V., McArthur, G., and Thompson, W. F. (2012). Discrimination of stress in speech and music: a mismatch negativity (MMN) study. *Psychophysiology* 49, 1590–1600. doi: 10.1111/j.1469-8986.2012.01472.x
- Sams, M., Paavilainen, P., Alho, K., and Näätänen, R. (1985). Auditory frequency discrimination and event-related potentials. *Electroencephalogr. Clin. Neurophysiol.* 62, 437–448. doi: 10.1016/0168-5597(85)90054-1
- Schirmer, A. (2004). Timing speech: a review of lesion and neuroimaging findings. *Brain Res. Cogn. Brain Res.* 21, 269–287. doi: 10.1016/j.cogbrainres.2004.04.003
- Schirmer, A. (2011). How emotions change time. *Front. Integr. Neurosci.* 5:58. doi: 10.3389/fnint.2011.00058
- Schirmer, A., and Escoffier, N. (2010). Emotional MMN: anxiety and heart rate correlate with the ERP signature for auditory change detection. *Clin. Neurophysiol.* 121, 53–59. doi: 10.1016/j.clinph.2009.09.029
- Schirmer, A., Escoffier, N., Zysset, S., Koester, D., Striano, T., and Friederici, A. D. (2008). When vocal processing gets emotional: on the role of social orientation in relevance detection by the human amygdala. *Neuroimage* 40, 1402–1410. doi: 10.1016/j.neuroimage.2008.01.018
- Schirmer, A., Kotz, S. A., and Friederici, A. D. (2002). Sex differentiates the role of emotional prosody during word processing. *Brain Res. Cogn. Brain Res.* 14, 228–233. doi: 10.1016/S0926-6410(02)00108-8
- Schirmer, A., Lui, M., Maess, B., Escoffier, N., Chan, M., and Penney, T. B. (2006). Task and sex modulate the brain response to emotional incongruity in Asian listeners. *Emotion* 6, 406–417. doi: 10.1037/1528-3542.6.3.406
- Schirmer, A., Ng, T., Escoffier, N., and Penney, T. B. (in press). *Emotional Voices Distort Time Behavioral and Neural Correlates. Timing Time Percept.*
- Schirmer, A., Seow, C. S., and Penney, T. B. (2013). Humans process dog and human facial affect in similar ways. *PLoS ONE* 8:e74591. doi: 10.1371/journal.pone.0074591
- Schirmer, A., Simpson, E., and Escoffier, N. (2007). Listen up! Processing of intensity change differs for vocal and nonvocal sounds. *Brain Res.* 1176, 103–112. doi: 10.1016/j.brainres.2007.08.008
- Schirmer, A., Striano, T., and Friederici, A. D. (2005). Sex differences in the preattentive processing of vocal emotional expressions. *Neuroreport* 16, 635–639. doi: 10.1097/00001756-200504250-00024
- Scott, S. K., Blank, C. C., Rosen, S., and Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain J. Neurol.* 123 (Pt 12), 2400–2406. doi: 10.1093/brain/123.12.2400
- Su, Y.-H., and Jonikaitis, D. (2011). Hearing the speed: visual motion biases the perception of auditory tempo. *Exp. Brain Res.* 214, 357–371. doi: 10.1007/s00221-011-2835-4
- Takegata, R., Tervaniemi, M., Alku, P., Ylinen, S., and Näätänen, R. (2008). Parameter-specific modulation of the mismatch negativity to duration decrement and increment: evidence for asymmetric processes. *Clin. Neurophysiol.* 119, 1515–1523. doi: 10.1016/j.clinph.2008.03.025
- Thönnessen, H., Boers, F., Dammers, J., Chen, Y.-H., Norra, C., and Mathiak, K. (2010). Early sensory encoding of affective prosody: neuromagnetic tomography of emotional category changes. *Neuroimage* 50, 250–259. doi: 10.1016/j.neuroimage.2009.11.082
- Tse, C.-Y., and Penney, T. B. (2006). Preattentive timing of empty intervals is from marker offset to onset. *Psychophysiology* 43, 172–179. doi: 10.1111/j.1469-8986.2006.389.x
- Van Rijn, H., Kononowicz, T. W., Meck, W. H., Ng, K. K., and Penney, T. B. (2011). Contingent negative variation and its relation to time estimation: a theoretical evaluation. *Front. Integr. Neurosci.* 5:91. doi: 10.3389/fnint.2011.00091
- Voyer, D., and Reuengrith, E. (2015). Perceptual asymmetries in a time estimation task with emotional sounds. *Laterality* 20, 211–231. doi: 10.1080/1357650X.2014.953956
- Warren, J. E., Sauter, D. A., Eisner, F., Wiland, J., Dresner, M. A., Wise, R. J. S., et al. (2006). Positive emotions preferentially engage an auditory-motor “mirror” system. *J. Neurosci.* 26, 13067–13075. doi: 10.1523/JNEUROSCI.3907-06.2006

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Schirmer, Escoffier, Cheng, Feng and Penney. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# A Roving Dual-Presentation Simultaneity-Judgment Task to Estimate the Point of Subjective Simultaneity

Kielan Yarrow<sup>1\*</sup>, Sian E. Martin<sup>1</sup>, Steven Di Costa<sup>2</sup>, Joshua A. Solomon<sup>3</sup> and Derek H. Arnold<sup>4</sup>

<sup>1</sup> Department of Psychology, City University London, London, UK, <sup>2</sup> Department of Psychology, UCL Institute of Cognitive Neuroscience, London, UK, <sup>3</sup> Centre for Applied Vision Science, City University London, London, UK, <sup>4</sup> School of Psychology, The University of Queensland, Brisbane, QLD, Australia

## OPEN ACCESS

### Edited by:

Hugo Merchant,  
Universidad Nacional Autónoma de  
México, Mexico

### Reviewed by:

Shigeru Kitazawa,  
Osaka University School of  
Biosciences, Japan  
María Herrojo Ruiz,  
Charité-University of Medicine Berlin,  
Germany

### \*Correspondence:

Kielan Yarrow  
kielan.yarrow.1@city.ac.uk

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 09 December 2015

**Accepted:** 09 March 2016

**Published:** 24 March 2016

### Citation:

Yarrow K, Martin SE, Di Costa S,  
Solomon JA and Arnold DH (2016) A  
Roving Dual-Presentation  
Simultaneity-Judgment Task to  
Estimate the Point of Subjective  
Simultaneity. *Front. Psychol.* 7:416.  
doi: 10.3389/fpsyg.2016.00416

The most popular tasks with which to investigate the perception of subjective synchrony are the temporal order judgment (TOJ) and the simultaneity judgment (SJ). Here, we discuss a complementary approach—a dual-presentation (2x) SJ task—and focus on appropriate analysis methods for a theoretically desirable “roving” design. Two stimulus pairs are presented on each trial and the observer must select the most synchronous. To demonstrate this approach, in Experiment 1 we tested the 2xSJ task alongside TOJ, SJ, and simple reaction-time (RT) tasks using audiovisual stimuli. We interpret responses from each task using detection-theoretic models, which assume variable arrival times for sensory signals at critical brain structures for timing perception. All tasks provide similar estimates of the point of subjective simultaneity (PSS) on average, and PSS estimates from some tasks were correlated on an individual basis. The 2xSJ task produced lower and more stable estimates of model-based (and thus comparable) sensory/decision noise than the TOJ. In Experiment 2 we obtained similar results using RT, TOJ, ternary, and 2xSJ tasks for all combinations of auditory, visual, and tactile stimuli. In Experiment 3 we investigated attentional prior entry, using both TOJs and 2xSJs. We found that estimates of prior-entry magnitude correlated across these tasks. Overall, our study establishes the practicality of the roving dual-presentation SJ task, but also illustrates the additional complexity of the procedure. We consider ways in which this task might complement more traditional procedures, particularly when it is important to estimate both PSS and sensory/decisional noise.

**Keywords:** multisensory perception, timing and time perception, temporal order, 2AFC, simultaneity judgment

## INTRODUCTION

Introspection suggests conscious experiences proceed successively. This is part of what we mean when we say that we have a sensation of the passage of time. Determining the relative timing at which two or more events occur would thus appear to be an important perceptual operation, and one that might underscore various higher-level inferences, such as the causal relationship between events, or the degree to which two events should be grouped perceptually. However, the

processes by which the brain determines relative timing require clarification. The problem appears particularly acute for multisensory events, where relevant neural signals might be dispersed widely in space and time. However, even within a single sense, the way in which temporal succession and overlap are determined is not yet established.

One of the most fundamental questions one can ask about relative timing is with what objective asynchrony the observer considers two events to be maximally synchronous. We usually investigate this issue by attempting to estimate a point of subjective simultaneity (PSS) from different experimental conditions. Hence, to make progress, we need good experimental procedures to lay bare our temporal qualia. In this paper we consider a complementary task for this purpose: The dual-presentation simultaneity judgment (2xSJ). This type of task has been used fairly infrequently in relative timing experiments (e.g., Allan and Kristofferson, 1974; Van de Par and Kohlrausch, 2000; Powers et al., 2009; Roseboom et al., 2011; Stevenson and Wallace, 2013). Here we will first argue that a roving-standard design is theoretically desirable, second describe an appropriate observer model for fitting and summarizing the data this task generates, and third attempt to benchmark data from this task against more common approaches in order to assess its strengths and limitations.

## TEMPORAL JUDGMENT TASKS

For explicit temporal judgments, two tasks are particularly popular: The temporal order judgment task (TOJ; e.g., Sternberg and Knoll, 1973) and the synchrony judgment task (SJ; e.g., Schneider and Bavelier, 2003). The former task asks of a participant “which came first” (or some variant) whereas the latter asks “were they simultaneous?” Another somewhat less utilized variant, the ternary order or SJ3 task (e.g., Ulrich, 1987), offers three response categories: “first,” “simultaneous,” and “second.” In all cases trial-by-trial data can be summarized via meaningful model parameters when an appropriate observer model is fitted. These tasks (and two further tasks which will be discussed shortly) are schematized in **Figure 1**. The most commonly derived parameter, the point of subjective simultaneity (or PSS), captures any bias to report one stimulus as having come earlier than the other.

For the temporal order judgment task, under typical observer models (e.g., Gibbon and Rutschmann, 1969) the PSS estimate can be inferred to represent a combination of, first, a difference in sensory delays for the two signals and, second, any decision-level bias in interpreting relative arrival times. Distinguishing these contributions is not generally possible, which raises interpretative issues, particularly if decision-level biases might reasonably be expected to change across experimental conditions. For example, “prior entry” (Titchener, 1908) describes an experimental finding wherein attended events are thought to be perceived more rapidly than unattended ones (see Spence and Parise, 2010, for review). Prior entry has often been assessed using TOJs, with two stimuli originating from different positions and/or sensory modalities, and attention directed preferentially toward one of the two

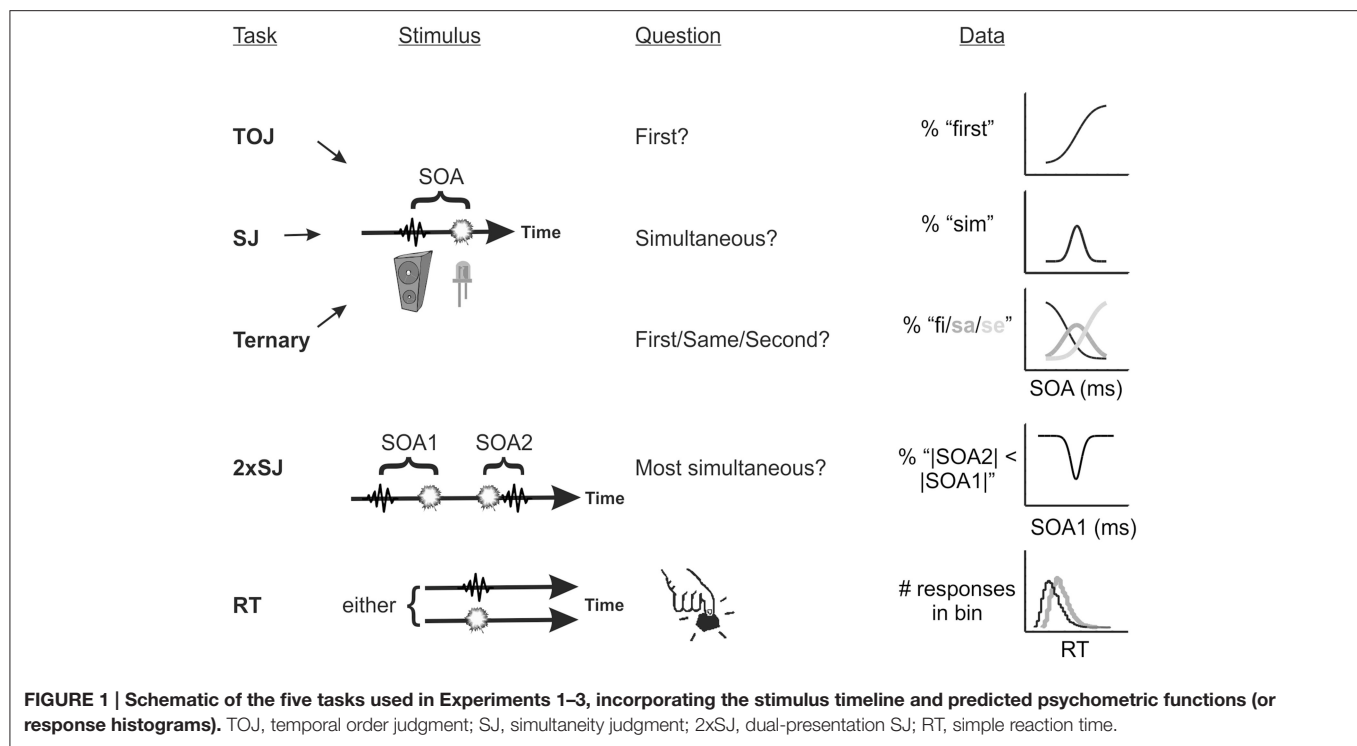
events. The demand characteristic, to attend preferentially to one of the two stimulus origins, has the potential to place that particular answer firmly in mind, which might bias responses at the decision level (Shore et al., 2001; Spence et al., 2001).

The simultaneity judgment and ternary order judgment tasks can also be used to recover a PSS, and in some cases, such as the prior entry effect, these tasks might be more appropriate in order to make the question less leading.<sup>1</sup> However, rather than a PSS, these tasks most naturally recover two boundaries around subjective simultaneity, reflecting points where judgments change from “A precedes B” to “simultaneous,” and from “simultaneous” to “A follows B” (Yarrow et al., 2011). Hence it is quite common to observe a plateau in the psychometric function, with simultaneity reported ubiquitously across several SOAs (for some examples, see García-Pérez and Alcalá-Quintana, 2012a, **Figures 6–8**; Yarrow et al., 2015, **Figure 4**). Inferring a single PSS from such data requires additional assumptions (e.g., whether the threshold for perceiving/judging simultaneity is the same when A follows B as when B follows A) which may be problematic, as there is no current consensus regarding the correct observer model for data in this form (e.g., Ulrich, 1987; Schneider and Bavelier, 2003; Yarrow et al., 2011; García-Pérez and Alcalá-Quintana, 2012a,b). In a sense, SJ and SJ3 tasks provide a temporal window within which the PSS lies, rather than a single point estimate.

## OBSERVER MODELS FOR CHARACTERIZING TEMPORAL JUDGEMENTS

So far we have made reference to observer models without specifying exactly what this means. In this paper we will use observer models derived from signal detection theory (SDT; Green and Swets, 1966; Macmillan and Creelman, 2005). Detection-theoretic approaches to temporal judgments are well-established (e.g., Baron, 1969; Gibbon and Rutschmann, 1969; Sternberg and Knoll, 1973; Allan, 1975; Ulrich, 1987; Schneider and Bavelier, 2003; Yarrow et al., 2011, 2013; García-Pérez and Alcalá-Quintana, 2012a,b). Models of this type generally assume that observers are accessing a (noisy) encoding of the *difference in arrival times* ( $\Delta t$ ) between two signals (somewhere in the brain) and using this quantity to make a decision. The key source of noise in these decisions is variability in terms of the latency with which signals arrive at a decisional mechanism, with each

<sup>1</sup>This assertion clearly needs to be examined case by case. For example, consider the literature on temporal recalibration (Fujisaki et al., 2004; Vroomen et al., 2004). This effect, presumed to be a form of adaptation, is revealed when participants are repeatedly exposed (i.e., adapted) to a non-zero asynchrony between different kinds of event (e.g., beeps that consistently lag after flashes). In this situation their PSS has been shown to change; participants now respond as though the relationship to which they have been adapted appears more synchronous than it seemed prior to the adaptation. However, if we repeatedly expose our participants to a particular asynchrony, it is not difficult to imagine that they might come to form a belief that this relationship is important, biasing their interpretation of any subsequently presented/judged asynchronies when forced to categorize them as simultaneous or not. Note, however, that some recent evidence suggests that temporal recalibration is not entirely the result of a decision-level bias (Roseboom et al., 2015).



signal contributing its (additive) variability to the distribution of encoded differences in arrival times across trials. This kind of model has been referred to as a general independent-channels model (Sternberg and Knoll, 1973) or a general-threshold model (Ulrich, 1987).

Specific variants of this general model vary mostly in terms of how many additional layers of complexity are included. For example, the simplest way to conceive of a temporal order judgment (TOJ) is that there is a single criterion used to divide the observed  $\Delta t$  into two possible order responses (Gibbon and Rutschmann, 1969). If the difference in arrival times falls below this criterion, event A is judged as having happened first; otherwise it is judged second. If, during an experiment, two stimuli are presented repeatedly but at varying physical stimulus onset asynchronies (SOAs), the model predicts a smooth function relating the SOA to the proportion of times one of the two orders is selected. The shape of this function reflects the form of latency noise, being, for example, a cumulative Gaussian if Gaussian latency noise is assumed (Baron, 1969).

Variants of these models can make predictions about other common temporal judgments in addition to the TOJ, such as the SJ3 task (before/same/after), considered in detail by Allan (1975) and Ulrich (1987), and the SJ task, considered for example by Schneider and Bavelier (2003) and by Yarrow et al. (2011, 2013). In these tasks the internal response  $\Delta t$  must be divided into three regions, rather than two (in order to demarcate “same” from “before” and “after”). This means there are two decision criteria, not one. In a variant of this kind of model, some authors (e.g., Venables, 1960; García-Pérez and Alcalá-Quintana, 2012a,b) consider that there might also be a zone near zero where no differentiation of timing is possible and observers must

guess. This inclusion, of a “guessing zone,” is a departure from classic SDT, which avoids the notion of a hard threshold. Instead, classic SDT presumes that encoded values are always recoverable. Another feature that can vary between models is the form of assumed latency noise (for example, exponential rather than Gaussian arrival time distributions can be assumed; García-Pérez and Alcalá-Quintana, 2012a,b).

## A ROVING DUAL-PRESENTATION SJ TASK

In this paper we consider a variant of the popular SJ task, which we refer to as a (roving) 2xSJ task. This task has a close structural similarity to some recent approaches in visual psychophysics (Morgan et al., 2013, 2015; Jogan and Stocker, 2014; García-Pérez and Peli, 2014). Although roving 2xSJ designs have occasionally been used in the literature on relative timing, their results have not been interpreted using formal observer models, something which we undertake here.

A note on our terminology seems appropriate at this point. The task we discuss here might reasonably be described as a two-alternative forced choice SJ task. However, taken literally, many tasks can be considered “two-alternative forced choice,” and indeed this description is sometimes applied to SJs and TOJs. Strictly speaking, in the tradition of signal detection theory, 2AFC has additional connotations. Specifically, it implies the presentation of two different exemplars on each trial, between which an observer must discriminate. Hence, in the context of temporal perception, a 2AFC simultaneity judgment would typically involve the presentation of one simultaneous pair of stimuli, and one non-simultaneous pair, in a random



sequential order (sometimes referred to as a 2IFC; two interval forced choice) with the requirement to select the synchronous (or, alternatively, the asynchronous) pair. However, the 2AFC designation is inherently ambiguous (regarding whether there are two presentations, or two possible choices) and has not always been used in a manner consistent with the SDT tradition. For this reason, we adopt the clearer “dual-presentation” terminology here.

What observer model might apply to this task? Under the simplest account (c.f. Baron, 1969) each presentation of a pair of stimuli offset by a fixed physical temporal extent would generate a subjective difference in arrival times, and these subjective differences would be variable across successive trials, generating a noisy Gaussian distribution of subjective arrival times,  $\Delta t$ . With two such pairs forming a trial, the observer's task is to compare the absolute subjective differences associated with each pair, to determine which is most simultaneous. Hence the decision variable is the difference in (unsigned) differences in subjective arrival time. To this value a criterion is applied (zero for an unbiased observer) and the observer concludes that the first or second pair is most simultaneous, depending on whether the decision variable falls above or below this criterion. This model and decision process is a special case of one described by García-Pérez and Peli (2014), but is applied here to the temporal rather than the spatial domain. If one pair is always simultaneous (the *standard*) and the other is varied in SOA (the *test*), and order is randomized, the psychometric function (plotting proportion of trials wherein the standard is judged as most synchronous, or equivalently where the test is judged more asynchronous, against the test SOA) is U shaped, with a minimum at the point of subjective simultaneity (see **Figure 2A**).

In the course of generating data via the 2xSJ described to this point, the experimenter must present a synchronous target on every trial. Unfortunately this seems perfectly designed to train the observer to recognize what a truly synchronous relationship feels like, which is not helpful when we are seeking their natural (potentially non-zero) PSS, or assessing whether this varies with some experimental manipulation. Fortunately a fairly straightforward solution to this problem exists: As experimenters, we should not exclusively use synchronous standards. If neither pair is guaranteed to be synchronous, it becomes difficult for a participant to learn what synchrony is, independent from the perception of synchrony.

At first glance this procedure seems wasteful, as trials without a synchronous standard do not contribute to the psychometric function shown in **Figure 2A**. Must they be discarded? The answer is no. The observer model also makes predictions about how often a  $-20$  ms SOA standard should appear more synchronous than a  $60$  ms SOA test, or any other combination. The only complication is that we must now move from a single SOA vs. proportion correct function to a set of functions, one for each standard (see **Figures 2A–F**). However, just as with a synchronous standard, each predicted function retains a minimum at the PSS (because a test presented exactly at the PSS will always be more likely than any other test value to be judged as most synchronous, regardless of what value it is being compared to). When the standard is zero, the slopes of

the psychometric function are determined largely by sensory noise. This too remains the case for functions predicted for non-zero standards (note the parallel slopes in **Figure 2E**). In fact, an observer model with just three parameters (one for PSS, one for sensory noise, and one capturing any preference to favor the first pair over the second or vice versa when reporting greater synchrony; see **Figure 3** for further explanation of this interval bias) predicts an entire family of psychometric functions. These functions vary in a yoked manner as the model's parameters are adjusted, so best-fitting parameters can be obtained by fitting data from all standards/tests at once.

In the wider literature, “roving” dual-presentation designs like this are sometimes presented as a means to minimize the influence of non-sensory biases, while still measuring a perceptual quality (e.g., Morgan et al., 2013). In brief, these tasks allow the experimenter to apply a contextual manipulation in *both* presentations of a trial, making it less plausible that the manipulation will directly bias the judgment (for example, by nudging a decision criterion in one or other direction). Revisiting the prior-entry example, if an observer must attend the same modality on both presentations, no simple rule such as “pick the stimulus I am attending to” presents itself. However, it is generally possible to conceive of more complex biasing strategies. For reasons of concision, we do not make bias minimization a major focus of our discussion here.

## THE PRESENT EXPERIMENTS

Our goal here is to demonstrate the use of the roving 2xSJ task as a measure of the PSS. We initially present data from two observers who engaged in a substantial number of audiovisual trials using this task. We use their data to illustrate the fitting procedure (**Figures 2, 3**), and to evaluate whether simple observer models are plausible. We then present two sets of data each collected from 24 participants, with a much smaller number of trials per participant (which is more representative of typical timing experiments). We additionally collect data in several other tasks, to test whether PSS estimates from different procedures are comparable, and assess correlations across subjects for derived parameters representing both bias (e.g., PSS) and (inverse) precision (i.e., the standard deviation of inferred latency distributions). In Experiment 2, we extend these analyses to judgments involving different combinations of visual, auditory and tactile stimuli. Finally, in Experiment 3, we attempt to replicate a classic effect in the relative-timing literature—endogenous prior entry, using both 2xSJ and TOJ tasks.

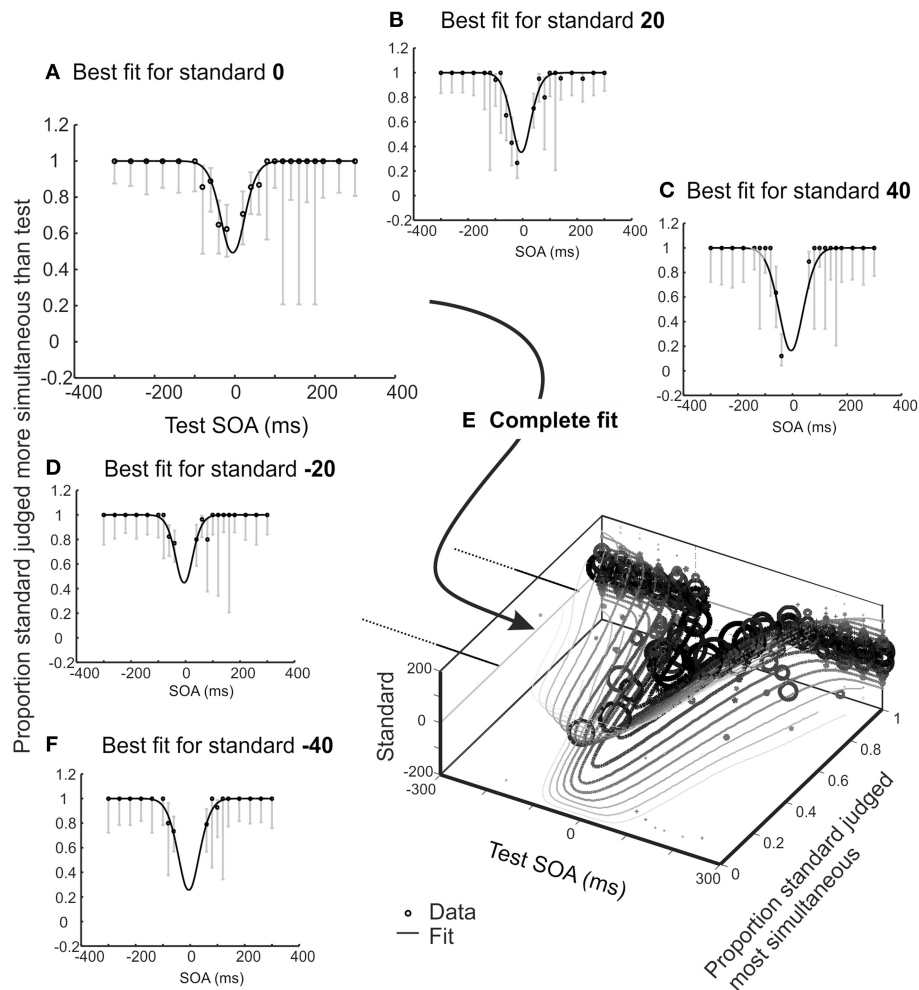
## EXPERIMENTS 1A, 1B, AND 1C

### Methods

#### Participants

Participants were recruited (and provided informed consent) according to procedures approved by the City University London Department of Psychology Ethics Committee.

There were two participants in Experiment 1a, one male author (KY) and one female author (SM), initially aged 38 and 24 respectively. Observer KY was highly experienced with



**FIGURE 2 | Illustrative results for one experienced psychophysical observer who completed 15 blocks (2280 trials) of an audio-visual 2xSJ task.** One SOA was selected from a wide range using the method of constant stimuli, the second from a range centered closer to synchrony, using an adaptive method for stimulus selection (providing the roving element; see methods for full details). In order to aid visualization, trials were extracted sequentially in sets associated with a particular standard. In this figure, data and model predictions are averaged for the two possible presentation orders, so the uninformative data points where standard and test are identical have been removed. **(A–D,F)** Observer model fit alongside several data subsets consisting of trials associated with standards ranging from  $-40$  to  $+40$  ms. Error bars show 95% Wilson score binomial confidence intervals, and provide an indication of the number of trials at each stimulus level. **(E)** All data displayed together with the overall model fit. Size of data points correlates with the number of contributing trials.

relative timing tasks, while observer SM had relatively limited psychophysical experience.

An opportunity sample of 27 naïve participants was tested in Experiment 1b. Of these, three were excluded from analysis (see Data Analysis, below) to yield a sample size of 24 (mean age = 24.3, range 18–51, six male). Another opportunity sample of 24 naïve participants was tested in Experiment 1c (mean age = 29.8, range 19–52, 11 male).

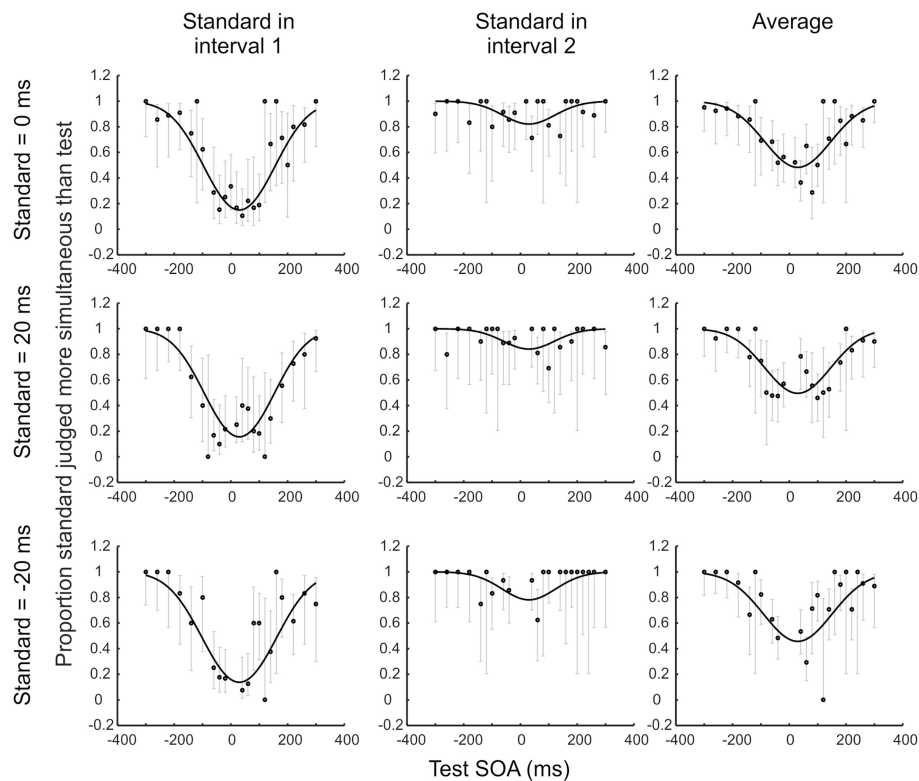
### Apparatus and Stimuli

A PC connected to a 20-inch CRT monitor was interfaced with one or more National Instruments A/D cards (DAQCard-6715; DAQPad-6015; X-series PCIe-6323) via a bespoke visual c++ program in order to generate signals and (for the RT task) record responses. Signals (beeps and flashes) were generated at 44100 Hz, and were 10 ms long, with onset and offset slightly

smoothed using a Hanning window across the first and last millisecond of the stimulus. The red visual LED signal was otherwise continuous ( $\sim 60$  mcd point source) while the sound was a 1000 Hz sine wave. The LED was placed immediately in front of the center of the monitor, at a distance of  $\sim 57$  cm from the eyes, so the light subtended a visual angle of  $\sim 0.5^\circ$ . Beeps were presented from a speaker located immediately to the left of the monitor ( $\sim 30^\circ$  from the LED) at a comfortable suprathreshold intensity. Responses were recorded via keyboard (for the temporal judgment tasks) or a digital button (for the RT task). Participants fixated the LED during stimulus presentation.

### Design and Procedure

In Experiment 1a both participants completed a 2xSJ task, followed by a TOJ task, followed by a combined SJ and 2xSJ task, with each task typically completed across several days. There



**FIGURE 3 | Illustrative results for one novice psychophysical observer who completed 15 blocks (2280 trials) of an audio-visual 2xSJ task.** Observer model fits are shown alongside several data subsets, consisting of trials associated with standards ranging from  $-20$  to  $+20$  ms. Data are plotted separately based on the order of stimulus presentation (and also in averaged format) across the columns of the figure, to show how an interval bias is captured by the model. For example, imagine that the participant is biased so that, given the question “which was most simultaneous,” they will only pick interval 1 if its duration is at least 30 ms less than that of interval 2 (whereas an unbiased observer would pick interval 1 if its duration was shorter by any margin). If the standard is in interval 1, the chance of picking the standard as more simultaneous is decreased relative to an unbiased observer. By contrast, if the standard is in interval 2, the chance of picking the standard as more simultaneous is increased relative to the unbiased observer. Averaging data from the two possible interval orders simplifies presentation, but the underlying model that is fit to data should ideally still include a parameter to capture any interval bias, because assuming no bias in the presence of an actual bias will lead to an overestimate of sensory noise (although the PSS will still be recovered adequately).

was a substantial separation between completing the tasks. In Experiment 1b participants completed three tasks (2xSJ, TOJ, and simple RT) in a single session, with task order counterbalanced across participants. In Experiment 1c participants completed a combined SJ and 2xSJ task. Participants completed 15–20 practice trials before each task, but received no feedback about the correctness of their responses at any time to avoid biasing subjective timing judgments.

In the 2xSJ task, participants in Experiment 1a each completed 15 blocks of 152 trials (i.e., 2280 trials in total), while those in Experiment 1b and 1c completed a single block of 152 trials. There were two flash-beep pairs, and thus two SOA values on each trial. One of the two pairs was selected at random via the method of constant stimuli from the following 19 SOAs (where positive = beep follows flash):  $-300$ ,  $-260$ ,  $-220$ ,  $-180$ ,  $-140$ ,  $-100$ ,  $-60$ ,  $-40$ ,  $-20$ ,  $0$ ,  $20$ ,  $40$ ,  $60$ ,  $100$ ,  $140$ ,  $180$ ,  $220$ ,  $260$ ,  $300$  ms. Each SOA was selected and presented four times in the first interval, and four times in the second, for a total of 152 trials. The SOA for the other flash-beep pair was selected at random using an adaptive method, such that it would generally be near

the PSS, but not always zero (so participants could not infer/learn true synchrony across the experiment). To achieve this, the SOA was drawn from a discrete probability distribution with steps of 20 ms. The initial shape of the distribution was uniform, spanning  $-60$  to  $+60$  ms. However, the distribution had the potential to include values from  $-300$  to  $+300$  ms, and it was updated after each trial based on which of the two presented asynchronies had been selected as more simultaneous. Specifically, the distribution was adjusted so that selection likelihood was increased for all asynchronies  $\pm 40$  ms from the asynchrony selected as most simultaneous on that trial. This approach is loosely based on the generalized Pólya urn model (Rosenberger and Grill, 1997) proposed for efficient sampling for temporal order judgments.

In summary—the 2xSJ task involved participants being presented with two flash-beep pairs, with neither SOA being predictable. The pairs were separated by a uniform random 1000–2000 ms interval and participants were required to respond to the question “Which pair was more simultaneous?” using arrow keys on the keyboard. This triggered the next stimulus presentation after 1000–2000 ms. Participants were also given the option to

cancel a trial due to inattention, in which case it was repeated at the end of the block.

In the combined 2xSJ and SJ task, stimulus selection and presentation was identical to the 2xSJ task with the following exceptions. Participants were required to make a simultaneity judgment after each stimulus pair, followed by a most simultaneous judgment after every two stimulus pairs. In Experiment 1a, observers completed eight blocks (1216 2xSJ judgments and 2432 SJ judgments). In Experiment 1c, they completed a single block (152 2xSJs and 304 SJs). Each response triggered the next stimulus presentation after 1000–1400 ms, except for the second SJ response, which was followed by the 2xSJ question after 500 ms.

In the TOJ task, for Experiment 1a participants completed 23 blocks of 100 trials each (2300 trials in total). SOA values were selected at random on each trial from an adaptive probability distribution. This distribution was uniform at the start of each block, containing SOAs from  $-225$  to  $+225$  ms in 5 ms increments, but was updated after each accepted trial according to the generalized Pólya urn model (Rosenberger and Grill, 1997;  $k = 32$ ) which attempts to generate test values that sample the full psychometric function efficiently. Distributions could expand to include SOAs from  $-450$  to  $+450$  ms. For Experiment 1b participants completed a single block of 100 trials. In this case the adaptive distribution initially contained SOAs from  $-140$  to  $+140$  ms in 20 ms increments, and could expand to include values from  $-300$  to  $+300$  ms via the generalized Pólya urn method ( $k = 8$ ). In both experiments, after each presentation, participants responded to the question “Which came first (beep or flash)?” using arrow keys on the keyboard. They also had an option to cancel and repeat the trial later. The midpoint of the next flash-beep pair came 1000–2000 ms after each response.

In the RT task (Experiment 1b only) each trial consisted of either a flash or beep, with 50 trials of each type intermixed in random order within a block of 100 trials. Participants responded to each stimulus as quickly as possible using a digital button, following a 1000–2000 ms uniform random response-stimulus interval.

## Data Analysis

For all temporal judgment tasks (2xSJ, SJ, and TOJ) Matlab (The MathWorks Inc.) was used to find maximum-likelihood fits to data (assuming binomially distributed data) with both a null (guessing) model and also a simple independent-channels observer model. The Nelder-Mead simplex algorithm was used to find the best fit. To avoid problems with local maxima, simplex searches were initiated from the factorial combination of several positions per parameter (i.e., a grid search seeded a set of simplex searches). Observer models incorporated a fixed 1% keyboard error/lapse rate, to model occasional errors without increasing parametric complexity (and also simplify the calculation of log likelihood).

Trial-by-trial data for the 2xSJ task consisted of pairs of SOAs, plus a judgment about which pair was most simultaneous. For all trials, the SOA nearest 0 was defined as the standard, as this designation will facilitate a compact presentation of data. All trials where the standard was 0 (i.e., simultaneous) were extracted

first, divided according to whether the standard was in the first or the second interval. Remaining trials were then examined, extracting all cases where the standard had an SOA of  $-20$ . This was repeated, looking for standards of  $+20$ , then  $-40$ , then  $+40$ , and so on until all trials with at least one SOA of less than  $\pm 200$  ms had been extracted. For each standard SOA occurring in each interval, data were plotted to show the proportion of times that the standard was judged as *more simultaneous* than the test (see **Figure 2** for examples following averaging across the two presentation intervals, and **Figure 3** for examples separated by presentation interval). These functions would be expected to have a minimum near the point of subjective simultaneity.

The tested observer model assumes each stimulus is accompanied by Gaussian noise that will affect its central arrival latency, and that the two stimuli comprising an AV pair may be delayed by neural processing to different extents (generating a non-zero PSS). For each stimulus pair in the 2xSJ, the noisy and delayed signals therefore arrive centrally with latency differences ( $\Delta t$ ) that form a Normal distribution of internal responses for any physical SOA:

$$\Delta t_{\text{standard}} \sim N(\text{SOA}_{\text{standard}} + \mu, \sigma^2) \quad (1)$$

$$\Delta t_{\text{test}} \sim N(\text{SOA}_{\text{test}} + \mu, \sigma^2) \quad (2)$$

Where  $\text{SOA}_{\text{standard}}$  is the standard SOA (i.e., the stimulus pair that is closest to synchrony),  $\text{SOA}_{\text{test}}$  is the test SOA (i.e., the other stimulus pair),  $\mu$  captures any asynchrony specific to the observer (i.e., the PSS) and  $\sigma^2$  is the variance contributed by each  $\Delta t$  distribution.

For the subsequent decision, the model assumes that  $\Delta t$  in each pair is converted to an absolute score and that the larger absolute score is judged as less simultaneous. The probability of selecting the standard is therefore:

$$\begin{aligned} \Pr(\text{“Standard”} \mid \text{SOA}_{\text{standard}}, \text{SOA}_{\text{test}}) \\ = \Pr(|\Delta t_{\text{standard}}| < |\Delta t_{\text{test}}|) \end{aligned} \quad (3)$$

Which, can be written:

$$\Pr(\text{“Standard”} \mid \text{SOA}_{\text{standard}}, \text{SOA}_{\text{test}}) = \Pr\left(\frac{\Delta t_{\text{standard}}^2}{\Delta t_{\text{test}}^2} < 1\right) \quad (4)$$

Note that  $\frac{\Delta t_{\text{standard}}^2}{\Delta t_{\text{test}}^2}$  is a random variable with a doubly non-central  $F$ -distribution. Its numerator’s non-centrality parameter is  $(\mu + \text{SOA}_t)^2/\sigma^2$ , its denominator’s non-centrality parameter is  $(\mu + \text{SOA}_n)^2/\sigma^2$ , and both numerator and denominator have one degree of freedom (Morgan et al., 2015). In our Matlab code (available at <http://www.hexicon.co.uk/Kielan/>) we made use of a saddle-point approximation to the doubly non-central  $F$  cumulative distribution function (Butler and Paoletta, 2002; Paoletta, 2007).

So far, the model simulates an unbiased observer, in the sense of having no preference for the first interval over the second or vice versa. However, Equation (4) can be modified to incorporate an interval bias. Let  $< \text{SOA}_{\text{standard}}, \text{SOA}_{\text{test}} >$  denote that the



standard was presented first and let  $< SOA_{test}, SOA_{standard} >$  denote that the test was presented first. Then:

$$\begin{aligned} \Pr(\text{"Standard"} | < SOA_{standard}, SOA_{test} >) \\ = \Pr\left(\frac{\Delta t_{standard}^2}{\Delta t_{test}^2} < \beta\right) \end{aligned} \quad (5)$$

$$\begin{aligned} \Pr(\text{"Standard"} | < SOA_{test}, SOA_{standard} >) \\ = \Pr\left(\frac{\Delta t_{standard}^2}{\Delta t_{test}^2} < \frac{1}{\beta}\right) \end{aligned} \quad (6)$$

Under this scheme, the interval bias is proportional, capturing a decision rule in which the observer selects interval 1 ( $I_1$ ) when  $|I_1| < c|I_2|$ , and  $c$  equals  $\beta^{1/2}$ . In words, the biased observer is selecting interval 1 when its duration is less than, e.g., one and a half times the duration of interval 2. This bias can be contrasted with the *constant* bias modeled by García-Pérez and Peli (2014) and presented in their Equations (5), (A5), and (A6) (pages 1676 and 1692). Under this scheme, the observer selects interval 1 when  $|I_1| - |I_2| < c$ . In words, the biased observer is selecting interval 1 when its duration exceeds that of interval 2 by less than, e.g., 50 ms. As we had no a priori reason to favor one form of interval bias over the other, we implemented fits using both, and retained the best fit using either model for each participant in our results.

Recent work by Patten and Clifford (2015) allowed us to derive a closed-form expression for the constant-bias model.<sup>2</sup> Although our fits were obtained using the (slower to evaluate) derivations described above, we include the new derivation here for completeness, as it is now the default option in our Matlab code:

$$\begin{aligned} \Pr(\text{"Standard"} | < SOA_{standard}, SOA_{test} >) = \\ \left\{ \begin{aligned} &\frac{1}{4} \left( 2 - \operatorname{erf}\left(\frac{\Delta t_{standard} - \Delta t_{test} + c}{2\sigma}\right) \left( \operatorname{erf}\left(\frac{2\mu + \Delta t_{standard} + \Delta t_{test} - c}{2\sigma}\right) - 1 \right) - \right. \\ &\quad \left. \operatorname{erf}\left(\frac{2\mu + \Delta t_{standard} + \Delta t_{test} - c}{2\sigma}\right) + \operatorname{erf}\left(\frac{2\mu + \Delta t_{standard} + \Delta t_{test} + c}{2\sigma}\right) + \right. \\ &\quad \left. \operatorname{erf}\left(\frac{-\Delta t_{standard} + \Delta t_{test} + c}{2\sigma}\right) \left( \operatorname{erf}\left(\frac{2\mu + \Delta t_{standard} + \Delta t_{test} + c}{2\sigma}\right) + 1 \right) \right), \quad \text{if } c \leq 0 \\ &\frac{1}{4} \left( 4 + \left( 1 - \operatorname{erf}\left(\frac{-\Delta t_{standard} + \Delta t_{test} + c}{2\sigma}\right) \right) \left( \left( 1 - \operatorname{erf}\left(\frac{2\mu + \Delta t_{standard} + \Delta t_{test} - c}{2\sigma}\right) \right) - 2 \right) - \right. \\ &\quad \left. \left( 1 - \operatorname{erf}\left(\frac{\Delta t_{standard} - \Delta t_{test} + c}{2\sigma}\right) \right) \left( 1 - \operatorname{erf}\left(\frac{2\mu + \Delta t_{standard} + \Delta t_{test} + c}{2\sigma}\right) \right) \right), \quad \text{if } c > 0 \end{aligned} \right. \end{aligned} \quad (7)$$

Where erf denotes the error function:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (8)$$

For our null model, we assumed participants might simply guess, but be biased to choose one or the other interval more often (a one-parameter model). This would lead to deviations from a 0.5 prediction at all test stimulus levels, depending on the interval in which the standard was presented.

To test participant compliance (for exclusion purposes) and the appropriateness of our observer model, we considered two metrics based on deviance of model fit (defined here as  $-2 \times$

the shortfall in log-likelihood relative to a saturated model). We first assessed whether the more complex (i.e., higher-parameter) observer model provided a significantly better fit than the guessing model. Asymptotically, for nested models the improvement in deviance expected by chance approximates a chi-squared distribution with d.f. equalling the difference in model parameters. We used this result to assess whether the more complex model provided a significantly better fit than its less complex counterpart (at one-tailed  $p < 0.05$ ). If not, there was little evidence that the participant was not simply guessing. No participants in Experiment 1 needed to be replaced on this basis.

For Experiment 1a, we also considered whether our observer model represented a reasonable approximation of the complete psychophysical process. For this purpose we turned to Monte-Carlo simulation. We fed the stimulus values each participant received across the entire experiment into the best-fitting observer model to generate a simulated set of responses. These were then maximum-likelihood fitted with the model, in order to establish a deviance score for the best-fitting model when that model had in fact generated the data. We repeated this operation 1000 times to create a distribution of expected deviances if the model were correct. Finally, we compared the deviance of the model when fitted to *real* data against the simulated distribution of expected deviances, to assess whether the model could be rejected as a full characterization of what observers were doing (two-tailed  $p < 0.05$ ; c.f. Wichmann and Hill, 2001).

For the TOJ task, the same basic observer model assumptions (i.e., Gaussian latency noise) along with the simplest conceivable decision rule (i.e., select order A when  $\Delta t$  is below a decision criterion, otherwise select order B)<sup>3</sup> predict a cumulative Gaussian psychometric function, where  $\mu$  is the PSS and  $\sigma$  is

the standard deviation of the  $\Delta t$  distribution. The corresponding guessing model has only a single free parameter (a bias for one order over the other, predicting a horizontal line crossing the y axis somewhere between 0 and 1) and is nested relative to the observer model. Hence we assessed whether the observer model provided a significantly better fit than the guessing model (at one-tailed  $p < 0.05$ ) by comparing the change in deviance to a chi-squared distribution with one degree of freedom. In Experiment 1b, three participants were rejected because their performance did not provide evidence to reject the guessing model (i.e., performance was not significantly different from chance). For Experiment 1a we also assessed whether the

<sup>2</sup>Thanks to Kai Schreiber for help with the derivation.

<sup>3</sup>Described as a "deterministic" decision rule by Sternberg and Knoll (1973).

observer model represented a reasonable approximation of the complete psychophysical process, using the resampling method described above for the 2xSJ task.

For the SJ task, the observer must partition the decision space in a slightly more complex manner than for the TOJ, using two decision criteria in order to report simultaneity only when  $\Delta t$  falls between them. This decision rule predicts a psychometric function that is the difference of two cumulative Gaussians, with their means at the positions of the two decision criteria and their (shared) standard deviation being that of the  $\Delta t$  distribution (Schneider and Bavelier, 2003). To derive a single PSS, a further assumption of some kind is required (for example that the decision boundaries are placed equidistant from subjective zero). Hence we generally prefer to report the two criteria themselves (Experiment 1a) but adopt the equidistance assumption for the purpose of generating a PSS value for correlation analyses (Experiment 1c).

This three-parameter SJ model produces a symmetric psychometric function, but asymmetries are sometimes observed in SJ data (e.g., Yarrow et al., 2011; García-Pérez and Alcalá-Quintana, 2012a,b). If we retain the assumption of Gaussian latency noise, one way to model such an asymmetry is to assume that the two decision criteria might also contribute (independent) noise to the decision (Ulrich, 1987). If the positions of the two decision criteria are considered Gaussian random variables, the resulting psychometric function is the difference of two cumulative Gaussians, but with separate  $\sigma$  parameters (hence a four parameter model; Yarrow et al., 2011).<sup>4</sup> In our analyses, we fitted both three and four-parameter variants of SJ models. We first asked whether the three-parameter model provided a significantly better fit than a two parameter cumulative Gaussian [deviance improvement,  $\chi^2_{(1)} < 0.05$ ]. We chose this model in place of a simpler guessing model as it can capture both guessing, and cases where the range of stimuli is sufficient to capture the decision boundary on one, but not both, sides of zero. In Experiment 1c, no participants were excluded on this basis. We then asked whether the four-parameter SJ model provided a significantly better fit than the three-parameter version. If so, we used parameters from the four-parameter fit, taking the lower of the two  $\sigma$  parameters as our measure of precision (as, under this model, it represents an upper bound on the standard deviation of the  $\Delta t$  distribution). This model was used for 7/24 participants.

For simple RT data from Experiment 1b, we first excluded trials with RTs  $< 100$  ms or  $> \text{mean} + (2.5 \times \text{SD})$  ms. The “PSS” was then calculated as the difference between the trimmed mean RT to light and the trimmed mean RT to sound. This gives a measure of the head start sound seems to have relative to light, which can then be compared with the PSS in temporal judgment tasks (Gibbon and Rutschmann, 1969). The starting points for a comparable measure of sensory noise were variances of response times for flashes and beeps in trimmed trials. To generate a measure equivalent to the one obtained in the temporal judgment

tasks (i.e., the standard deviation of the  $\Delta t$  distribution) variances for sound RTs and light RTs were summed then square rooted.

For a subset of temporal judgments, we derived bootstrap confidence intervals on best-fitting model parameters. Bootstrap procedures were non-parametric and based on 1999 resamples, using the bias-corrected and accelerated (BCa) method (Efron and Tibshirani, 1994). When considering inferential statistics at the group level, we observed numerous violations of parametric assumptions (e.g., non-normality in difference distributions, Shapiro-Wilks  $p > 0.05$ ). We therefore generally used paired-sample permutation  $t$ -tests when assessing differences (based on 10,000 permutations) with a  $t_{\max}$  correction for multiple comparisons when three or more conditions were compared. To assess associations, we used the Pearson correlation coefficient, but when there was evidence of non-normality in either of the contributing distributions (Shapiro-Wilks  $p < 0.05$ ) we assessed significance via bootstrap confidence intervals. Unless otherwise noted, we used an alpha value of 0.05 and two-tailed tests.

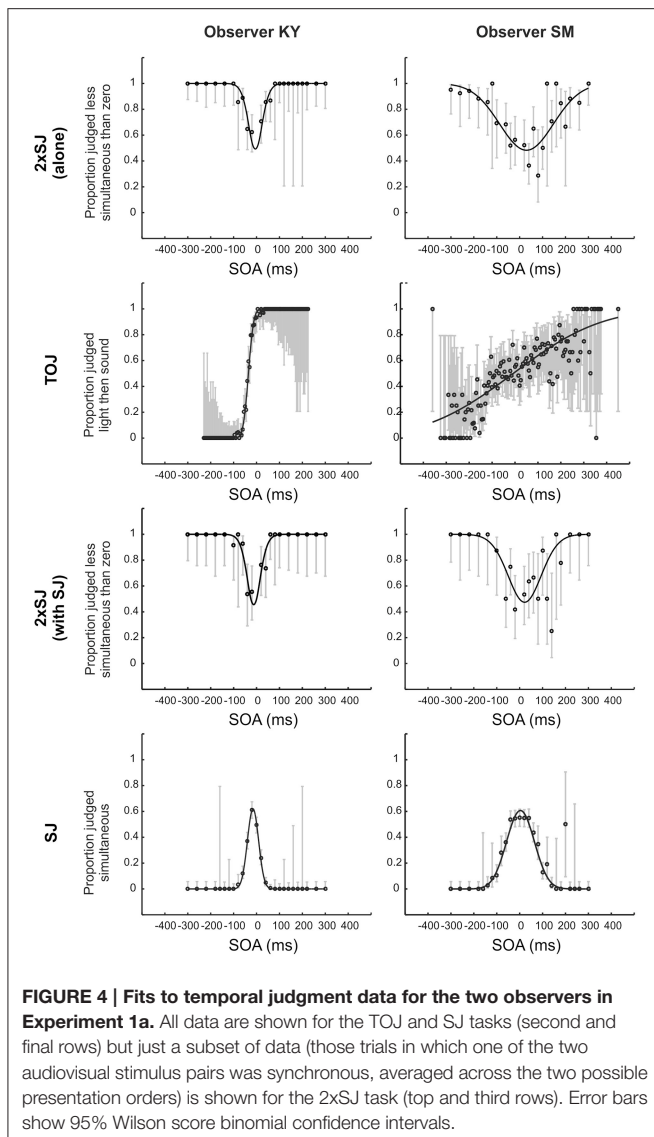
## Results

**Figure 4** shows raw data for the TOJ and SJ tasks and a subset of the raw data (specifically that with a zero-SOA standard) for the 2xSJ tasks, alongside best-fitting model predictions for both observers in Experiment 1a. By eye, the fits look fairly good. For observer SM, the 2xSJ task simulation suggested that our simple observer model could plausibly be the generating model for the data when the task was performed alongside the SJ, but not when performed alone, as the deviance of the best-fitting model differed significantly from its expected value in the latter case ( $p = 0.048$ ). The observer model described data well for the TOJ task ( $p > 0.05$ ) but not for the SJ task ( $p = 0.03$ ). For observer KY, for the 2xSJ task deviance of the best-fitting model was significantly greater than predicted if our simple observer model were a complete generating model, both when the task was performed alone and alongside the SJ ( $ps < 0.001$ ). However, for the TOJ and SJ tasks, the observer model was plausible ( $p > 0.05$ ).

Parameters derived from these fits are presented in **Table 1**. PSS values were close to zero for both observers, while latency noise was considerably lower for experienced participant KY than for the more novice participant SM. For SM, noise was much lower in the 2xSJ task than in the TOJ task, despite the model-based equivalence of the two measures (which both correspond to the standard deviation of the difference in arrival times for auditory and visual signals,  $\sigma$ ). Confidence intervals indicate this is unlikely to be a chance result. However, noise was very similar between SJ and 2xSJ tasks. It seems SM exhibited a learning effect, as noise was lower for her second run on the 2xSJ task despite the additional requirements of the concurrent SJ task. PSS, however, was similar on both runs. The PSS from the 2xSJ was somewhat higher than that derived from the TOJ, and also than the mid-point of the two boundaries in the SJ (which was 3 ms).

For experienced observer KY, noise estimates were very similar in all tasks. As for SM, the PSS from the 2xSJ was somewhat higher than that derived from the TOJ task, and also than the mid-point of the two boundaries obtained in the SJ

<sup>4</sup>This model occasionally breaks, because differential levels of criterion noise and tight decision criteria imply that the two component cumulative Gaussians overlap. This can be resolved by turning to simulation and requiring that the decision boundaries never take an illogical order, but here we instead simply assigned zero likelihood to fits generating impossible probabilities.



(which was  $-16$  ms). For both observers, confidence intervals were non-overlapping for PSS estimates from the TOJ and 2xSJ tasks, with a more positive PSS in the 2xSJ task. The pattern was similar but slightly less clear cut for the PSS implied by the midpoint of the two decision boundaries in the SJ task. Widths of confidence intervals imply that the PSSs (and boundary estimates for the SJ) were similarly well-estimated by all tasks, whereas the 2xSJ and SJ tasks provided greater confidence regarding true values of latency noise than the TOJ, but specifically for observer SM. Finally, interval bias parameters in the 2xSJ task suggest that both observers showed a bias to favor the second interval.

Moving to the group results from Experiment 1b, **Figure 5** shows mean parameters derived from individual fits to data for the 24 participants who successfully completed the experiment. Average PSS estimates were similar for both temporal judgment tasks (TOJ and 2xSJ) and for simple RTs (all pairwise comparison  $p$ s  $> 0.05$ ) and in all cases were near zero but slightly positive (i.e., auditory RT  $<$  visual RT and simultaneity perceived when

audition trails vision), a fairly common finding in audiovisual timing (van Eijk et al., 2008). By contrast, average estimates of latency noise differed significantly across the three tasks [RT vs. TOJ,  $t_{(23)} = 6.35$ ,  $p < 0.001$ ; RT vs. 2xSJ,  $t_{(23)} = 3.49$ ,  $p = 0.005$ ; 2xSJ vs. TOJ,  $t_{(23)} = 4.79$ ,  $p < 0.001$ ]. Noise was highest in the TOJ task, lower in the 2xSJ task, and lowest in the RT task. For the 2xSJ task, 13/24 participants showed a bias to favor the second interval.

We also looked at the mean width of the 95% confidence intervals around estimates derived from the 2xSJ and TOJ tasks. Given that the 2xSJ task included more trials than the TOJ task in Experiment 1b, and would therefore be expected to provide tighter confidence intervals, for this comparison we looked at fits based only on the first 100 trials of the 2xSJ (which still gave mean estimates very similar to those shown in **Figure 5**, which were based on all 152 trials). For the PSS, confidence limits around estimates were similar for the two tasks [mean widths of 106 ms for 2xSJ vs. 132 ms for TOJ,  $t_{(23)} = 0.94$ ,  $p > 0.05$ ], while for latency noise confidence regions were significantly tighter regarding the lower estimates produced by the 2xSJ task (106 ms for 2xSJ vs. 380 ms for TOJ,  $t_{(23)} = 2.39$ ,  $p < 0.001$ ).

Importantly, Experiment 1b also provided the opportunity to see whether tasks agreed regarding individual differences in bias (PSS) and precision. **Figure 6** shows correlations across participants between equivalent measures obtained with each task. There was a significant correlation between the PSS values estimated from the 2xSJ task and those estimated from the TOJ task (bootstrap  $p < 0.05$ ), but neither correlated significantly with simple RT estimates. For measures of latency noise, correlations between RT and TOJ tasks and between TOJ and 2xSJ tasks were marginally significant (one-tailed bootstrap  $p < 0.05$ ).

The results of Experiment 1c, where a group of participants made SJs and 2xSJs concurrently, are shown in **Figure 7**. This illustrates correlations between the two tasks on both PSS and latency noise. Three participants, shown in gray, were clearly outliers in terms of their (in)ability to perform the two tasks, with very high estimates of sensory latency noise. Probably as a consequence of this, their PSS values are also extreme and outlying, suggesting that they have been poorly estimated (note the different axis scales for PSS in **Figure 7** compared to **Figure 6**). We therefore performed correlations both with and without (denoted in gray and black respectively) these outlying participants included. Correlations were significant between tasks on both measures (bootstrap  $p < 0.05$ ) with the exception of the PSS when outlying values were retained.

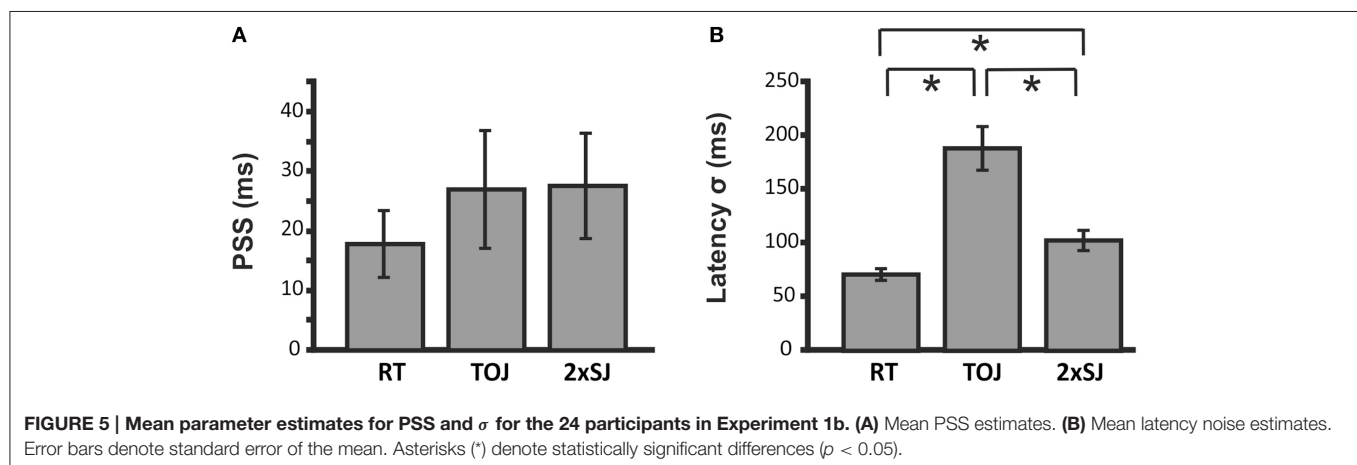
Concurrent performance of the SJ and 2xSJ tasks yielded mean parameter estimates which did not differ across tasks regardless of whether outlying participants were included in the analysis or not [with outliers: mean SJ PSS = 31 ms, mean 2xSJ PSS = 43 ms,  $t_{(23)} = 0.59$ ,  $p > 0.05$ ; mean SJ latency noise = 97 ms, mean 2xSJ latency noise = 108 ms,  $t_{(23)} = 0.91$ ,  $p > 0.05$ ; without outliers: mean SJ PSS = 24 ms, mean 2xSJ PSS = 18 ms,  $t_{(20)} = 1.18$ ,  $p > 0.05$ ; mean SJ latency noise = 76 ms, mean 2xSJ latency noise = 69 ms,  $t_{(20)} = 1.37$ ,  $p > 0.05$ ]. For the 2xSJ task, 20/24 participants showed a bias to favor the second interval (binomial  $p < 0.05$ ).

**TABLE 1 | Results of TOJ, 2xSJ, and SJ tasks from Experiment 1a, showing estimated model parameters such as point of subjective simultaneity (PSS) and Precision (latency noise;  $\sigma$ ) with accompanying 95% confidence intervals (CI) for each observer.**

Observer	Task	Parameters								
		PSS (ms)			Precision (ms)			Interval Bias		
		Low CI	PSS	High CI	Low CI	$\sigma$	High CI	Low CI	Bias	High CI
SM	2xSJ (alone)	20	30	39	68	78	86	56	64 <sup>a</sup>	78
KY		-8	-5	-2	21	24	27	1.57	1.79 <sup>b</sup>	2.05
SM	TOJ	-23	-8	9	273	307	360			
KY		-37	-35	-33	17	19	22			
SM	2xSJ (with SJ)	9	22	22	41	55	55	2	14 <sup>a</sup>	14
KY		-15	-12	-9	19	22	26	1.51	1.85 <sup>b</sup>	2.27

		Lower boundary (ms)			Precision (ms)			Upper boundary (ms)		
		Low CI	B <sub>Low</sub>	High CI	Low CI	$\sigma$	High CI	Low CI	B <sub>High</sub>	High CI
SM	SJ	-48	-42	-37	50	54	58	44	49	55
KY		-37	-34	-32	20	21	24	1	3	6

<sup>a</sup>Best fitting interval bias is constant.<sup>b</sup>Best fitting interval bias is proportional.

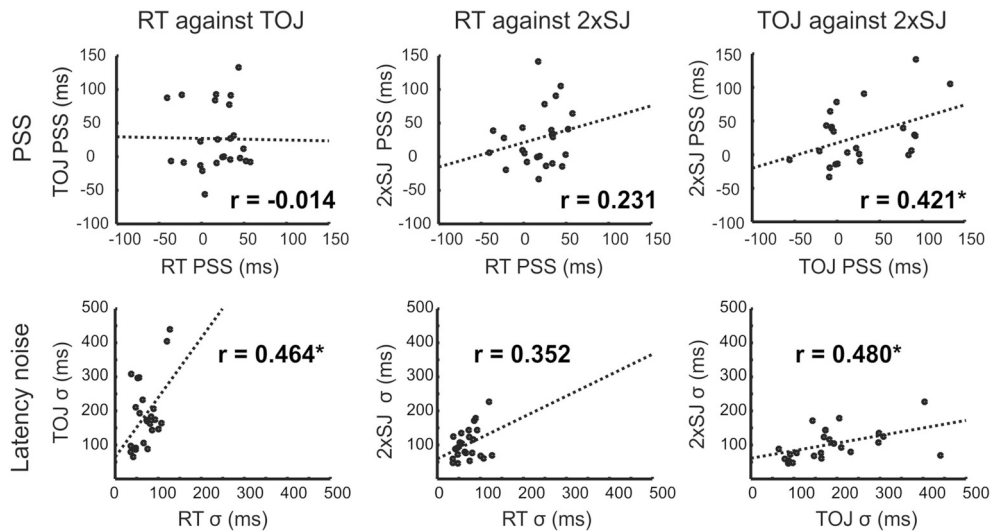
We noticed that, compared to our previous experiences recording SJs on their own, participants appeared to be applying more conservative decision criteria in the SJ task from Experiment 1c. We wondered if the presence of the additional (2xSJ) question was prompting them to be more conservative. As an informal test of this hypothesis, we retrieved a recent data set from 22 participants who completed a baseline SJ task very similar to that used here (but prior to several rather different conditions involving temporal adaptation; Yarrow et al., 2015). Stimuli were virtually identical to those employed here except that the LED flash was green, rather than red. To assess the liberal vs. conservative use of the simultaneous response, we calculated the distance between low and high decision criteria (based on the same four-parameter model fit in both data sets). Data met the assumptions of an independent-samples  $t$ -test, which revealed that participants placed their decision criteria closer together in the current data set incorporating a concurrent 2xSJ question than in our previous data set with only an SJ question [mean

distance with SJ task alone = 440 ms, mean distance with SJ and 2xSJ = 260 ms,  $t_{(44)} = 4.47$ ,  $p < 0.001$ ].

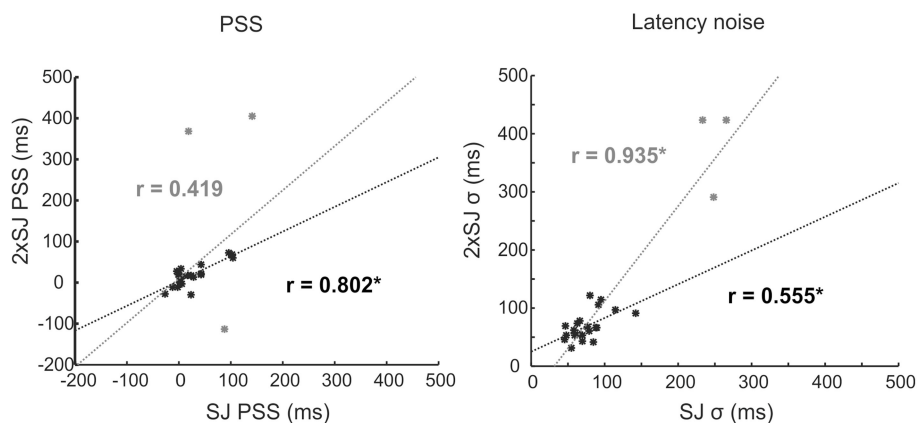
## Discussion

We fitted around 2300 trials from each of two motivated observers and 100–300 trials from each of two sets of 24 typical psychology participants, using simple but plausible models of the TOJ, SJ, and roving 2xSJ tasks. We also recorded simple RTs for one of these groups. Our latency models described the data fairly well for the two observers, but the models were demonstrably incomplete as data were significantly overdispersed. For individual observers, PSS values were more positive when estimated from 2xSJ data than from TOJ and SJ data, but at the group level we found no significant differences between PSS values estimated using our tasks. Group-level estimates of differential latency noise were significantly higher for the 2xSJ task than for the simple RT task, and for the TOJ task than for the 2xSJ task, with the latter result mirrored for





**FIGURE 6 | Scatter plots for correlations in parameter estimates across the 24 participants in Experiment 1b, along with lines of best fit. Asterisks (\*) denote significance ( $p < 0.05$ ; PSS, top) or marginal significance (one-tailed  $p < 0.05$ ;  $\sigma$ , bottom).**



**FIGURE 7 | Scatter plots for correlations in parameter estimates across the 24 participants in Experiment 1c, along with lines of best fit. Data for outlying participants (and correlations/fits for data incorporating those participants) are shown in gray. Data, correlations and fits with outliers excluded are shown in black. Asterisks (\*) denote significance ( $p < 0.05$ ).**

our naive observer, but not for our highly experienced observer. Estimates of latency noise were very similar for 2xSJ and SJ tasks when completed concurrently. At the group level, PSS estimates correlated for the TOJ and 2xSJ tasks and for the SJ and 2xSJ tasks, at least when extreme PSS estimates were removed.

The similar and correlated estimates of PSS provided by 2xSJ and TOJ tasks, and by the 2xSJ and SJ tasks, all of which have good face-validity as measures of temporal perception, provide some degree of cross validation for our 2xSJ procedures, and suggest that these tasks are accessing broadly similar cognitive processes. However, the differences in noise parameter estimates, all of which theoretically measure the same quantity ( $\sigma$ ), suggest latency variability is not the only source of noise in these tasks, as our modeling naively assumed. The lowest estimate was provided by the simple-RT task, but realistically this must already be an

overestimate because the RT task inherits some variability from the motor system that we did not consider formally.<sup>5</sup> The simple RT task might also rely on sensory pathways somewhat distinct from those used in other timing tasks, but assuming substantial overlap, lower RT noise suggests that the 2xSJ task might gain substantial noise at the decision level, or perhaps as a result of higher memory demands. We can however rule out an interval bias as a possible cause: Although we observed an interval preference, and such biases can have the effect of increasing noise estimates in 2AFC tasks, we explicitly modeled the interval bias for the 2xSJ and thus our estimates are uncontaminated in this respect.

<sup>5</sup>It is, however, possible that we underestimated noise slightly in the RT task, as we relied on data trimming to exclude outliers.

The increase in estimated noise from the 2xSJ to the TOJ task was even more striking than the increase for 2xSJ over the RT task. One possible explanation is that keying errors were more frequent in the TOJ task. It is possible to fit models with additional parameters to describe such errors (Wichmann and Hill, 2001; García-Pérez and Alcalá-Quintana, 2012a,b). However, to do so effectively it is necessary to sample extensively at extreme SOAs where performance asymptotes; for example, fitting SM's data (**Figure 4** second panel down on the right) with the lapse rate free to vary actually yielded the same estimate (1%) that had been fixed/assumed in our original fit, and hence also the same estimate of noise. In any case, it is not clear how much explanatory value this kind of account really has, even if it can provide a more appropriate measure of sensory noise, as it still begs the question of why participants are so prone to keying errors in the TOJ task.

The increased noise in the TOJ task might reflect additional processing steps for TOJ over and above those for SJs involving, for example, the binding of event content with event timings (Fujisaki and Nishida, 2005). Another possibility is that values of  $\Delta t$  (i.e., the subjective SOA) near zero cannot be recovered by observers, forcing them to guess in this region (García-Pérez and Alcalá-Quintana, 2012a,b). In this case a lower estimate of sensory noise might be obtained by fitting a TOJ model that explicitly models this low threshold. However, it is worth noting that such operations appear to have had only a limited impact for our more highly experienced observer. For novice participants these operations seem to provide a significantly greater challenge than the extra decision processes inherent in the 2xSJ (which requires that individual SOAs be remembered and compared). The fact that the 2xSJ returns lower estimates of noise is not trivial from a practical perspective, as these values are also better estimated (i.e., sit within tighter confidence intervals) relative to the TOJ. This may make the 2xSJ procedure a more useful task when assessing changes in noise across conditions, although the SJ also appears strong in this regard, and explicit modeling of additional processes might improve estimates for the TOJ.

While there may be some value in employing the 2xSJ in place of the TOJ, the SJ provided similarly low estimates of noise and is clearly a simpler and quicker task to implement. However, we have illustrated how 2xSJ data might be collected at the same time, and our preliminary comparison with previous SJ data (collected without a concurrent 2xSJ task) suggests the additional 2xSJ task encourages participants to use more constrained decision criteria for their SJs. This is potentially valuable, as when participants use very liberal criteria in the SJ, so that many SOAs are judged synchronous almost 100% of the time, any derived PSS value becomes more contentious. Specifically, it will depend to a greater extent on modeling assumptions, for example that participants place their decision criteria at equal distances from subjective time zero. However, our (informal) result would benefit from a more rigorous test, as our data sets differed in respects other than the presence or absence of the concurrent 2xSJ question. Although the set up was broadly similar, LED color, number of trials,

and SOA sampling scheme all differed between the data sets we compared.

Having obtained preliminary evidence that the 2xSJ task provides estimates of PSS and latency noise that are broadly compatible with those found using more established tasks, we next determined whether similar correlations could be obtained using stimuli from different modalities (i.e., all combinations of visual, tactile and auditory stimuli) and also with another common temporal judgment task closely related to the TOJ and the SJ, the ternary (SJ3) judgment task.

## EXPERIMENT 2

### Methods

Methods in Experiment 2 were identical to those in Experiments 1b with the following exceptions.

#### Participants

An opportunity sample of 6 participants was tested, including two authors (mean age = 27.7, range 20–37, three male).

#### Apparatus and Stimuli

Tactile stimuli were vibrotactile sine waves, identical to auditory stimuli except that their frequency was 200 Hz. Vibrotactile stimuli were delivered via a small (~1 cm diameter) ceramic piezoelectric disk coated in plastic. The disk was driven from a custom-built amplifier, and did not produce audible noises with the stimuli we used. It was gripped comfortably between index finger and thumb of the left (non-responding) hand, which rested on participants' laps, around 30 cm from the visual and auditory stimuli.

#### Design and Procedure

A  $4 \times 3$  factorial repeated-measures design manipulated both the temporal task (RT, TOJ, 2xSJ, and ternary) and the modality pairing that participants were judging (AV, audiovisual; AT, audiotactile; VT, visuotactile). The four tasks were presented in separate blocks within a single session, always in the same order (ternary, then TOJ, then RT, then 2xSJ). The three modality pairings were completed in separate sessions, with order counterbalanced across participants. For the ternary task, stimulus selection was as per the TOJ task, but in addition to the two order response options, participants could now opt to respond "simultaneous." If they did so, they were subsequently prompted to take a guess about order (used to update the adaptive distribution from which SOAs were being selected, and discourage excessive use of the simultaneous response option) but these responses were not analyzed.

#### Data Analysis

Data in the ternary task were fitted with the same model described previously for simultaneity judgments, except that model predictions were expressed for the three possible response categories, and maximum-likelihood fitting assumed a multinomial data model. To check for sensible responding, goodness of fit was compared against a two-parameter guessing model incorporating guess rates for two out of three response options.

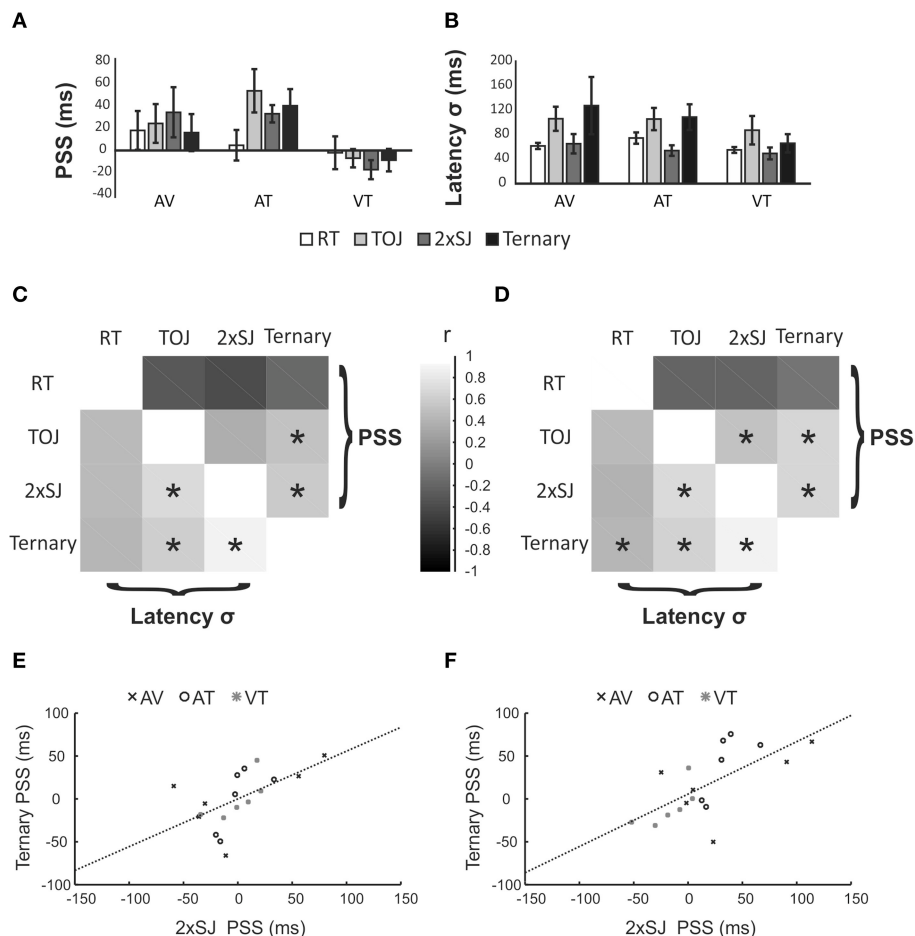
## Results

**Figure 8** provides an overview of the results from Experiment 2. Group average PSS and latency noise values are presented in **Figures 8A,B** respectively. PSS values were once again slightly positive for AV conditions, a trend that was exacerbated for AT conditions but reversed for VT conditions. Different tasks gave quite similar PSS estimates on average (with the possible exception of RT in the AT condition). Estimates of latency noise were similar between modality pairings, but appeared lower in RT and 2xSJ tasks compared to TOJ and ternary tasks. However, no differences between tasks reached statistical significance for either PSS or latency noise (perhaps reflecting the small sample size in this experiment).

In order to increase power to detect correlations, we combined data from all six observers and three modality pairings into 18 points. Differences between pairings might lead to a clustering of data into three sets. Hence any correlation would be driven in part by the common effect of a particular modality pairing on measures from two or more tasks. Although we consider

this essentially legitimate (i.e., if a change of modality pairings affects the PSS from two tasks in the same way, this is reasonable evidence that the two tasks are indexing similar mental operations) we also performed correlations after first normalizing data within each modality pairing. We did this by subtracting the mean for that pairing, so that only differences relative to the mean remained to be correlated between tasks.

Correlations are shown in **Figures 8C–F**. **Figures 8C,D** summarize correlations between all four tasks for both PSS and latency noise. Broadly, correlations between equivalent measures of latency noise are positive for all task pairs, whereas correlations between measures of PSS are generally low and slightly negative between the RT task and the other tasks, but high and positive between the three temporal judgment tasks. Focussing on the critical correlations between the 2xSJ task and the other tasks (and omitting marginal and non-significant results), with normalization there was a significant PSS correlation between the 2xSJ task and the ternary task ( $r = 0.575$ ,  $p = 0.013$ ), and a significant latency noise correlation between the 2xSJ task and



**FIGURE 8 | Results of Experiment 2. (A,B)** Mean parameter estimates (PSS and  $\sigma$ ) for the six participants in each of four (task)  $\times$  three (modality pairing) conditions. AV, Audiovisual; AT, audiotactile; VT, visuotactile. Error bars denote standard error of the mean. **(C,D)** Between-task correlations for PSS values (upper right) and  $\sigma$  values (lower left). Data were pooled across all modality pairings, but in **(C)** they were first normalized to the mean value within each modality pairing to remove variance associated with this manipulation. Asterisks (\*) denote statistically significant differences ( $p < 0.05$ ). **(E,F)** Scatter plots for one illustrative correlation (for the PSS, between 2xSJ and ternary tasks). Data are the same in both plots, but have been normalized (as for **C**, above) in **(E)**.

both the ternary task ( $r = 0.881$ , bootstrap  $p < 0.05$ ) and the TOJ task ( $r = 0.707$ ,  $p = 0.001$ ). Without normalization, correlations were generally slightly higher. Here, there were significant PSS correlations between the 2xSJ task and both the ternary task ( $r = 0.623$ ,  $p = 0.003$ ) and the TOJ task ( $r = 0.527$ ,  $p = 0.025$ ). Similarly, for latency noise there were significant correlations between the 2xSJ task and both the ternary task ( $r = 0.877$ , bootstrap  $p < 0.05$ ) and the TOJ task ( $r = 0.708$ ,  $p = 0.001$ ). The scatter plot for the correlation between the ternary and 2xSJ tasks is shown in **Figure 8** parts E and F (for normalized and non-normalized, data respectively).

## Discussion

In Experiment 2, we had observers make temporal judgments and rapid button presses in response to audiovisual, audiotactile, and visuotactile stimuli. The overall pattern of mean PSS values we recovered using four different tasks was similar across tasks. A simple reading would be that the auditory pathway is somewhat shorter than both the visual and tactile pathways, with the difference being greatest between auditory and tactile pathways. However, this result is likely to be stimulus specific and other interpretations are possible. For our purposes, the more important result is that the 2xSJ task provided results comparable to other temporal judgments tasks, and correlated with them for both PSS and latency noise measures (although as in experiment 1b, correlations with RT were lower for latency noise and absent for PSS). In particular, the new correlation between PSSs obtained using 2xSJ and ternary judgment tasks (based on several modality pairings) corroborates those previously obtained in Experiment 1b and 1c using TOJ and SJ tasks (with only AV stimuli).

Having found further evidence for the utility of the 2xSJ task when assessing a baseline PSS, we wanted to determine if it can also provide sensible estimates of *changes* in PSS across experimental conditions. For this purpose we attempted to recreate a classic experimental effect from the literature—cross-modal prior entry—tested with both 2xSJ and TOJ tasks.

## EXPERIMENT 3

### Methods

Methods in Experiment 3 were identical to those in Experiments 1a–c with the following exceptions.

### Participants

An opportunity sample of 11 naive participants was tested, with three excluded from further analysis as one or both of the observer models failed to fit their data better than the relevant chance model in one or more conditions. This yielded a sample size of 8 (mean age = 33.5, range 18–52, two male).

### Apparatus and Stimuli

Auditory stimuli were delivered through headphones (Sennheiser PX360). In order to manipulate the allocation of attention, a subset of stimuli were modified to become targets in a (secondary) detection task. In contrast to the usual stimulus duration of 10 ms, these stimuli had durations of 17 ms (for auditory targets) or 25 ms (visual targets).

## Design and Procedure

A  $2 \times 2$  factorial repeated-measures design manipulated both the temporal task (TOJ vs. 2xSJ) and the modality that participants had to monitor for targets in an additional detection task (auditory vs. visual). The four conditions were presented in separate blocks, with order counterbalanced across participants in a nested fashion (i.e., four possible orders, where each task could be completed first or second, and nested within that ordering each modality could be attended first or second). In addition to the two response options for temporal judgments (outlined in Experiments 1a and 1b), participants now received a third alternative—to indicate that a target had been present (in which case they were told not to worry about the temporal judgment). Accurate feedback was provided regarding the secondary detection task, flagging hits and misses on target-present trials and false alarms on temporal-judgment trials.

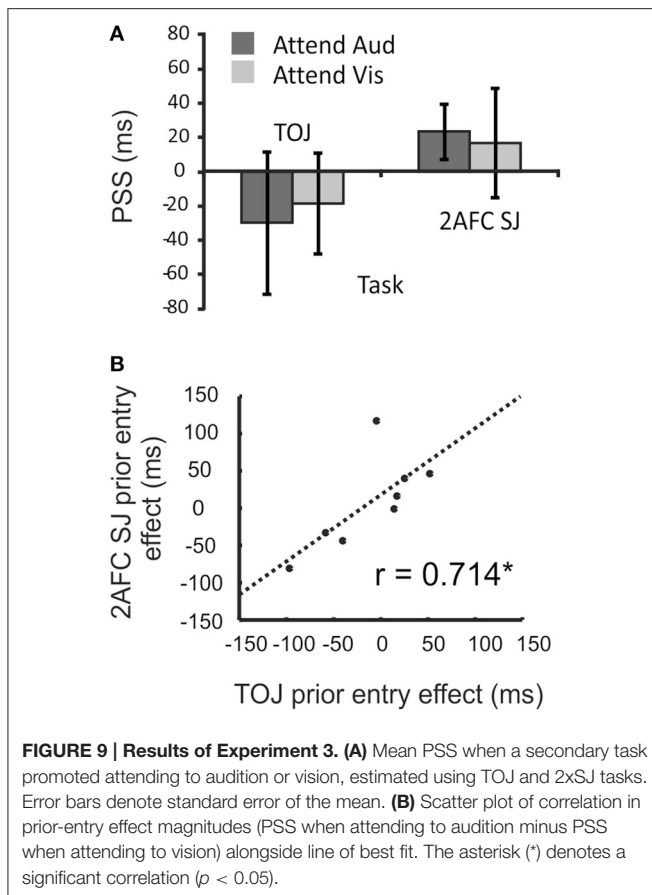
Blocks contained 190 trials, with 80% non-target (i.e., temporal-judgment) trials and 20% target trials. Targets were presented only in the monitored modality. The extra dual-task requirement made the temporal tasks more difficult. To counter this, for the 2xSJ task SOAs ranged more widely. One stimulus was drawn from the following 19 SOAs:  $-375$ ,  $-325$ ,  $-275$ ,  $-225$ ,  $-175$ ,  $-125$ ,  $-75$ ,  $-50$ ,  $-25$ ,  $0$ ,  $25$ ,  $50$ ,  $75$ ,  $125$ ,  $175$ ,  $225$ ,  $275$ ,  $325$ ,  $375$  ms (with each SOA occurring five times in each interval across a block of trials). The second SOA was drawn from a discrete probability distribution with steps of 25 ms, initially uniform, spanning  $-75$  to  $+75$  ms, but potentially expanding to  $\pm 375$  ms in an adaptive manner. For the TOJ task, SOA values from  $-450$  to  $+450$  ms were used (in 30 ms steps). This distribution was initially uniform across this entire range except for the two most extreme values, which were nine times more likely to occur than each of the 29 other SOAs (prior to adaptive updating).

## Results

In the secondary task, participants tended to detect targets successfully, but performance was imperfect. Hits and false alarms were converted to  $d'$ -prime ( $d'$ ) values (Green and Swets, 1966), with average  $d'$  for the group ranging from 2.22 for visual-target TOJ trials (79.3% hits, 6.3% false alarms) to 4.44 for auditory-target 2xSJ trials (96.1% hits, 0.3% false alarms). Hence there was an incentive to attend the modality containing detection targets.

We expected to see the PSS become more positive when participants attended audition relative to when they attended vision (as the auditory signal should be sped in the brain, and thus require a physical delay to seem synchronous). However, as shown in **Figure 9A**, on average the target modality (and thus the presumed allocation of attention) had no effect on PSS values estimated via either the TOJ task [ $t_{(11)} = 0.66$ ,  $p > 0.05$ ] or the 2xSJ task [ $t_{(11)} = 0.35$ ,  $p > 0.05$ ]. Furthermore, there was no evidence for a different magnitude of prior-entry effect between TOJ and 2xSJ blocks [with effect magnitude being the difference in PSS between auditory and visual-target conditions;  $t_{(11)} = 1.24$ ,  $p > 0.05$ ]. However, when we examined the prior-entry effect on a participant-by-participant basis, comparing the effect's magnitude derived using the TOJ task with that obtained using





**FIGURE 9 | Results of Experiment 3. (A)** Mean PSS when a secondary task promoted attending to audition or vision, estimated using TOJ and 2xSJ tasks. Error bars denote standard error of the mean. **(B)** Scatter plot of correlation in prior-entry effect magnitudes (PSS when attending to audition minus PSS when attending to vision) alongside line of best fit. The asterisk (\*) denotes a significant correlation ( $p < 0.05$ ).

the 2xSJ task, a significant correlation emerged ( $r = 0.71$ ,  $p < 0.05$ ; see **Figure 9B**).

## Discussion

In Experiment 3 we manipulated attention, directing it toward either the visual or auditory modality via a strategic incentive (to maximize performance on a concurrent detection task), while measuring changes in PSS via both a TOJ task and a roving 2xSJ task. We failed to obtain a prior-entry effect on average across participants, but did obtain a significant correlation between attentional influences on our two timing tasks.

It is not uncommon to fail to find cross-modal prior entry, particularly with manipulations of endogenous attention (e.g., Cairney, 1975). We used a 100% predictive instruction (i.e., targets always came only in the attended modality) so cannot offer any independent evidence that attention was allocated as we envisaged, but we think it likely on strategic grounds. Our manipulation of attention could be considered to be either between modalities or between spatial locations or, most likely, between both of these (as target stimuli came from either a fixated LED or via headphones). However, this manipulation had no significant effect on the PSS for our sample. Perhaps there really is no consistent effect to find, or perhaps the average effect is very small (e.g., associated latency changes in ERP components are tiny; Vibell et al., 2007) and we lacked power to demonstrate it.

We did, however, find evidence for a correlation in the (non-uniform) effects of attention on PSS estimates across participants. This correlation is interesting for two reasons. First, it demonstrates that while the experimental manipulation did not have a consistent effect on all participants, it influenced each participant's PSS in an individually reliable fashion (as revealed by the matched effects obtained using two different temporal judgment tasks in separate blocks of trials). Second, it provides further evidence that TOJ and 2xSJ tasks tap similar temporal processes.

## GENERAL DISCUSSION

In this paper, we (1) considered the merit of a roving 2xSJ task for estimating the bias and precision of temporal judgments; (2) provided predictions for a simple but theoretically-derived observer model, and; (3) benchmarked the task against more established TOJ, SJ, and ternary tasks when estimating both baseline PSS and (for the TOJ) changes in PSS. We found that the 2xSJ task was manageable for typical psychology participants; that the observer model was a somewhat useful approximation (albeit a simplification) of the full psychological process of temporal judgment; and that the 2xSJ task provides estimates of the PSS that are comparable to those obtained using other temporal judgments. The 2xSJ task is, however, likely to provide lower and less variable estimates of sensory noise than the TOJ, at least when the TOJ is modeled without additional cognitive operations such as guessing. On this basis we believe that the 2xSJ task has validity as a supplementary measure of temporal experience. From a practical perspective, we would recommend that researchers primarily consider using it in concert with the classic SJ (i.e., as an additional question) when one of the following two conditions apply:

- (1) When an estimate of the PSS is desirable between two quite different (e.g., bimodal) stimuli and there are reasons to believe that the SJ task will give rise to “synchronous” reports over a fairly broad range of SOAs (e.g., when stimuli are naturalistic/noisy or participants are not practiced). In this case, the 2xSJ should encourage the use of more conservative criteria regarding simultaneity (benefitting the fitting and interpretation of the SJ data) and provide an additional and less theoretically dependent point estimate of the PSS.
- (2) When a single-presentation task such as the TOJ or SJ seems likely to encourage decision biases, for example when the experimental manipulation could be seen to suggest one of the two possible answers when the observer is uncertain (e.g., a directive to preferentially attend one of two events).

In these situations, the 2xSJ should result in low and fairly stable estimates of sensory noise alongside a PSS that is less dependent upon the placement of decision criteria. However, such benefits must be weighed against the increased experimental time necessary to complete each trial.

We have considered five tasks here, but our main focus was the 2xSJ task. We obtained a significant correlation between PSS estimates from this task and other temporal judgment tasks,

and also between changes in PSS estimates across conditions for the 2xSJ and TOJ. Several previous studies have attempted to find correlations between PSS values estimated via more than one task. For example, both van Eijk et al. (2008) and Love et al. (2013) failed to find any correlation between the PSS estimated from a temporal order judgment and that estimated from a synchrony judgment, while Freeman et al. (2013) found a surprising negative correlation between PSS for audiovisual speech (estimated via TOJ) and the maxima of the function describing the probability of McGurk integration across different SOAs. We suggest previous failures to obtain correlations between TOJ and SJ tasks might reflect the different decision processes in these two tasks. In particular, the SJ is fundamentally a method for obtaining a *region* of subjective simultaneity, rather than a *point* of subjective simultaneity. To infer a PSS, it is necessary to make some assumption about how the two criteria for demarcating synchrony from asynchrony are selected (e.g., that they are placed symmetrically about a subjective  $\Delta t$  value of zero). However, participant strategies might vary, with a concomitant effect on the inferred PSS. Such strategic variability might make correlations difficult to detect. Our 2xSJ task, although still based on a judgment of simultaneity, forces observers to decide which SOA seems *most* synchronous, which might be more comparable to the SOA at which their impression of order switches in TOJs<sup>6</sup>.

Many studies have also reported differences between PSS estimates obtained using TOJ and SJ tasks. For example, Linares and Holcombe (2014) found that for four of seven participants, confidence intervals around the PSS did not overlap for TOJ and SJs. We obtained a similar result for two observers in Experiment 1a when comparing TOJ and 2xSJ parameters, although this difference was not apparent in the group data from Experiment 1b. Interestingly, Linares and Holcombe (2014) also found differences between PSS estimates obtained using a TOJ task and those obtained using a AV-VA (or VA-AV) duration comparison task, which shares a broad structural similarity with our 2xSJ task, but uses just a few rather longer durations. Differences in PSS values obtained using these kinds of tasks might reflect a common decision-level bias in the TOJ (e.g., observers tend toward one response when uncertain). Alternatively (or additionally) there might be an asymmetry in the transducer function that relates objective to subjective time for AV intervals relative to VA intervals (e.g., if AV time accrues more quickly than VA time at a subjective level, perhaps due to differences in arousal or attention). This could bias PSS estimates derived using both 2xSJ and interval-comparison tasks.

We did not obtain correlations between PSS estimates from our temporal judgment tasks and that estimated from simple RT, although there was some evidence for a correlation in estimates of noise involving RTs and other tasks. There is a previous literature examining the extent to which simple RT and TOJ tasks rely on the same sensory representations, with the main focus being the tendency for PSS estimated from TOJs to dissociate from that estimated using simple RT following experimental manipulations

such as changes in stimulus intensity (see e.g., Jaskowski, 1999, for a review of the early work). Here too dissociations may be explicable in terms of different decision strategies being applied to different tasks (Miller and Schwarz, 2006; Cardoso-Leite et al., 2007) rather than implying a complete mechanistic separation. Our mean PSS estimates actually matched fairly well between RT and temporal judgments, and the absence of a correlation is perhaps explicable in terms of fairly low experimental power combined with differences between these tasks from the decision level onwards.

Estimates of latency noise were often correlated between our various tasks. While this finding is consistent with a common sensory stage accessed by different tasks, as implied by independent-channels models, it might also have resulted from quite general cognitive factors, such as the ability to maintain focussed attention during a long, boring task. Stevenson and Wallace (2013) have previously reported correlations between measures of a construct known as the temporal binding window, derived using several of the tasks we assess here. They constructed this measure by fitting one or more sigmoids in a piecewise manner to their data, and calculating the difference between threshold values. It is rather difficult to map this kind of measure, which is likely to conflate latency noise and decision criteria (to different extents depending on the exact task) onto the model-based measures we derive here, but our findings are broadly consistent with theirs.

Fitting observer models to data is generally preferable to fitting arbitrary functions, as derived parameters will have clearly defined meanings. However, this is only true to the extent that the models are accurate. The observer models we develop and use here are very simple (too simple in several cases) but seem a reasonable starting point. There are many more complex variants that might be considered, and indeed some such variants have been shown to perform well for TOJ, SJ, and ternary tasks (García-Pérez and Alcalá-Quintana, 2012a,b, 2015). One example of the additional complexity we have omitted is the well-known scalar property (the variant of Weber's law that applies to time) as we have assumed constant noise alongside an affine transformation from objective to subjective SOAs. It remains to be seen whether more complex models that incorporate such features will provide a significantly better fit to temporal judgment data when their additional parametric flexibility is taken into consideration.

## CONCLUSIONS

We have outlined methods and analysis procedures for implementing a roving 2xSJ task, useful for determining both a point of subjective simultaneity and associated judgment precision estimates for subjective timing. This task returns PSS estimates that seem largely consistent with those returned by more traditional tasks, but in some cases provides lower and more constrained estimates of sensory noise, perhaps indicative of a more straightforward decision process. It does so while explicitly requiring participants to decide which alternative timing relationship is *most* synchronous on any given trial (rather than revealing what range of relationships are sometimes

<sup>6</sup>Of course the TOJ faces its own issues as a measure of maximal synchrony perception, not least the fact that it doesn't actually ask about synchrony directly, only about perceived order.

described as synchronous). It can also easily be combined with judgments about each stimulus. It therefore provides a useful complement to existing methods for investigating subjective timing.

## AUTHOR CONTRIBUTIONS

KY and DA conceived and designed the work. SM and SD acquired the data. JS, KY, SM, and SD analyzed the data. All

authors interpreted the data and drafted and approved the manuscript.

## ACKNOWLEDGMENTS

DA and KY's collaboration is supported by an Australian Research Council Discovery Grant. SM's RA position is funded by BBSRC grant BB/K01479X/1 to KY and JS. We thank Sara Shapiro for assistance with data collection.

## REFERENCES

- Allan, L. G. (1975). The relationship between judgments of successiveness and judgments of order. *Percept. Psychophys.* 18, 29–36. doi: 10.3758/BF03199363
- Allan, L. G., and Kristofferson, A. B. (1974). Successiveness discrimination: Two models. *Percept. Psychophys.* 15, 37–46. doi: 10.3758/BF03205825
- Baron, J. (1969). Temporal ROC curves and the psychological moment. *Psychon. Sci.* 15, 299–300. doi: 10.3758/BF03336309
- Butler, R. W., and Paoletta, M. S. (2002). Calculating the density and distribution function for the singly and doubly noncentral F. *Stat. Comput.* 12, 9–16. doi: 10.1023/A:1013160019893
- Cairney, P. T. (1975). Bisensory order judgement and the prior entry hypothesis. *Acta Psychol. (Amst.)*, 39, 329–340. doi: 10.1016/0001-6918(75)90025-6
- Cardoso-Leite, P., Gorea, A., and Mamassian, P. (2007). Temporal order judgment and simple reaction times: Evidence for a common processing system. *J. Vis.* 7:11. doi: 10.1167/7.6.11
- Efron, B., and Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.
- Freeman, E. D., Ipser, A., Palmabaha, A., Paunoiu, D., Brown, P., Lambert, C., et al. (2013). Sight and sound out of synch: Fragmentation and renormalisation of audiovisual integration and subjective timing. *Cortex* 49, 2875–2887. doi: 10.1016/j.cortex.2013.03.006
- Fujisaki, W., and Nishida, S. (2005). Temporal frequency characteristics of synchrony-asynchrony discrimination of audio-visual signals. *Exp. Brain Res.* 166, 455–464. doi: 10.1007/s00221-005-2385-8
- Fujisaki, W., Shimojo, S., Kashino, M., and Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat. Neurosci.* 7, 773–778. doi: 10.1038/nn1268
- García-Pérez, M. A., and Alcalá-Quintana, R. (2012a). On the discrepant results in synchrony judgment and temporal-order judgment tasks: A quantitative model. *Psychon. Bull. Rev.* 19, 820–846. doi: 10.3758/s13423-012-0278-y
- García-Pérez, M. A., and Alcalá-Quintana, R. (2012b). Response errors explain the failure of independent-channels models of perception of temporal order. *Front. Psychol.* 3:94. doi: 10.3389/fpsyg.2012.00094
- García-Pérez, M. A., and Alcalá-Quintana, R. (2015). Converging evidence that common timing processes underlie temporal-order and simultaneity judgments: A model-based analysis. *Atten. Percept. Psychophys.* 77, 1750–1766. doi: 10.3758/s13414-015-0869-6
- García-Pérez, M. A., and Peli, E. (2014). The bisection point across variants of the task. *Atten. Percept. Psychophys.* 76, 1671–1697. doi: 10.3758/s13414-014-0672-9
- Gibbon, J., and Rutschmann, R. (1969). Temporal order judgement and reaction time. *Science* 165, 413–415. doi: 10.1126/science.165.3891.413
- Green, D. M., and Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York, NY: Wiley.
- Jaskowski, P. (1999). "Reaction time and temporal-order judgment as measures of perceptual latency: The problem of dissociations," in *Cognitive Contributions to the Perception of Spatial and Temporal Events*, eds G. Aschersleben, T. Bachmann and J. Mueseler (Amsterdam: Elsevier), 265–282.
- Jogan, M., and Stocker, A. A. (2014). A new two-alternative forced choice method for the unbiased characterization of perceptual bias and discriminability. *J. Vis.* 14:20. doi: 10.1167/14.3.20
- Linares, D., and Holcombe, A. O. (2014). Differences in perceptual latency estimated from judgments of temporal order, simultaneity and duration are inconsistent. *Iperception* 5, 559–571. doi: 10.1068/i0675
- Love, S. A., Petrini, K., Cheng, A., and Pollick, F. E. (2013). A psychophysical investigation of differences between synchrony and temporal order judgments. *PLoS ONE* 8:e54798. doi: 10.1371/journal.pone.0054798
- Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide, 2nd Edn*. New York, NY: Lawrence Erlbaum Associates.
- Miller, J., and Schwarz, W. (2006). Dissociations between reaction times and the planning of vertical saccadic eye movements. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 394–412. doi: 10.1037/0096-1523.32.2.394
- Morgan, M., Grant, S., Melmoth, D., and Solomon, J. A. (2015). Tilted frames of reference have similar effects on the perception of gravitational vertical and the planning of vertical saccadic eye movements. *Exp. Brain Res.* 233, 2115–2125. doi: 10.1007/s00221-015-4282-0
- Morgan, M., Melmoth, D., and Solomon, J. (2013). Linking hypotheses underlying class A and class B methods. *Vis. Neurosci.* 30, 197–206. doi: 10.1017/S095252381300045X
- Paoletta, M. (2007). *Intermediate Probability: A Computational Approach*. Chichester, UK: John Wiley & Sons.
- Patten, M. L., and Clifford, C. W. G. (2015). A bias-free measure of the tilt illusion. *J. Vis.* 15:8. doi: 10.1167/15.15.8
- Powers, A. R., Hillock, A. R., and Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *J. Neurosci.* 29, 12265–12274. doi: 10.1523/JNEUROSCI.3501-09.2009
- Roseboom, W., Linares, D., and Nishida, S. (2015). Sensory adaptation for timing perception. *Proc. R. Soc. Ser. B Biol. Sci.* 282:20142833. doi: 10.1098/rspb.2014.2833
- Roseboom, W., Nishida, S., Fujisaki, W., and Arnold, D. H. (2011). Audio-visual speech timing sensitivity is enhanced in cluttered conditions. *PLoS ONE* 6:e18309. doi: 10.1371/journal.pone.0018309
- Rosenberger, W. F., and Grill, S. E. (1997). A sequential design for psychophysical experiments: an application to estimating timing of sensory events. *Stat. Med.* 16, 2245–2260.
- Schneider, K. A., and Bavelier, D. (2003). Components of visual prior entry. *Cogn. Psychol.* 47, 333–366. doi: 10.1016/S0010-0285(03)00035-5
- Shore, D. I., Spence, C., and Klein, R. M. (2001). Visual prior entry. *Psychol. Sci.* 12, 205–212. doi: 10.1111/1467-9280.00337
- Spence, C., and Parise, C. (2010). Prior-entry: a review. *Conscious. Cogn.* 19, 364–379. doi: 10.1016/j.concog.2009.12.001
- Spence, C., Shore, D. I., and Klein, R. M. (2001). Multisensory prior entry. *J. Exp. Psychol.* 130, 799–832. doi: 10.1037/0096-3445.130.4.799
- Sternberg, S., and Knoll, R. L. (1973). "The perception of temporal order: Fundamental issues and a general model," in *Attention and Performance IV*, ed S. Kornblum (London: Academic Press), 629–686.
- Stevenson, R. A., and Wallace, M. T. (2013). Multisensory temporal integration: task and stimulus dependencies. *Exp. Brain Res.* 227, 249–261. doi: 10.1007/s00221-013-3507-3
- Titchener, E. B. (1908). *Lectures on the Elementary Psychology of Feeling and Attention*. New York, NY: Macmillan.
- Ulrich, R. (1987). Threshold models of temporal-order judgments evaluated by a ternary response task. *Percept. Psychophys.* 42, 224–239. doi: 10.3758/BF03203074
- Van de Par, S., and Kohlrausch, A. (2000). "Sensitivity to auditory-visual asynchrony and to jitter in auditory-visual timing," in *Proceedings of SPIE: Human Vision and Electronic Imaging V*, Vol. 3959, eds B. E. Rogowitz and T. N. Pappas (Bellingham, WA: SPIE Press), 234–242.

- van Eijk, R. L., Kohlrausch, A., Juola, J. F., and van de Par, S. (2008). Audiovisual synchrony and temporal order judgments: effects of experimental method and stimulus type. *Percept. Psychophys.* 70, 955–968. doi: 10.3758/PP.70.6.955
- Venables, P. H. (1960). Periodicity in reaction time. *Br. J. Psychol.* 51, 37–43. doi: 10.1111/j.2044-8295.1960.tb00722.x
- Vibell, J., Klinge, C., Zampini, M., Spence, C., and Nobre, A. C. (2007). Temporal order is coded temporally in the brain: Early event-related potential latency shifts underlying prior entry in a cross-modal temporal order judgment task. *J. Cogn. Neurosci.* 19, 109–120. doi: 10.1162/jocn.2007.19.1.109
- Vroomen, J., Keetels, M., de Gelder, B., and Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Brain Res. Cogn. Brain Res.* 22, 32–35. doi: 10.1016/j.cogbrainres.2004.07.003
- Wichmann, F. A., and Hill, N. J. (2001). The psychometric function: I. fitting, sampling, and goodness of fit. *Percept. Psychophys.* 63, 1293–1313. doi: 10.3758/BF03194544
- Yarrow, K., Jahn, N., Durant, S., and Arnold, D. H. (2011). Shifts of criteria or neural timing? the assumptions underlying timing perception studies. *Conscious. Cogn.* 20, 1518–1531. doi: 10.1016/j.concog.2011.07.003
- Yarrow, K., Minaei, S., and Arnold, D. H. (2015). A model-based comparison of three theories of audiovisual temporal recalibration. *Cogn. Psychol.* 83, 54–76. doi: 10.1016/j.cogpsych.2015.10.002
- Yarrow, K., Sverdrup-Stueland, I., Roseboom, W., and Arnold, D. H. (2013). Sensorimotor temporal recalibration within and across limbs. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 1678–1689. doi: 10.1037/a0032534

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Yarrow, Martin, Di Costa, Solomon and Arnold. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Monkeys Share the Human Ability to Internally Maintain a Temporal Rhythm

Otto García-Garibay, Jaime Cadena-Valencia, Hugo Merchant and Victor de Lafuente \*

*Instituto de Neurobiología, Universidad Nacional Autónoma de México, Querétaro, Mexico*

## OPEN ACCESS

### Edited by:

Lars Muckli,  
University of Glasgow, UK

### Reviewed by:

Rodrigo Laje,  
National University of Quilmes,  
Argentina  
Angus Paton,  
University of Glasgow, UK

### \*Correspondence:

Victor de Lafuente  
lafuente@unam.mx

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 11 February 2016

**Accepted:** 05 December 2016

**Published:** 23 December 2016

### Citation:

García-Garibay O, Cadena-Valencia J,  
Merchant H and de Lafuente V (2016)  
Monkeys Share the Human Ability to  
Internally Maintain a Temporal  
Rhythm. *Front. Psychol.* 7:1971.  
doi: 10.3389/fpsyg.2016.01971

Timing is a fundamental variable for behavior. However, the mechanisms allowing human and non-human primates to synchronize their actions with periodic events are not yet completely understood. Here we characterize the ability of rhesus monkeys and humans to perceive and maintain rhythms of different paces in the absence of sensory cues or motor actions. In our rhythm task subjects had to observe and then internally follow a visual stimulus that periodically changed its location along a circular perimeter. Crucially, they had to maintain this visuospatial tempo in the absence of movements. Our results show that the probability of remaining in synchrony with the rhythm decreased, and the variability in the timing estimates increased, as a function of elapsed time, and these trends were well described by the generalized law of Weber. Additionally, the pattern of errors shows that human subjects tended to lag behind fast rhythms and to get ahead of slow ones, suggesting that a mean tempo might be incorporated as prior information. Overall, our results demonstrate that rhythm perception and maintenance are cognitive abilities that we share with rhesus monkeys, and these abilities do not depend on overt motor commands.

**Keywords:** rhythm, timing, rhesus, Weber fraction, model of time perception

## INTRODUCTION

The ability to estimate time intervals is fundamental to behavior. Motor actions performed outside their intended temporal window often have reduced effectiveness or a complete loss of purpose. However, the mechanisms allowing the brain to time future sensory and motor events are not yet completely understood (Merchant and de Lafuente, 2014). Human, and to a certain extent, monkey subjects can repeatedly tap in synchrony with sensory stimuli (*synchronization*), and they can continue tapping in the absence of external stimuli (*continuation*) (Wing and Kristofferson, 1973; Ivry and Hazeltine, 1995; Zarco et al., 2009; Repp and Su, 2013). The increase in variability of the tapping responses that define time intervals is well described by the generalized Weber's law:

$$\sigma^2 = k \cdot T^2 + \sigma_{indep}^2 \quad (1)$$

in which  $T$  is elapsed time,  $k$  approaches the square root of the Weber fraction at long elapsed times, and the term  $\sigma_{indep}^2$  represents a basal variance that does not increase with time (Getty, 1975; Killeen and Weiss, 1987; Gibbon et al., 1997; Bizo et al., 2006; Merchant et al., 2008; Zarco et al., 2009; Laje et al., 2011).

However, the capacity of human and non-human primates to maintain a rhythm in the absence of sensory cues, or a motor action such as tapping, has been less studied

(Grahn, 2009; Patel et al., 2009; Fitch, 2013; Repp and Su, 2013). A particularly important question that remains unanswered is whether monkeys are able to perceive and maintain a rhythm in the absence of overt motor actions (Bispham, 2006; Merchant and Honing, 2014). Here we characterize the behavior of human and rhesus subjects in a task in which they have to estimate the tempo of a periodic sensory event and then maintain that rhythm in the absence of movements. We hypothesize that human and monkey subjects share the ability to maintain a temporal rhythm in working memory, and that this is not dependent on overt motor actions. This will support the notion that rhythmic interval timing is a higher cognitive function not tied to particular motor actions, which is shared among primates.

We developed a rhythm task in which subjects had to observe a visual stimulus that periodically changed its location along a circular perimeter. After this *presentation* period, the stimulus disappeared and subjects had to internally follow its location as a function of elapsed time. Importantly, at a random time during this *continuation* phase, subjects were asked to indicate the estimated position of the stimulus (*Go-time*). Thus, this task generated a visuospatial rhythm defined by the time interval between location changes (Doherty et al., 2005), much like the rhythm defined by the motion of a discretely moving second hand in a clock. To correctly estimate the stimulus position subjects must first adjust their internal chronometers to the pace of the visual stimulus and then use that internal rhythm to predict the position during the continuation phase. Since we know that the variability of the timing estimates increases with elapsed time we expect the probability of correct responses to decline as a function of time.

Whether subjects time single intervals independently or they estimate total elapsed time is an important open question that we address in human subjects by analyzing the pattern of errors and also by fitting *continuous time* and a *reset time* models.

An important question in timing research is whether intervals of different lengths are timed by a single mechanism or whether different intervals use distinct chronometers. There is evidence that the standard Weber fraction is not constant for intervals larger than approximately 1.2 s (Hinton and Rao, 2004; Bizo et al., 2006; Lewis and Miall, 2009; Grondin, 2012, 2014; Allman et al., 2014), and this could be a sign that different clocks or timing processes are used to time intervals of different durations (Bangert et al., 2011; Rammsayer and Troche, 2014). We approach this issue by calculating the traditional Weber fraction for intervals of different duration, and also by fitting a model of the generalized Weber fraction (Equation 1). The results show that the Weber fraction diminishes as a function not only of total elapsed time, but also as a function of the interval length subdividing that total time (Grondin et al., 1999). Our results demonstrate that the generalized Weber law provides a satisfactory description of behavioral patterns such as the proportion of correct responses, the increase in variability as a function of time, and the systematic pattern of timing errors. The evidence suggests that short (0.5 s), medium (0.75 s), and long intervals (1.0 s) seem to be timed by mechanisms with increasingly large time-independent variance.

## METHODS

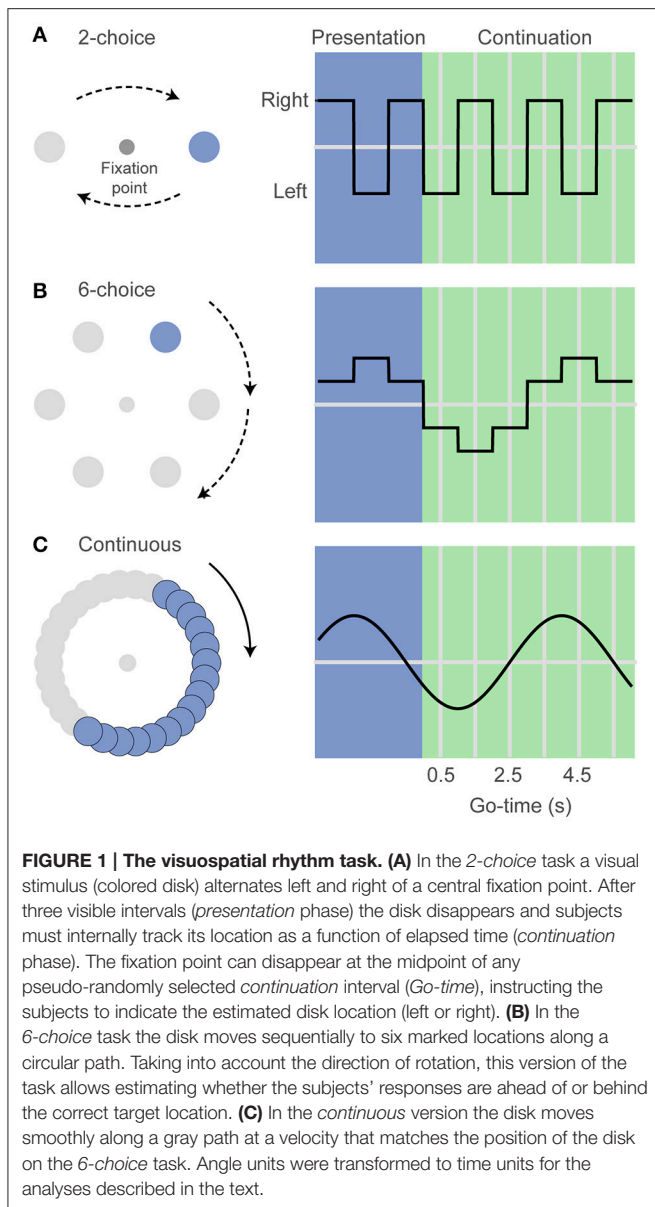
### Behavioral Tasks

In our visuospatial rhythm task the human subjects were asked to maintain their eyes in a fixed position (fixation) and to keep a mouse cursor at the center of a computer monitor while attending a peripheral disk that periodically changed location (**Figure 1**). After the presentation of 3 filled intervals (*presentation* phase), the disk disappeared and subjects had to covertly predict its position as a function of elapsed time (*continuation* phase). After 1–6 *continuation* intervals (uniform distribution, pseudo-randomly selected) the fixation point disappeared (*Go-time*), instructing the subjects to move the cursor and click over estimated position of the disk at the *Go-time*. It is important to note that the rhythm stops at *Go-time* and subjects can calmly click over the estimated position. In other words, it is not an interception task in which reaction time and hand movement should be taken into account when executing the behavioral response. The interval duration was chosen pseudo-randomly on each trial (0.50, 0.75, or 1 s for monkey and 8-choice datasets; 0.50 or 1 s for the rest of the datasets). Instead of using a mouse, monkeys were trained to maintain their right hand at the center of a touchscreen and at the *Go-time*, to perform a reach movement to touch the estimated location of the disk. They were rewarded with a drop of water on correct responses. An infrared camera (200 Hz, Applied Science Laboratories) was used to monitor eye position within 1.5° around the fixation point (**Figure 1**).

Monkeys were first trained in a 6-choice version of the task but then we decided to simplify it to a 2-choice task that is more suitable for the acquisition of neurophysiological data that we plan to carry after the behavioral tests presented in this report. (**Figure 1**). In addition to the 2-choice task, human subjects performed a 6-choice, an 8-choice, and also a *continuous* version of the task. The 6-choice and 8-choice versions of the task were included in the human experiments to accurately estimate how the variance of the behavioral responses changes as function of elapsed time. The use of 6 or 8 targets make it possible to measure whether responses are ahead of or behind the true stimulus position. This is not possible in the 2-choice task because there is only one correct and only one incorrect target.

In the *continuous* version of the task, the disk moved smoothly along a gray path. The disk moved at the same speeds as those in the 6-choice task. A response was defined as correct if the mouse click was within 30° of the correct position (this divides the gray circular path into six regions, analogous to the 6-choice task). We developed the *continuous* task as a control experiment in which timing is required to estimate the position of an invisible target (O'Reilly et al., 2008), but it does not depend on the rhythm imposed by the repetition of isochronous intervals.

To correctly predict the stimulus position subjects must rely on an internal chronometer whose variability increases with elapsed time, as described by the generalized Weber's law. Thus, the *Go-time* is a key experimental variable determining how well the subjects can estimate the disk location. Short *Go-times* will likely result in correct responses, while at long *Go-times* subjects are more likely to miss the correct disk location (they can get



ahead or behind the true location). Note that the spatial location of the stimulus (angle) and the spatial location of the behavioral responses (angle) were expressed in time units (seconds).

We describe behavioral performance with four variables, and we plot these as a function of *Go-time* (Figure 4): (1) The probability of a correct response  $p(\text{correct})$ , indicating the proportion of trials in which subjects correctly estimated the position of the disk; (2) the standard deviation (*Std*) of the responses, expressed in time units; (3) the traditional Weber fraction, defined as the standard deviation (*Std*) divided by the mean generated time (mean spatial location of the behavioral responses, converted to time units); and (4) the constant error, or bias, defined as the difference between the true and the estimated position of the disk, expressed in time units. It must be noted that the constant error can only be estimated in the 6-choice, 8-choice,

and the *continuous* versions of the task. The 2-choice version of the task allows recording correct and incorrect responses, but precludes determining whether an incorrect response was ahead of or behind the true stimulus position. The columns of Figure 4 show these four behavioral parameters for each dataset, grouped by interval duration, and plotted as a function of *Go-time*. The *Go-cue* (disappearance of the eye fixation point in humans or disappearance of hand fixation point in monkeys) occurred at the middle of 1–6 continuation intervals (pseudo-randomly selected; 1–4 continuation intervals in monkeys). Thus, *Go-times* were 0.5, 1.5, 2.5, 3.5, 4.5, 5.5, for the 1 s interval and 0.25, 0.75, 1.25, 1.75, 2.25, 2.75 for the 0.5 s interval. For monkeys, the first four of those *Go-times* were used, and an additional interval of 0.75 s was also tested (*Go-times* 0.38, 1.13, 1.88, 2.63 s).

## Participants, Apparatus, and Training

Thirteen human subjects were tested in this study and were paid for their participation (8 females, median age 25, Std 4.1). They were right-handed, had normal or corrected-to-normal vision, and were naive about the purpose of the experiment. All subjects reported no systematic musical training for more than a year. Each subject volunteered and gave informed consent for this study, which complied with the Declaration of Helsinki and was approved by the National University of Mexico Institutional Review Board. In addition to a minimum monetary compensation, human subjects were also compensated for every correct trial (feedback was provided on each trial by flashing the correct target position). Two male monkeys (*Macaca mulatta*, 5–7 kg, ages 5, and 6) were used. Animal experimental procedures were approved by the National University of Mexico Institutional Animal Care and Use Committee and conformed to the principles outlined in the Guide for Care and Use of Laboratory Animals (NIH, publication number 85-23, revised 1985). Human subjects were seated comfortably on a chair facing a computer monitor (LCD screen, 60 Hz refresh rate, model S27C350H) in a quiet room. Stimuli were generated and data were collected with custom software written in Matlab and the Psychophysics Toolbox (Brainard, 1997). Subjects came to the lab on separate days to perform each task type (2-choice, 6-choice, 8-choice, *continuous*). The order of the task type was counterbalanced between subjects. In each session subjects performed 48 training trials followed by a 15 min rest period, and then 288 test trials (6 *Go-times*, 2 interval durations, 24 repetitions) with 15 min rest periods every 98 trials. Monkeys spent ~4 months progressively learning the task structure, and another ~6 months for their performance to reach asymptotic levels. To make sure monkeys learned to estimate a rhythm (*presentation* phase) and then being able to use that rhythm to predict the stimulus position as a function of the elapsed time (*continuation* phase), we first trained them in a version of the 6-choice task in which interval length was chosen from a continuous distribution (300–1200 ms, uniform distribution) and the number of *presentation* intervals was variable (1–4, uniform distribution). This variation in initial conditions minimized the possibility of monkeys learning a simple association between elapsed time and a fixed stimulus position on the screen. We then moved to the 2-choice task that

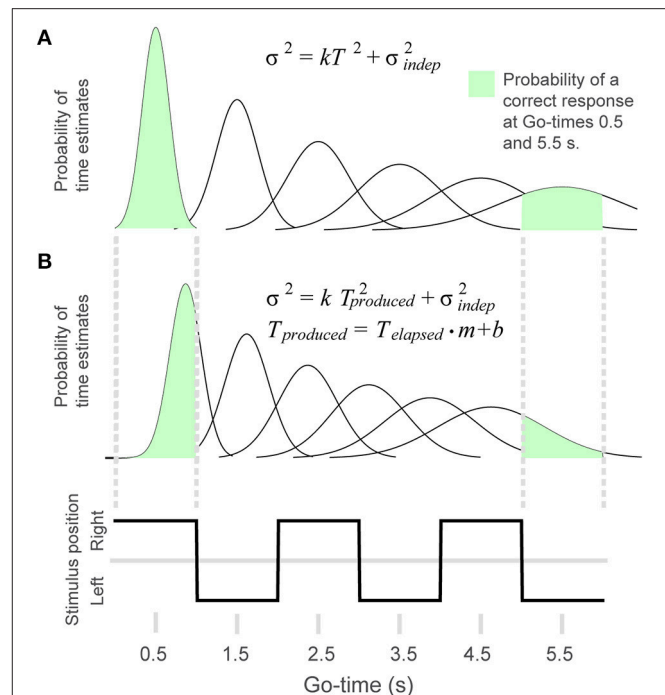
we present here and that will be used in physiological recordings in the future. The 2-choice version of the task is better suited for the acquisition and analysis of neuronal data because it has fewer conditions and variables. For example, it has only two possible starting and end locations of the stimulus. The behavioral decision is thus binary, allowing us to record many repetitions of the same type of trials and the underlying neurophysiological data. On each training day monkeys performed 3–6 runs with approximately 130 trials per run. The data analyzed here was obtained from 358 sessions in a ~4 month period following training (109 sessions monkey I; 249 sessions monkey M; 47,235 total trials).

## Fitting the Generalized Weber Law

To test the extent to which behavioral performance conformed to the scalar property of timing, we adapted the generalized Weber law to the discrete time intervals that define our rhythm task (Figures 2A,B). We generated a model in which the probability of a correct response  $p(\text{correct})$  was defined as the area under a Gaussian distribution that is comprised within the limits of the time interval corresponding to a given *Go-time* (this distribution represents the variability of the internal time estimates). For example, Figure 2A shows that the area comprised within the first *continuation* interval (*Go-time* = 0.5 s) is close to 1, whereas the area comprised within the sixth memory interval approximates 0.5 (*Go-time* = 5.5 s). In this manner, as described by the generalized Weber law (Equation 1), the time-dependent increase in variability results in a reduced proportion of correct responses as a function of *Go-time*, and the steepness of this decrease is modulated by the  $k$  parameter of Equation 1.

In its traditional form, the Weber fraction determines the slope with which the standard deviation of time estimates grows as a function of elapsed time:  $\sigma = k \cdot T$ ; where  $k$  is the Weber fraction,  $\sigma$  stands for standard deviation and  $T$  is elapsed time. However, it has been found that the addition of a time-independent noise constant better describes how  $\sigma$  grows as a function of time:  $\sigma = k \cdot T + \sigma_{\text{indep}}$ ; in which  $\sigma_{\text{indep}}$  represents this time-independent source of variability. The addition of this constant results in the traditional Weber fraction ( $\sigma / \text{mean}$ ) not being constant as a function of elapsed time: it is higher at short times and it decreases as time elapses. This is because at short times ( $T$  is small) the total variability is dominated by  $\sigma_{\text{indep}}$ , and as time elapses the total variability is mainly due to the  $k \cdot T$  product. Thus, at longer elapsed times the term  $k$  in Equation 1 approximates the traditional Weber fraction in which the variability is accounted by  $k \cdot T$ . When variability is expressed as variance and time is also squared, the resulting equation for the generalized Weber fraction is Equation (1).

In addition to fitting  $p(\text{correct})$ , our model also fit the standard deviation (*Std*) of the behavioral responses. However, it is not possible to directly fit Equation 1 to our data because of the discrete nature of the behavioral responses (2-, 6-, 8-choice), i. e., Equation (1) varies continuously whereas the subjects' responses vary within a finite number of options. Thus, the model calculates *Std* from the expected proportion of responses distributed across the discrete time intervals. In the case of the 2-choice task, for example, the discrete nature of the responses



**FIGURE 2 | To model behavioral performance we adapted the generalized Weber law to the rhythm task. (A)** The Gaussian distributions illustrate the time-dependent increase in variability of the timing estimates. The probability of a correct response was calculated as the area of the Gaussian curve comprised within the interval defined by a given *Go-time*. The probability of a correct response for *Go-times* 0.5 and 5.5 s is illustrated in green. **(B)** In the 4-parameter equation, produced time is modified by multiplicative and additive factors. This allows the model to capture systematic errors like shortening or lengthening of elapsed time. The figure illustrates the distributions resulting from a positive displacement and a shortening of time estimates.

causes the standard deviation to saturate at long elapsed times, when behavior is at random chance and the behavioral responses are distributed equally between the two choices (Figure 4, second column). Thus, the 0.5 s saturating value is the expected standard deviation of a random variable taking the values 0 and 1 s (as in the behavioral responses corresponding to correct and incorrect responses in the 1 s time interval trials).

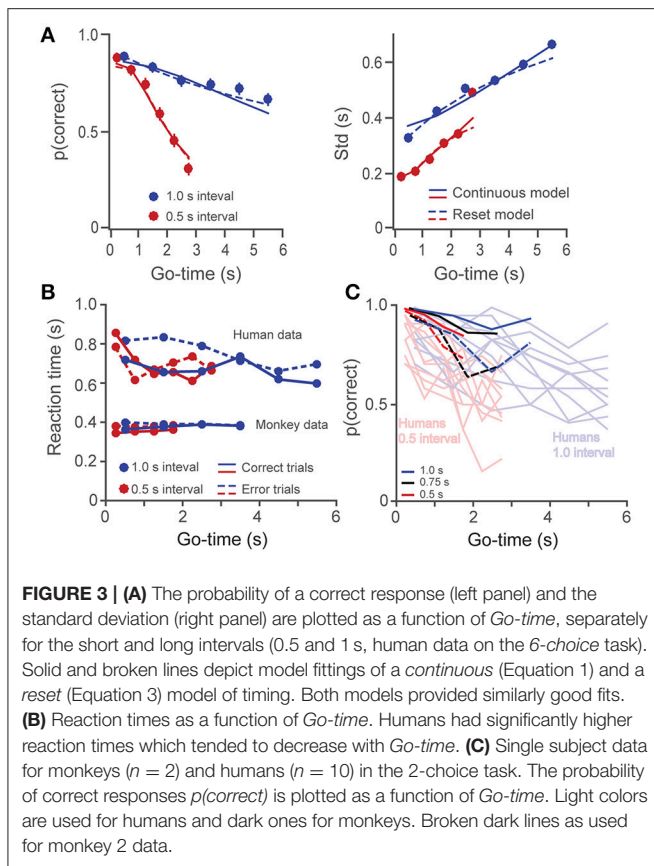
The generalized Weber law describes how variance changes as a function of time. However, it cannot account for systematic trends in the constant error that is, it cannot capture whether a subject's estimate of time is ahead of or behind true elapsed time. For our model to capture systematic differences between real and estimated time (constant error, Figure 4, rightmost column) we made use of two additional parameters ( $m$ ,  $b$ ):

$$\sigma^2 = k \cdot T_{\text{produced}}^2 + \sigma_{\text{indep}}^2$$

$$T_{\text{produced}} = T_{\text{elapsed}} \cdot m + b \quad (2)$$

These parameters allowed our model to take into account biases such as a constant time displacement ( $b$ ), and the shortening or lengthening of produced time ( $m$ ) (Figure 2B). Equation 2 was used for the fits shown in Figures 4C,D. However, when





comparing parameters  $k$  and  $\sigma_{indep}^2$  across tasks we used the two-parameter generalized Weber's model (Equation 1, this is because the 2-choice tasks do not allow to calculate the constant error).

As was done by Buonomano and colleagues (Laje et al., 2011), we tested a *reset* version of the generalized Weber law in which, instead of variance increasing in proportion with total time squared (term  $k \cdot T^2$ , Equation 1), it increased with the sum of the squares of each interval duration:

$$\sigma^2 = k \cdot (T_1^2 + T_2^2 + \dots + T_{Go-time}^2) + \sigma_{indep}^2 \quad (3)$$

Thus, in the *reset* version of the model, variance increases linearly rather than quadratically with time. By plotting Std as a function of *Go-time*, this trend can be observed as a saturating effect at large *Go-times* (Figure 3A, right panel). Whether subjects time individual intervals separately, or they time total elapsed time is an important question in timing research (Hinton and Rao, 2004; Hinton et al., 2004; Laje et al., 2011; Narkiewicz et al., 2015). We found that the *continuous* (Equation 1) and *reset* (Equation 3) models provided statistically similar fits to our data ( $p = 0.13$ , paired  $t$ -test on the Fisher-transformed correlation coefficients between behavioral data and model estimates,  $t_{(24)} = -1.6$ ; Laje et al., 2011; Figure 3A). For simplicity, our model used Equations 1, 2 to fit the behavioral data.

Fitting was performed with the function *fmincon* in Matlab R2014b by minimizing the error between estimates from the model and the behavioral results, simultaneously for parameters  $p(\text{correct})$ , Std, and constant error (one fit for each interval duration). Because of the difference in scale and measurement units [probability in  $p(\text{correct})$ , seconds in Std, and constant error], these quantities were standardized to values between 0 and 1 before calculating the total fitting error.

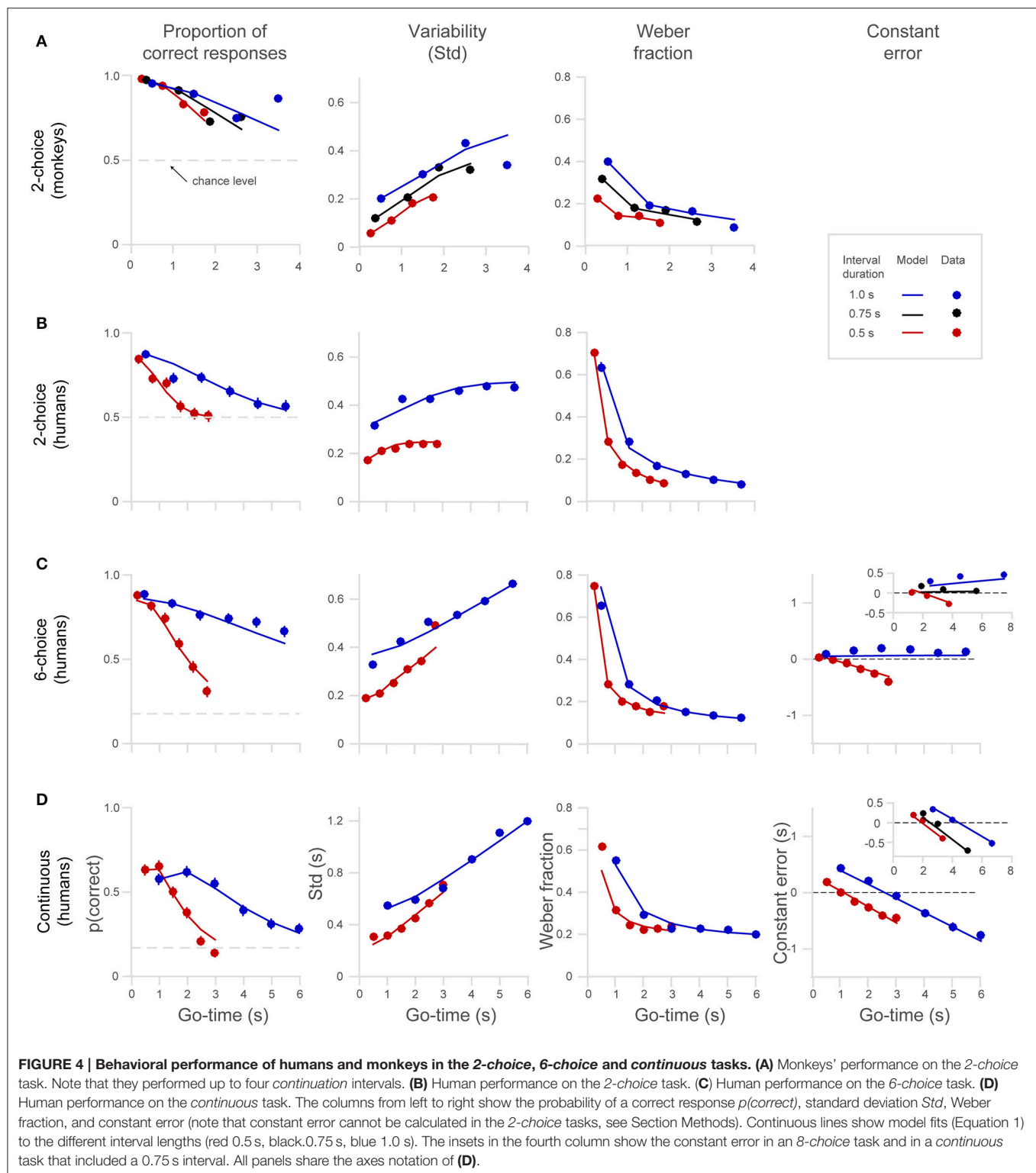
## RESULTS

Humans and monkeys learned to perform the timing tasks, and their behavior showed consistent patterns. We show single subject data for the 2-choice task in Figure 3C and mean data for the different datasets in Figure 4. The proportion of correct responses (Figure 4, first column) decreased as a function of *Go-time*, a trend well captured by our model of the generalized Weber law (continuous lines). Monkeys' performance (Figure 4A) was better than that of humans (Figure 4B) as can be readily appreciated by the larger proportion of correct responses, the lower variability (Std), and the lower Weber fraction. In humans, the proportion of correct responses approached random performance around the 5–6<sup>th</sup> intervals, and the standard deviation saturated at 0.5 and 0.25, the maximum possible values for the 1 and 0.5 s intervals on the 2-choice task (Figure 4B, see Section Methods).

In addition to a better performance, monkeys also showed significantly faster reaction times to the *Go-cue* (Figure 3B,  $p < 0.01$ , two-sample  $t$ -test on the pooled data for humans against the pooled data for monkeys, i.e., all *Go-times*, correct and incorrect responses;  $t_{(38)} = -17.9$ ). It is likely that increased performance and faster reaction times are a consequence of the longer training the monkeys received (Methods and Discussion). Human subjects showed a trend of diminishing reaction times as a function of *Go-time* (linear regression, slope =  $-20$  ms/s,  $p < 0.05$ ), which could reflect the anticipation of trial termination (increasing hazard rate).

Compared to the 6-choice task, the proportion of correct responses in the *continuous* task was significantly lower (Figures 4C,D, to formalize this observation we performed a paired  $t$ -test comparing each  $p(\text{correct})$  across tasks for each *Go-time* and each interval length  $t$ -test,  $t_{(11)} = 11.2$ ,  $p = 2.4e-07$ ). As described in Section Methods the region defining a correct response in the *continuous* task was a window of  $\pm 30^\circ$  around the correct location, comprising a 6<sup>th</sup> of the circle, just as in the 6-choice task. However, it is likely that the larger variability and the resulting lower proportion of correct responses observed in the *continuous* task is explained by the absence of six defined choices. With six defined choices, there is less uncertainty about the correct target position.

As can be observed on the panels of the first column of Figure 4, the decreased proportion of correct responses is more pronounced for the short interval (0.5 s, red dots and lines), and this is observed in all versions of the task (2-, 6-choice, and continuous). To formalize the observation that  $p(\text{correct})$  decreases more rapidly for short intervals (0.5 s) than for long



ones (1.0 s) we compared  $p(\text{correct})$  at similar intermediate *Go-times* for each dataset (one  $p(\text{correct})$  for each interval, i.e., comparison of two proportions for each dataset). We set  $p < 0.01$  and then we corrected for multiple comparisons (Bonferroni

correction, new significant  $p < 0.0025$ ). That  $p(\text{correct})$  decreases more rapidly for fast intervals is an expected trend because the temporal window for a correct response is narrower for short intervals. That is, even if timing variability at a given elapsed time

is equal for short and long intervals, a reduction in the probability of correct response is expected for narrower time intervals.

We observed, as had other studies before, that the Weber fraction is not constant but declines exponentially as a function of time (**Figure 4**, third column; Laje et al., 2011). This trend is explained by the presence of time-independent variability (y-intercept on the Std graphs, term  $\sigma_{indep}^2$  of the model). This basal variability has a large influence at short elapsed times. At longer elapsed times the y-intercept has less impact on the ratio *Std/mean* that defines Weber fraction. The fact that the generalized Weber law satisfactorily fits the behavioral data is strong evidence supporting the presence of time-independent variance in the timing mechanism (**Figure 4**, third column).

## Constant Errors and its Relation to Timing Strategy

In addition to the proportion of correct responses  $p(\text{correct})$ , variability (*Std*), and Weber fraction, the *6-choice* and *continuous* tasks allowed us to estimate the constant error, i.e., the difference between estimated and true elapsed time. This is easily computed by taking into account the direction of disk rotation (clockwise or counterclockwise) and then calculating the difference between the true disk position and the subject's estimated disk position. This angle difference is then expressed in time units. We observed a marked difference in the pattern of errors between the *6-choice* and the *continuous* versions of the task, and this difference can be useful to determine whether subjects are timing each individual interval or total elapsed time.

When the target jumps fast across the six choices (0.5 s interval, red dots, **Figure 4C**), the pattern of negative errors indicates that subjects increasingly lag behind the real target position, thus signifying that the subject's internal chronometer was running slower than the intended pace (**Figure 4C**, last column, red line and dots). Showing the opposite trend, subjects tended to get ahead of a slowly jumping target (1 s interval, blue dots, **Figure 4C**), indicating that their internal chronometer was running faster than the intended 1 s intervals. As can be observed in the insert, the same pattern of errors was observed in an *8-choice* task in which three interval durations were tested (0.5, 0.75, and 1 s). Importantly, the insert shows that the behavioral responses for the middle interval duration (0.75 s) were unbiased, suggesting that the subjects' internal chronometer tends to pace at the rate that is the mean of the distribution of interval durations (Jazayeri and Shadlen, 2010, 2015). We performed a one-way analysis of covariance (ANCOVA) on the mean constant errors and found that slopes are significantly affected by the "interval duration" factor. This analysis also revealed that the slope for the 1 s interval is significantly positive, ( $p < 0.01$ ,  $t = 7.7$ , d.f. = 3; inset on **Figure 4C**, blue dots), the slope of the 0.75 s interval is not significantly different from zero [ $p = 0.61$ ,  $t_{(3)} = 0.6$ ] and finally, that the slope of the 0.5 interval is significantly negative [ $p < 0.01$ ,  $t_{(3)} = -6$ ].

Compared to *6-choice*, the *continuous* task shows a different pattern of errors as can be readily appreciated in the last column of **Figure 4D**. Instead of a bias that progressively accumulates with a positive slope for long intervals and a negative slope

for short ones, what it is observed is that all interval durations generate constant errors with negative slopes. Moreover, for all interval durations, short elapsed times (*Go-time*) generate positive errors while long elapsed times result in negative errors. The same trend can be observed in the insert depicting a *continuous* experiment in which three disk speeds were used (matching the position of the disk in the *6-choice* task at the 0.5 and 1.0 s intervals, with an additional interval of 0.75 s).

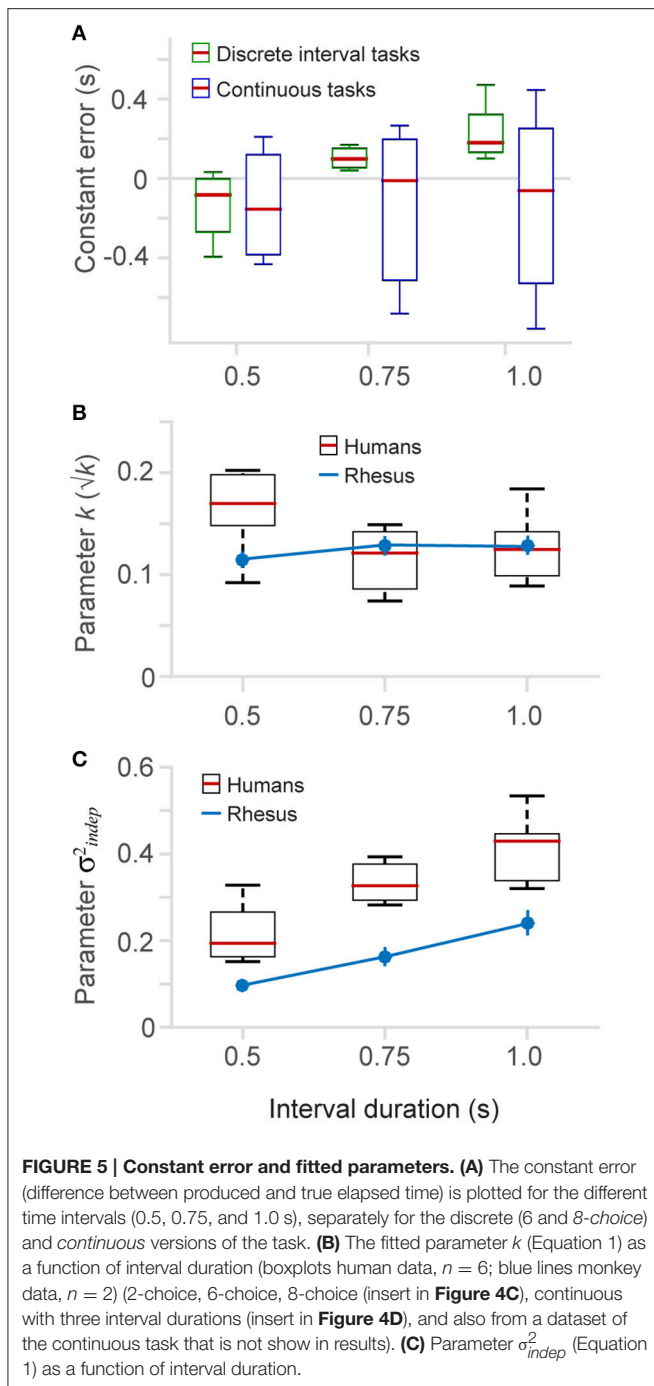
Constant errors are plotted as a function of interval length in **Figure 5A**, separately for the discrete and *continuous* tasks. It can be seen that the constant errors in the discrete tasks (pooled 6- and 8-choice; averaged across *Go-times*) change with a positive slope as interval length increases, whereas in the *continuous* tasks they span both positive and negative values for all interval durations. We conducted a linear regression on each dataset (*continuous* and discrete) and found that constant errors on the discrete task have a significantly positive slope (0.74, [0.47 1.0] 95% C.I., d.f. = 19), and a significantly negative y-intercept (0.4958, [0.7039–0.2877] 95% C.I., d.f. = 19), i.e., they go from negative to positive values as interval duration increases. Conversely, the regression on the *continuous* task shows that the slope and intercept are not statistically different from zero, i.e., they are scattered around zero for the three interval durations (slope =  $-0.02$ , [ $-0.74$  0.71] 95% C.I., d.f. = 19; intercept =  $-0.12$ , [ $-0.68$ , 0.45] 95% C.I., d.f. = 19).

The error patterns differ between the *continuous* and *discrete* tasks, suggesting that in the discrete *8-choice*, and *6-choice* tasks subjects are timing individual intervals and that their estimates are biased toward a mean interval. Conversely, in the *continuous* task the pattern of errors indicates that subjects were timing the total duration of the *continuation* phase, and their time estimates are biased toward the mean total duration (Jazayeri and Shadlen, 2010; Acerbi et al., 2012). Our finding that the *continuous* and *discrete* tasks exhibit different error patterns is important because it allows us to determine whether subjects are timing individual intervals or total elapsed time (see Section Discussion).

## Time Dependent and Time-Independent Variance

The brain might use a single chronometer to time a range of durations or, conversely, make use of different chronometers for different behaviorally relevant intervals. This question can be approached by comparing the classical Weber fraction in long- and short-interval trials, as illustrated in **Figure 4** (third column), and also by comparing the coefficients  $k$  and  $\sigma_{indep}^2$  (Equation 2) resulting from fitting the model to the behavioral data, separately for each time interval. If a single chronometer underlies timing of short and long intervals, we would expect similar Weber fractions and similar  $k$  and  $\sigma_{indep}^2$  values for the different interval durations. Significant differences in these parameters would lend support to the notion that multiple chronometers could be used to time different intervals.

As described by Weber's law, the standard deviation of the timing estimates linearly increases with elapsed time. The human data on the *6-choice* and *continuous* tasks indicate that this increase in variability has different slope and intercept values



for long- and short-interval trials (**Figures 4C,D**, Std graphs). Short-interval trials (0.5 s) have smaller variability but a larger slope, while long-interval trials (1 s) show a larger variability that grows at a lower rate (variability patterns on the 2-choice version of the task are not informative because they have an upper limit at long elapsed times, and this limit is different for long and short intervals, see Section Methods). We found that, the traditional Weber fraction decreases as a function of elapsed time (Laje et al., 2011), and additionally, that short-interval trials show lower Weber fractions for elapsed times up to 3 s.

To quantitatively assess whether variability differs across interval durations (0.5, 0.75, and 1 s) we fit our datasets with the two parameter model (Equation 1) to estimate the  $k$  and  $\sigma_{indep}^2$  parameters. **Figures 5B,C** plot the fitted parameters as a function of interval duration. In humans, we observed a tendency of  $k$  to be larger for the short interval (**Figure 5B**). However, this tendency was not present in the monkey data, indicating either a difference between species or possibly an effect of training on the  $k$  parameter. We speculate that human subjects showed a larger  $k$  parameter because they performed fewer trials of the timing task (as presented next, this is also the case for the  $\sigma_{indep}^2$  parameter, an observation also made by Laje et al., 2011).

Our results show a positive correlation between the  $\sigma_{indep}^2$  parameter and interval duration (**Figure 5C**). Longer time intervals show larger  $\sigma_{indep}^2$ , and this trend is observed in humans as well as in monkeys. Monkeys, however, have lower  $\sigma_{indep}^2$  values, probably due to an effect of additional training and the total number of trials they performed (see Section Methods). We tested this correlation by a linear regression and found that for panel 5B the slopes for monkeys and humans are not statistically different from zero, meaning that there is no influence of the interval length on the  $k$  parameter slope for human data:  $-0.07$ ,  $[-0.17 \ 0.02]$  95% C.I., d.f. = 13; slope for monkey data:  $0.02$ ,  $[-0.20 \ 0.25]$ , 95% C.I., d.f. = 1. For **Figure 5C** we found that both linear regressions have statistically significant positive slopes (slope for human data:  $0.4$ ,  $[0.23 \ 0.57]$  95% C.I., d.f. = 13; slope for monkey data:  $0.28$ ,  $[0.12 \ 0.44]$ , 95% C.I., d.f. = 1), meaning that the basal standard deviation (parameter  $\sigma_{indep}^2$ ) increases as a function of interval duration.

## DISCUSSION

In summary, the main novel observations of the present study are that (1) monkeys were as capable as humans to follow visuo-spatial rhythms with different tempos, and they were able to internally maintain those rhythms without overt movements; (2) both species showed an increase in temporal variability that followed the generalized Weber law, where the time-independent variability changed as a function of the tempo (interval length); and (3) the pattern of constant errors across tempos indicated that human subjects were resetting their clock each interval instead of measuring continuous elapsed time.

## Monkeys and Humans Can Internally Maintain a Temporal Rhythm

Our experiments show that monkeys and humans are able to perceive visuo-spatial rhythms of different paces, and they can internally maintain those rhythms without overt movements. This important finding indicates that rhythm perception and maintenance is a higher cognitive function that we share with other primates and that it does not depend on the execution of motor commands.

The pattern of constant errors (**Figures 4C,D**, last column) calculated from the human data suggests that subjects were timing individual intervals in the discrete task, but total duration



in the *continuous* task. Additionally, timing errors in the 6- and 8-choice tasks show that subjects were lagging behind fast rhythms and getting ahead of slow ones (although the errors in 1s interval of the 6-choice task do not increase linearly they are all positive. The increasing trend is better appreciated in the insert of **Figure 4C**). The fact that a rhythm of intermediate pace generated no bias supports the notion that the timing mechanism calibrates itself to the distribution of interval durations it has to measure, as has been shown by previous research (Jones and McAuley, 2005; Jazayeri and Shadlen, 2010, 2015; Acerbi et al., 2012). Conversely, timing total elapsed time generates a pattern of errors that are positive for short elapsed times and negative for long elapsed times. This suggests that subjects' time estimates were biased toward the mean total duration. We propose that the different patterns of constant errors are a reliable signature that could help to distinguish whether subjects are timing individual intervals or total elapsed time.

The tendency to produce intervals that are closer to the mean is a well-established observation often named the "central-tendency" effect or Vierordt's law (Roy and Christenfeld, 2008; Bangert et al., 2011; Shi et al., 2013). Our results show that in keeping rhythms of different paces the central tendency effect is observed as a bias toward the mean frequency of the rhythms instead of toward the mean total duration. Incorporating prior information such as the mean value of a range of intervals is a mechanism that helps to reduce the effect of noise in time estimation and production, and in our case, rhythm maintenance.

We did not test our monkeys on the continuous task so whether they show the same pattern of errors as human remains an open question. However, it is important to consider that monkeys and humans showed the same patterns of behavioral responses in the 2-choice task, and also the same model satisfactorily accounted for the behavior of human and monkeys.

The question whether subjects time individual intervals or total duration has been addressed before in humans (Hinton and Rao, 2004; Hinton et al., 2004; Laje et al., 2011; Narkiewicz et al., 2015). Buonomano and colleagues used a spatiotemporal task in which subjects had to perform a series of button presses with an elaborated spatial and temporal structure. They found that, although subjects were generating a series of individual intervals, a *continuous* time model was a better fit to their behavioral results. On the contrary, our data from the 6-choice task suggest that subjects were resetting their clocks after each individual time interval. We believe these seemingly contradictory results arise from the different experimental designs. In our rhythm task, subjects could be asked to indicate the position of the target at any given interval (*Go-time*), so they were prepared to generate a behavioral response for each interval. If the *Go* cue didn't arrive by the middle of an interval they had to start timing the next interval and so on. In contrast, on the rhythm task of Buonomano and colleagues subjects had to perform a complete series of intervals for each trial, and this might have compelled them to time total elapsed time. We think that variable *Go-times*, that is, the possibility of terminating the trial at any interval, prompted the subjects to time each interval independently.

It might seem contradictory that the pattern seen in the constant errors suggests that subjects were timing individual intervals whereas the model we fit was based on variance growing with total elapsed time (Equations 1, 2, human data). However, we must note that the difference between a *reset* and a *continuous* model, from the point of view of how variability grows, is a difference in the shape of the curve of *Std* vs. time (**Figure 3A**). The *reset* model predicts that *Std* grows sub-linearly while the *continuous* model predicts a linear increase. We found that, with our current data, these two models could not be distinguished. Our results showed, however, that *continuous* and a *reset* mode of timing could be discerned from the pattern of constant errors (**Figures 4C,D**, last column).

## Basal Variance Depends on Interval Length

Monkeys and humans showed performance parameters well captured by the generalized Weber law. Monkeys, however, showed less timing variability and a higher proportion of correct responses. We speculate that this superior performance is due to the longer training the monkeys received (the monkey dataset was collected after 4–6 months of training). It is likely that increased performance and faster reaction times are a consequence of the longer training the monkeys received. However, it is also possible that differences in reward value and motor planning also contribute to these differences (humans used a mouse cursor while monkeys directly touched the screen to communicate their choices). Previous studies in humans have shown that the Weber fraction quickly decreases after just a few practice sessions (Laje et al., 2011). We speculate that due to their extensive training the Weber fraction of our monkey subjects was at its asymptotic value, but this might not have been the case of our human subjects who performed only one practice session. We expect that with enough training, human subjects could have performed the rhythm task as accurately as the monkey subjects. Our model fittings revealed that humans and monkeys had similar *k*-values (Equations 1, 2, **Figure 5B**), and that the lower variability of the monkeys' time estimates was due mainly to a lower time-independent variance (**Figure 5C**, blue line).

The term  $\sigma_{indep}^2$  showed a tendency to increase as a function of interval duration in both species, indicating that different time intervals have different amounts of time-independent noise. This observation suggests that different chronometers or time mechanisms could time different interval durations. We favor the view that training in timing tasks induces the formation of multiple time templates that match the range and distribution shape of the behaviorally relevant time intervals. Indeed, previous psychophysical and physiological studies support the notion of neural circuits tuned to different interval durations (Nagarajan et al., 1998; Meegan et al., 2000; Bartolo and Merchant, 2009; Merchant et al., 2013; Bartolo et al., 2014). It is also known that timing different types of movement, biological vs. non-biological for example, is performed by different brain structures that can be selectively manipulated (Avanzino et al., 2015), and this is also consistent with the idea that there is no central general-purpose chronometer.

Our data shows that Weber fraction decreases exponentially as a function of elapsed time and that this is due to the  $\sigma_{indep}^2$

term, that is, to the presence of a basal variability (y-intercept on the Std vs. *Go-time* graphs, **Figure 4**). As can be observed in the graphs, the effect of this basal variability reduces at longer elapsed times. However, a recent study in which Grondin and colleagues asked subjects to count at different speeds showed that Weber fraction increased in proportion to the interval length used to subdivide a large total time and that this effect persisted for elapsed times of up to 24 s (Grondin et al., 2015). Their results also showed that mean produced time was always shorter than real elapsed time. Contrary to Grondin's findings, our data predicts that no differences should be observed for long and short subdividing intervals when total elapsed times are larger than  $\sim 3$  s and that errors should not be all negative but instead should be positive for long subdividing intervals and negative for short intervals. We suggest that these differences could be explained by differences in the experimental design. We used interleaved trials in which total elapsed time (*Go-time*) and interval length were pseudo-randomly selected while Grondin and colleagues used a blocked design in which subjects performed the trials of different subdividing intervals in separate sessions. We believe this is an important difference because it has been demonstrated that subjects adjust their internal chronometer according to the distribution of timing intervals they must estimate (Jazayeri and Shadlen, 2010). As was the case in Buonomano's task, subjects in Grondin's experiments had to count up to a predetermined total number of intervals, prompting them to measure total elapsed time.

It is well known that subdividing a long interval into smaller ones decreases the total variance of the estimated elapsed time. Although our task was not designed to explore this phenomenon our results show that subdividing total elapsed time into 0.5 s intervals reduces the timing variability as compared to subdividing with 1 s intervals. This can be observed in **Figure 4C** by comparing the variability of the red and blue lines. We note however, that the beneficial effect of subdividing elapsed time into 0.5 s intervals is limited to total elapsed time of 3–4 s.

It is known that macaque monkeys do not easily entrain to temporal rhythms and that training them in rhythmic tapping tasks might take up to a year (Zarco et al., 2009; Merchant and Honing, 2014; Patel and Iversen, 2014). We speculate that the spatial component of our visuospatial task was an important sensory element that helped the monkeys better perceive and maintain rhythms of different paces. There is evidence that macaques rely more on visual than on auditory cues to control their timing behavior (Zarco et al., 2009; Merchant and Honing, 2014). Nevertheless, the timing behavior of monkeys followed the same pattern of temporal variability and constant errors than humans in a synchronization-continuation tapping task (Zarco et al., 2009).

A possible alternative explanation is that monkeys did not engage the visuo-spatial rhythm but relied instead only on an association between elapsed time and target position. However, we consider this possibility unlikely. The association between *Go-times* and target position was not fixed. In the 2-choice task, for example, the stimulus randomly initiates on the left or the right. In the 6-choice task the stimulus randomly initiates in any of the 6 positions and can rotate either clock wise or counterclockwise.

This variation in initial conditions (remember that interval length and *Go-times* are also selected pseudo-randomly) makes it highly unlikely that subjects were mapping a given *Go-time* with a fixed target position. We would like to mention that, although we only report the behavior in the 2-choice task, monkeys were initially trained in a version of the 6-choice task in which the interval length was chosen from a continuous distribution (300–1200 ms, uniform distribution). During this phase of training, the number of *presentation* intervals was also variable (1–4, uniform distribution). Thus, the stimulus position at any given elapsed time was dependent on (1) the position of the first *presentation* interval, (2) the direction of stimulus rotation, (3) the number of *presentation* intervals, (4) the interval length (chosen randomly from a continuous distribution), and finally (5) the *Go-time* itself. This variation in initial conditions makes it practically impossible for the monkeys to learn all possible *Go-time* and stimulus position combinations and instead encourages them to use the rhythmic motion of the stimulus to predict its future position once it is no longer visible (Coull and Nobre, 2008; Coull, 2009).

The neuronal mechanisms underlying our perception of time and our ability to predict periodic sensory events are not yet completely understood (Roux et al., 2003; Ivry and Spencer, 2004; Eagleman et al., 2005; Coslett et al., 2009; Coull et al., 2011; Wittmann, 2013). It is only recently that the physiological correlates of timing have begun to be systematically investigated in primates (Ghose and Maunsell, 2002; Leon and Shadlen, 2003; Janssen and Shadlen, 2005; Genovesio et al., 2006; Fiorillo et al., 2008; Lebedev et al., 2008; Mita et al., 2009; Machens et al., 2010). It is known that neuronal correlates of timing can be found in parietal, motor, and pre-motor cortices of the primate cerebral cortex (Roux et al., 2003; Merchant et al., 2011; Jazayeri and Shadlen, 2015). These studies revealed distinct groups of neurons whose activity dynamics correlate either with elapsed time from the last motor or sensory event, or with the time remaining to the next motor command.

It is our goal to contribute to the understanding of the neural mechanisms of time estimation and time reproduction. We developed the visuospatial timing task in non-human primates to use it as an experimental model for studying the neuronal correlates of timing. This rhythm task is an ideal experimental setting because it lacks any movement during the *continuation* phase and it will let us study the neuronal correlates of timing without interference by movement or sensory-related activity.

## AUTHOR CONTRIBUTIONS

Vd, OG, and HM conception and design of research; Vd and OG performed experiments; Vd, OG and JC analyzed data; Vd, OG, HM, and JC interpreted results of experiments; Vd prepared figures; Vd, HM, OG, and JC edited and revised manuscript; Vd drafted manuscript.

## ACKNOWLEDGMENTS

We thank Mehrdad Jazayeri for early discussions about the experimental design, Edgar Bolaños and Luis Prado for technical assistance, and Dorothy Pless for proofreading. OG is a doctoral

student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and received fellowship 331516 from Consejo Nacional de Ciencia y Tecnología (CONACYT). This research was supported by

Dirección del Personal Académico de la Universidad Nacional Autónoma de México (HM: IN201214-25, Vd: IN201115) and CONACYT (HM: 236836; Vd: 254313, 247200, Fronteras de la Ciencia 245).

## REFERENCES

- Acerbi, L., Wolpert, D. M., and Vijayakumar, S. (2012). Internal representations of temporal statistics and feedback calibrate motor-sensory interval timing. *PLoS Comput Biol.* 8:e1002771. doi: 10.1371/journal.pcbi.1002771
- Allman, M. J., Teki, S., Griffiths, T. D., and Meck, W. H. (2014). Properties of the internal clock: first-and second-order principles of subjective time. *Ann. Rev. Psychol.* 65, 743–771. doi: 10.1146/annurev-psych-010213-115117
- Avanzino, L., Bove, M., Pelosin, E., Ogliastro, C., Lagravinese, G., and Martino, D. (2015). The cerebellum predicts the temporal consequences of observed motor acts. *PLoS ONE* 10:e0116607. doi: 10.1371/journal.pone.0116607
- Bangert, A. S., Reuter-Lorenz, P. A., and Seidler, R. D. (2011). Dissecting the clock: understanding the mechanisms of timing across tasks and temporal intervals. *Acta Psychol. (Amst)* 136, 20–34. doi: 10.1016/j.actpsy.2010.09.006
- Bartolo, R., and Merchant, H. (2009). Learning and generalization of time production in humans: rules of transfer across modalities and interval durations. *Exp. Brain Res.* 197, 91–100. doi: 10.1007/s00221-009-1895-1
- Bartolo, R., Prado, L., and Merchant, H. (2014). Information processing in the primate basal ganglia during sensory guided and internally driven rhythmic tapping. *J. Neurosci.* 34, 3910–3923. doi: 10.1523/JNEUROSCI.2679-13.2014
- Bispham, J. (2006). Rhythm in music: what is it? who has it? and why?. *Music Percept.* 24, 125–134. doi: 10.1525/mp.2006.24.2.125
- Bizo, L. A., Chu, J. Y., Sanabria, F., and Killeen, P. R. (2006). The failure of Weber's law in time perception and production. *Behav. Processes* 71, 201–210. doi: 10.1016/j.beproc.2005.11.006
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Coslett, H. B., Shenton, J., Dyer, T., and Wiener, M. (2009). Cognitive timing: neuropsychology and anatomic basis. *Brain Res.* 1254, 38–48. doi: 10.1016/j.brainres.2008.11.015
- Coull, J. T. (2009). Neural substrates of mounting temporal expectation. *PLoS Biol.* 7:e1000166. doi: 10.1371/journal.pbio.1000166
- Coull, J. T., Cheng, R. K., and Meck, W. H. (2011). Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology* 36, 3–25. doi: 10.1038/npp.2010.113
- Coull, J. T., and Nobre, A. C. (2008). Dissociating explicit timing from temporal expectation with fMRI. *Curr. Opin. Neurobiol.* 18, 137–144. doi: 10.1016/j.conb.2008.07.011
- Doherty, J. R., Rao, A., Mesulam, M. M., and Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *J. Neurosci.* 25, 8259–8266. doi: 10.1523/JNEUROSCI.1821-05.2005
- Eagleman, D. M., Peter, U. T., Buonomano, D., Janssen, P., Nobre, A. C., and Holcombe, A. O. (2005). Time and the brain: how subjective time relates to neural time. *J. Neurosci.* 25, 10369–10371. doi: 10.1523/JNEUROSCI.3487-05.2005
- Fiorillo, C. D., Newsome, W. T., and Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. *Nat. Neurosci.* 11, 966–973. doi: 10.1038/nn.2159
- Fitch, W. T. (2013). Rhythmic cognition in humans and animals: distinguishing meter and pulse perception. *Front. Syst. Neurosci.* 7:68. doi: 10.3389/fnsys.2013.00068
- Genovesio, A., Tsujimoto, S., and Wise, S. P. (2006). Neuronal activity related to elapsed time in prefrontal cortex. *J. Neurophysiol.* 95, 3281–3285. doi: 10.1152/jn.01011.2005
- Getty, D. J. (1975). Discrimination of short temporal intervals: a comparison of two models. *Percept. Psychophys.* 18, 1–8. doi: 10.3758/BF03199358
- Ghose, G. M., and Maunsell, J. H. (2002). Attentional modulation in visual cortex depends on task timing. *Nature* 419, 616–620. doi: 10.1038/nature01057
- Gibbon, J., Malapani, C., Dale, C. L., and Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: advances and challenges. *Curr. Opin. Neurobiol.* 7, 170–184. doi: 10.1016/S0959-4388(97)80005-0
- Grahn, J. A. (2009). Neuroscientific investigations of musical rhythm: recent advances and future challenges. *Contemp. Music Rev.* 28, 251–277. doi: 10.1080/07494460903404360
- Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: evidence from interval discrimination, reproduction, and categorization. *J. Exp. Psychol.* 38, 880. doi: 10.1037/a0027188
- Grondin, S. (2014). “About the (non) scalar property for time perception,” in *Neurobiology of Interval Timing*, eds H. Merchant and V. de Lafuente (New York, NY: Springer), 17–32.
- Grondin, S., Laflamme, V., and Mioni, G. (2015). Do not count too slowly: evidence for a temporal limitation in short-term memory. *Psychon. Bull. Rev.* 22, 863–868. doi: 10.3758/s13423-014-0740-0
- Grondin, S., Meilleur-Wells, G., and Lachance, R. (1999). When to start explicit counting in a time-intervals discrimination task: a critical point in the timing process of humans. *J. Exp. Psychol.* 25:993. doi: 10.1037/0096-1523.25.4.993
- Hinton, S. C., Harrington, D. L., Binder, J. R., Durgerian, S., and Rao, S. M. (2004). Neural systems supporting timing and chronometric counting: an fMRI study. *Cogn. Brain Res.* 21, 183–192. doi: 10.1016/j.cogbrainres.2004.04.009
- Hinton, S. C., and Rao, S. M. (2004). “One-thousandone... one-thousand two...”: chronometric counting violates the scalar property in interval timing. *Psychon. Bull. Rev.* 11, 24–30. doi: 10.3758/BF03206456
- Ivry, R. B., and Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: evidence for a common timing mechanism. *J. Exp. Psychol.* 21:3. doi: 10.1037/0096-1523.21.1.3
- Ivry, R. B., and Spencer, R. M. (2004). The neural representation of time. *Curr. Opin. Neurobiol.* 14, 225–232. doi: 10.1016/j.conb.2004.03.013
- Janssen, P., and Shadlen, M. N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nat. Neurosci.* 8, 234–241. doi: 10.1038/nn1386
- Jazayeri, M., and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nat. Neurosci.* 13, 1020–1026. doi: 10.1038/nn.2590
- Jazayeri, M., and Shadlen, M. N. (2015). A neural mechanism for sensing and reproducing a time interval. *Curr. Biol.* 25, 2599–2609. doi: 10.1016/j.cub.2015.08.038
- Jones, M. R., and McAuley, J. D. (2005). Time judgments in global temporal contexts. *Percept. Psychophys.* 67, 398–417. doi: 10.3758/BF03193320
- Killeen, P. R., and Weiss, N. A. (1987). Optimal timing and the weber function. *Psychol. Rev.* 94, 455–468. doi: 10.1037/0033-295X.94.4.455
- Laje, R., Cheng, K., and Buonomano, D. V. (2011). Learning of temporal motor patterns: an analysis of continuous versus reset timing. *Front. Integr. Neurosci.* 5:61. doi: 10.3389/fnint.2011.00061
- Lebedev, M. A., O'doherty, J. E., and Nicolelis, M. A. (2008). Decoding of temporal intervals from cortical ensemble activity. *J. Neurophysiol.* 99, 166–186. doi: 10.1152/jn.00734.2007
- Leon, M. I., and Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* 38, 317–327. doi: 10.1016/S0896-6273(03)00185-5
- Lewis, P. A., and Miall, R. C. (2009). The precision of temporal judgement: milliseconds, many minutes, and beyond. *Philos. Trans. R. Soc. Lond. B* 364, 1897–1905. doi: 10.1098/rstb.2009.0020
- Machens, C. K., Romo, R., and Brody, C. D. (2010). Functional, but not anatomical, separation of “what” and “when” in prefrontal cortex. *J. Neurosci.* 30, 350–360. doi: 10.1523/JNEUROSCI.3276-09.2010
- Meegan, D. V., Aslin, R. N., and Jacobs, R. A. (2000). Motor timing learned without motor training. *Nat. Neurosci.* 3, 860–862. doi: 10.1038/78757

- Merchant, H., and de Lafuente, V. (eds.). (2014). "Introduction to the neurobiology of interval timing," in *Neurobiology of Interval Timing* (New York, NY: Springer), 1–13.
- Merchant, H., and Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Front. Neurosci.* 7:274. doi: 10.3389/fnins.2013.00274
- Merchant, H., Perez, O., Zarco, W., and Gamez, J. (2013). Interval tuning in the primate medial premotor cortex as a general timing mechanism. *J. Neurosci.* 33, 9082–9096. doi: 10.1523/JNEUROSCI.5513-12.2013
- Merchant, H., Zarco, W., Pérez, O., Prado, L., and Bartolo, R. (2011). Measuring time with different neural chronometers during a synchronization-continuation task. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19784–19789. doi: 10.1073/pnas.1112933108
- Merchant, H., Zarco, W., and Prado, L. (2008). Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. *J. Neurophysiol.* 99, 939–949. doi: 10.1152/jn.01225.2007
- Mita, A., Mushiaki, H., Shima, K., Matsuzaka, Y., and Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nat. Neurosci.* 12, 502–507. doi: 10.1038/nn.2272
- Nagarajan, S. S., Blake, D. T., Wright, B. A., Byl, N., and Merzenich, M. M. (1998). Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. *J. Neurosci.* 18, 1559–1570.
- Narkiewicz, M., Lambrechts, A., Eichelbaum, F., and Yarrow, K. (2015). Humans don't time subsecond intervals like a stopwatch. *J. Exp. Psychol.* 41, 249. doi: 10.1037/a0038284
- O'Reilly, J. X., Mesulam, M. M., and Nobre, A. C. (2008). The cerebellum predicts the timing of perceptual events. *J. Neurosci.* 28, 2252–2260. doi: 10.1523/JNEUROSCI.2742-07.2008
- Patel, A. D., and Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Front. Syst. Neurosci.* 8:57. doi: 10.3389/fnsys.2014.00057
- Patel, A. D., Iversen, J. R., Bregman, M. R., and Schulz, I. (2009). Studying synchronization to a musical beat in nonhuman animals. *Ann. N.Y. Acad. Sci.* 1169, 459–469. doi: 10.1111/j.1749-6632.2009.04581.x
- Rammsayer, T. H., and Troche, S. J. (2014). In search of the internal structure of the processes underlying interval timing in the sub-second and the second range: a confirmatory factor analysis approach. *Acta Psychol. (Amst)* 147, 68–74. doi: 10.1016/j.actpsy.2013.05.004
- Repp, B. H., and Su, Y. H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon. Bull. Rev.* 20, 403–452. doi: 10.3758/s13423-012-0371-2
- Roux, S., Coulmance, M., and Riehle, A. (2003). Context-related representation of timing processes in monkey motor cortex. *Eur. J. Neurosci.* 18, 1011–1016. doi: 10.1046/j.1460-9568.2003.02792.x
- Roy, M. M., and Christenfeld, N. J. (2008). Effect of task length on remembered and predicted duration. *Psychon. Bull. Rev.* 15, 202–207. doi: 10.3758/PBR.15.1.202
- Shi, Z., Church, R. M., and Meck, W. H. (2013). Bayesian optimization of time perception. *Trends Cogn. Sci. (Regul. Ed.)* 17, 556–564. doi: 10.1016/j.tics.2013.09.009
- Wing, A. M., and Kristofferson, A. B. (1973). Response delays and the timing of discrete motor responses. *Percept. Psychophys.* 14, 5–12. doi: 10.3758/BF03198607
- Wittmann, M. (2013). The inner sense of time: how the brain creates a representation of duration. *Nat. Rev. Neurosci.* 14, 217–223. doi: 10.1038/nrn3452
- Zarco, W., Merchant, H., Prado, L., and Mendez, J. C. (2009). Subsecond timing in primates: comparison of interval production between human subjects and rhesus monkeys. *J. Neurophysiol.* 102, 3191–3202. doi: 10.1152/jn.00066.2009

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer AP and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2016 García-Garibay, Cadena-Valencia, Merchant and de Lafuente. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Meta-Analysis of Functional Neuroimaging and Cognitive Control Studies in Schizophrenia: Preliminary Elucidation of a Core Dysfunctional Timing Network

Irene Alústiza<sup>1,2\*</sup>, Joaquim Radua<sup>3,4,5</sup>, Anton Albajes-Eizagirre<sup>4,5</sup>, Manuel Domínguez<sup>1,2</sup>, Enrique Aubá<sup>1,2</sup> and Felipe Ortuño<sup>1,2</sup>

<sup>1</sup> Department of Psychiatry and Clinical Psychology, Clínica Universidad de Navarra, Pamplona, Spain, <sup>2</sup> Instituto de Investigación Sanitaria de Navarra, Navarra, Spain, <sup>3</sup> Department of Psychosis Studies, Institute of Psychiatry, Kings College, London, UK, <sup>4</sup> FIDMAG Germanes Hospitalaries Hospital Sant Rafael, Barcelona, Spain, <sup>5</sup> Centro de Investigación Biomédica en Red de Salud Mental, Barcelona, Spain

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College, USA

### Reviewed by:

Martin Wiener,  
George Mason University, USA  
Federica Piras,  
Santa Lucia Foundation, Italy

### \*Correspondence:

Irene Alústiza  
ilalustiza@unav.es

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 29 October 2015

**Accepted:** 31 January 2016

**Published:** 17 February 2016

### Citation:

Alústiza I, Radua J,  
Albajes-Eizagirre A, Domínguez M,  
Aubá E and Ortuño F (2016)  
Meta-Analysis of Functional  
Neuroimaging and Cognitive Control  
Studies in Schizophrenia: Preliminary  
Elucidation of a Core Dysfunctional  
Timing Network.  
Front. Psychol. 7:192.  
doi: 10.3389/fpsyg.2016.00192

Timing and other cognitive processes demanding cognitive control become interlinked when there is an increase in the level of difficulty or effort required. Both functions are interrelated and share neuroanatomical bases. A previous meta-analysis of neuroimaging studies found that people with schizophrenia had significantly lower activation, relative to normal controls, of most right hemisphere regions of the time circuit. This finding suggests that a pattern of disconnectivity of this circuit, particularly in the supplementary motor area, is a trait of this mental disease. We hypothesize that a dysfunctional temporal/cognitive control network underlies both cognitive and psychiatric symptoms of schizophrenia and that timing dysfunction is at the root of the cognitive deficits observed. The goal of our study was to look, in schizophrenia patients, for brain structures activated both by execution of cognitive tasks requiring increased effort and by performance of time perception tasks. We conducted a signed differential mapping (SDM) meta-analysis of functional neuroimaging studies in schizophrenia patients assessing the brain response to increasing levels of cognitive difficulty. Then, we performed a multimodal meta-analysis to identify common brain regions in the findings of that SDM meta-analysis and our previously-published activation likelihood estimate (ALE) meta-analysis of neuroimaging of time perception in schizophrenia patients. The current study supports the hypothesis that there exists an overlap between neural structures engaged by both timing tasks and non-temporal cognitive tasks of escalating difficulty in schizophrenia. The implication is that a deficit in timing can be considered as a trait marker of the schizophrenia cognitive profile.

**Keywords:** timing, cognition, neuroimaging studies, cognitive control, schizophrenia, SDM-meta-analysis

## INTRODUCTION

Temporal processing is central to many aspects of human cognition. Accurate judgment of elapsed time is associated with a broad range of activities from relatively basic tasks, such as planning or sequencing, to the higher order processes involved in driving a car, doing sport, playing music etc. According to Navon (1978), time occupies the highest level of the hierarchy of dimensions that forms our perception of the world. In view of the primacy of timing in human cognition, it has been suggested that timing dysfunction lies at the root of deficits (such as planning and aspects of decision making) observed in schizophrenia (Volz et al., 2001; Macar and Vidal, 2009).

That a temporal processing deficit exists in schizophrenia is increasingly recognized on the basis of phenomenological, clinical, and neurobiological observations. Although this deficit was described at the beginning of the last century, it continues to be of interest in current research.

Over the last decade, study of timing in schizophrenia has been fostered by two main factors. First, different models of schizophrenia pathogenesis implicate time perception. For example, Andreasen's theory of cognitive dysmetria (1999) conceptualizes schizophrenia as "*a misconnection in the fluid and coordinated sequences of thought and action stemming from a dysfunction of the cortico-cerebellar-thalamic-cortical circuit.*" Thus, this theory proposes that a disturbance in temporal coordination of information processing may underlie many symptoms of this mental disease (Davalos et al., 2011). Alternatively, Franck et al. (2005) maintained that schizophrenia is related to an excessive temporal integration of events, which leads to classic symptoms. To the degree that psychopathological dimensions of delusions, hallucinations and disorganized speech and behavior can be conceptualized as expressions of dysfunctional neural timing, findings related to such psychopathological dimension are relevant to our understanding of the pathophysiology of schizophrenia. From a phenomenological perspective, schizophrenia can be regarded as a structural breakdown of time consciousness (Vogeley and Kupke, 2007).

The second big reason for research into timing in schizophrenia stems from the idea that the real-life functional difficulties experienced by patients are better accounted for by timing impairment than by dysfunctions in executive control (Volz et al., 2001; Davalos et al., 2003). The potential impact of timing disturbance on cognition and daily behavior is great, and so knowledge of the etiology of timing deficits in schizophrenia may provide important insights into disease pathology.

Controversy remains, however, regarding the existence of a genuine timing disorder in schizophrenia. It is unclear whether disruptions in timing are due to primary disturbances in central temporal processes (perceptual or biological) or to secondary well-known disease-related cognitive impairments that include attention, declarative and working memory or executive functions. Regarding secondary disruptions, deficits associated with a cognitively controlled timing mechanism (for measuring duration in the order of seconds) would be expected to be different from deficits in an automatic mechanism (for

measuring duration at the sub second scale). In fact, performance is found to be equally impaired for both duration ranges, and this suggests that the timing deficit in schizophrenia is essential and primary. The deficit seems to be independent of the length of duration that needs to be timed and also independent of more generalized cognitive impairments (Ciullo et al., 2015).

On the other hand, recent meta-analyses indicate that temporal processing is mediated by a cognitive task requirement (Radua et al., 2014a). Task requirements have an effect on the timing process engaged and on the neural substrate involved, not only in schizophrenia patients but also in normal human timing (Wiener et al., 2010).

Since temporal cognition is a fundamental "basic unit of ability" on which other cognitive and behavioral processes are based (Allman and Meck, 2012), complex cognitive functioning depends on underlying temporal constraints (von Steinbüchel and Pöppel, 1993). Accordingly, temporal processing plays a role in determining a wide range of cognitive processes.

According to the Scalar Expectancy Theory (SET), time processing involves multiple cognitive processes: an internal clock, short- and long- term memory, and decisional processes (Gibbon et al., 1984). In this sense, cognitive processes comprise allocation of attentional resources to the perception and encoding of incoming temporal information, storage and retrieval of the temporal percept in long-term memory, and comparison with other percepts in working memory (Piras et al., 2014). Alterations in any stage or aspect of the system are expected to result in individual and pathophysiological differences. Neuroimaging studies have focused on microanalysis of specific and independent brain networks related to each of the three SET subcomponents (Allman and Meck, 2012).

To the degree that timing is related to other cognitive domains such as attention and working memory, how timing is carried out and the neural networks responsible are relevant to our understanding of healthy cognition (Buhusi and Meck, 2005). The relationship has been proposed to stem from the involvement in timing not only of cortical structures such as the dorsolateral prefrontal cortex but also of other regions, such as the supplementary and pre-supplementary motor areas. The cortical structures are known to play a role in normal cognitive functions, and the motor areas have been identified as crucial for linking cognition to action (Basso et al., 2003). In view of the relationship, there has been increased interest from researchers of schizophrenia in the study of timing.

Timing and other cognitive processes are known to share brain networks (Gómez et al., 2014); some of the cognitive processes involved are attention, automatic or controlled behavior change, working memory, and the degree of concentration required depending on the level of difficulty of the task.

Functional magnetic resonance imaging (fMRI) is well suited to further investigate the nature of timing deficits in schizophrenia since it provides information about the task-related responses across the whole brain. To date there are only a few fMRI studies that examine timing in schizophrenia (Volz et al., 2001; Davalos et al., 2011), but findings suggest that timing deficits in schizophrenia might be due to a combined

impairment of timing mechanisms in the basal ganglia or thalamus and impaired attentional or mnemonic resources organized in prefrontal cortices (Davalos et al., 2011).

Examination of timing in schizophrenia patients, who exhibit cognitive dysfunctions, can be regarded as a valid heuristic approach to explain the essence of the relationship between timing and cognition (i.e., whether the interrelation results from specific co-variation of common temporal processes or from coincidental co-variation in the cognitive components shared by the two functions; Piras et al., 2014).

Cognitive effort is an aspect of every cognitive process and refers to the level of difficulty of the cognitive task and to the consequent mental effort that individuals need to apply to achieve the cognitive aim (Radua et al., 2014a).

Daily tasks demand different levels of cognitive control, and therefore, continuous modulation of the level of effort is needed. Changes in cognitive load require the participation of common cerebral networks.

The neural mechanisms of timing are related to other cognitive functions; cognitive control and accurate executive functioning require the participation of functional and neuroanatomical components of time perception (Radua et al., 2014a). In the context of schizophrenia, neurobiological dysfunctions, or cognitive impairments have been demonstrated to interfere with certain levels of temporal processing, for example, in interval discrimination tasks (Roy et al., 2012).

Given the interrelation between timing and neuropsychological processes, and assuming that pathophysiological distortions in time can depend on and reflect neuropsychological deficits characteristic of neuropsychiatric disorders, the study of timing can be a way to research cognitive dysfunction. Impaired timing has been reported in diseases associated primarily with dopaminergic and fronto-striatal dysfunctions such as in schizophrenia. It has been suggested that the study of timing in schizophrenia can reveal important information on the core cognitive disturbances of this disorder (Matell and Meck, 2004; Wiener et al., 2011).

The study of timing is relevant to our understanding of neurobiological and cognitive abnormalities in schizophrenia. Brain lesion and neuroimaging studies have shown that the cortico-cerebellar-thalamic circuit engaged in temporal processing is involved, in terms of impaired activity coordination among the different brain regions, in the disease's pathophysiology (Andreasen et al., 1999). The cortico-cerebellar-thalamic network involves the bilateral pre-supplementary and supplementary motor area (SMA), the right middle frontal region, the right inferior parietal region, the insula, the left putamen, the right posterior cerebellum, the superior temporal gyrus, the right thalamus, the right middle frontal gyrus, and the left superior temporal gyrus (Volz et al., 2001; Ivry and Spencer, 2004; Ortuño et al., 2011). Hypothetically, cognitive deficits, which in turn lead to impaired timing, can be interpreted as being the result of a disturbance in the functioning of the cortico-striatal pathways; and this same disturbance contributes to a variety of other symptoms associated with schizophrenia (Ward et al., 2011).

We hypothesize that an impaired temporal/cognitive control network underlies the dysfunctional cognition of higher processes in schizophrenia. Our emerging hypothesis is that timing structures are activated either by increased demand on working memory; by the need to shift attention from lower, automatic, levels to higher, controlled, levels; or by certain complex mental operation tasks. Thus, a dysfunctional time estimation network may be linked with other critically impaired functions in schizophrenia.

In the current study, we seek to determine whether schizophrenia patients present a dysfunctional activity pattern in a cognitive control circuit and whether such a pattern matches the pattern involved in timing. To these two ends, we conducted a SDM meta-analysis of published neuroimaging data and then performed a multimodal analysis to identify common brain regions in the findings of that SDM meta-analysis and our previously published activation likelihood estimate (ALE) meta-analysis of neuroimaging of time perception in schizophrenia patients.

## MATERIALS AND METHODS

### Meta-Analysis of Cognitive Difficulty

Two electronic bibliographic databases were searched (PubMed and Web of Science) to identify fMRI studies reporting brain activation patterns associated with changes in cognitive control and effort. This search was limited to literature published between January 2012 and December 2014. Based on preliminary searches, this timeframe was deemed to yield a sufficient number of studies for testing the hypothesis of the present work (please note that we did not intend to conduct an exhaustive meta-analysis). We compare the findings of our SDM meta-analysis to those obtained through our previously published ALE meta-analysis. The previous meta-analysis comprised only three studies, and so inclusion of a disparately large number of studies in the new meta-analysis was not a priority. Keywords were (fMRI) AND (attention OR working memory OR executive functions OR controlled processes) AND (schizophrenia).

Inclusion criteria were: (1) use of a standardized or experimental designed cognitive task; (2) samples composed of healthy volunteers and/or patients with schizophrenia; (3) availability of peak coordinates or statistical parametric maps, either in the published article or after contacting the authors; (4) use of whole brain analyses; (5) use of a constant threshold in the different regions of the brain.

Exclusion criteria were (1) studies from which peak coordinates or statistical parametric maps could not be retrieved from the published article or after contacting the authors; (2) studies whose analyses were limited to specific regions of interest; (3) studies in which different thresholds were used in different regions of the brain; (4) functional neuroimaging studies with techniques other than fMRI (e.g., PET, SPECT); (5) studies that did not specify at least two levels of difficulty of cognitive task or did not use levels with a clear difference in difficulty; (6) studies that considered a resting state or baseline as the lower

level of difficulty; (7) studies based on *Independent Component Analysis* (ICA); (8) case reports, qualitative studies, reviews, and meta-analyses.

No language restrictions were imposed.

Two reviewers independently assessed the studies against the inclusion/exclusion criteria in a standardized manner. Keywords were initially screened in the title and abstract. Afterwards, the full text of eligible studies was analyzed. Any conflicts in reviewers' decisions about inclusion vs. exclusion were resolved through discussion between the two reviewers.

For each selected study, the following information was extracted: number of participants (patients and controls), cognitive tasks and contrasts (**Table 1**), and peak coordinates (MNI or Talarach) and their effect size ( $t$  statistic,  $z$  score,  $p$ -value).

Data were spatially summarized with anisotropic effect-size signed differential mapping software (ES-SDM, <http://www.sdmproject.com>; Radua and Mataix-Cols, 2009; Radua et al., 2011, 2014b), a novel quantitative voxel-based meta-analytic method. First, peak coordinates and their  $t$ -values were used to recreate an effect-size map of the BOLD response for each contrast. These maps included both activations (easy > difficult) and deactivations (difficult > easy; Radua and Mataix-Cols, 2010).

We applied multi-source pre-processing of the data in order to obtain more accurate and thorough recreations of the statistical tridimensional maps of the comparisons between patients and controls for the difficult vs. easy contrast. For each study, we used signed differential mapping (SDM) and the reported peak coordinates and  $t$ -values to separately recreate:

1. the map of the difficult-easy contrast in patients (we will refer to this map as the *patients-only map*), where values were positive for activations (difficult > easy) and negative for deactivations (difficult < easy);
2. the map of the same contrast in healthy controls (we will refer to this map as the *controls-only map*); and
3. the map of the comparison between patients and controls in this contrast (we will refer to this map as the *combined map*), where values were positive for hyperactivations (patients > controls in difficult > easy) or for failures of deactivation (patients < controls in difficult < easy), and negative for hypoactivations (patients < controls in difficult > easy) or hyperdeactivations (patients > controls in difficult < easy).

Results of the pre-processing were inspected to ensure that the recreated maps coincided reasonably well with the results reported in the studies. When two or more contrasts involved overlapping samples, they were combined into a single average map with decreased variance (Rubia et al., 2014; Alegria et al., Submitted).

Next, controls-only maps were subtracted from patients-only maps to obtain *subtraction maps*. This calculation took into account that the maps were not means but  $t$ -values:

$$t_{\text{Patients-Controls}} = \sqrt{\frac{n_{\text{Controls}}}{N}} \cdot t_{\text{Patients}} - \sqrt{\frac{n_{\text{Patients}}}{N}} \cdot t_{\text{Controls}}$$

Note that the recreation of tridimensional maps from peak information requires that the combined maps contain more accurate information in voxels close to the peaks of the differences between groups. Conversely, the subtraction maps have accurate information in voxels close to peaks of activation or deactivation in one or both groups. Thus, a more accurate map can be obtained by merging combined maps with subtraction maps. Such merging consisted in averaging the maps, weighting by the accuracy of each of them:

$$t_{\text{Final}} = w_{\text{Between-groups}} \cdot t_{\text{Between-groups}} + w_{\text{Patients-Controls}} \cdot t_{\text{Patients-Controls}}$$

The weight of a combined map ranged from 1 at peaks to 0 in voxels far from any peak. Similarly, the weight of a subtraction map ranged from 1 at peaks found in both patient and control maps to 0 in voxels far from any patient or control map peak. More specifically, weights were calculated as follows: (a) SDM pre-processing was carried out with all peaks set to 1 to derive the degree of accuracy of each map, (b) averaging of patient and control maps of accuracy was carried out to derive subtraction accuracy maps, and (c) scaling of the combined and subtraction accuracy maps was carried out in order that they sum to unity.

Finally, the effect-size and the effect-size variance maps of all studies were introduced into a meta-analytical random-effects model, which takes intra-study variability, sample-size, and between-study heterogeneity into account. Assessment of statistical significance was based on a distribution-free permutation test (Radua et al., 2011).

## Multimodal Meta-analysis of Cognitive Difficulty and Time Perception

We performed a multimodal meta-analysis to combine the findings from the above-described SDM meta-analysis of studies comparing two levels of cognitive difficulty with those from an ALE meta-analysis on three neuroimaging studies exploring time perception in schizophrenia (see Supplementary Material, **Table 1**). This latter was previously published by our team (Ortuño et al., 2011).

The aim of this multimodal analysis was to detect brain regions that are activated or deactivated by both cognitive difficulty and time perception tasks. We, therefore, overlapped the map of the BOLD response to cognitive difficulty with the map of the BOLD response to time perception. This was conducted using a modification of the probability of the union of the maps (Radua et al., 2013), rather than a simple overlap of them, as the former has been shown to deal with the presence of error in the  $p$ -values of the individual meta-analysis. The combination of the ALE and the SDM meta-analysis was then computed as the union of their probabilities (Radua and Mataix-Cols, 2012). Final results were thresholded with voxel  $p < 0.01$ , peak  $p < 0.001$ , and cluster extent > 10 voxels.

## RESULTS

The search strategy identified 1134 citations. Duplicated papers were removed. From the remaining studies 1091 were excluded



**TABLE 1 | Studies of cognitive control included in our SDM meta-analysis.**

Authors	Sample	Task	Included contrast
1. Anticevic et al., 2012	28 SZ24 HC	Simple perceptual decision task	Negative vs. Neutral distraction
2. Avsar et al., 2013	14 SZ14 HC	Delay-discounting task	Delay-discounting vs. Sensorimotor control; Hard vs. Easy trial difficulty
3. Bender et al., 2013	14 SZ13 HC	Volitional and visually guided saccades task	Simple volitional vs. Visually guided saccade
4. Bjorkquist and Herbener, 2013	14 SZ14 HC	Social perception task	Social vs. Nonsocial images
5. Bleich-Cohen et al., 2014	16 SZ20 HC	N-back WM task	2-back vs. 0-back
6. Das et al., 2012	20 SZ19 HC	ToM task	ToM animation vs. Random animation
7. de la Fuente-Sandoval et al., 2012	12 SZ13 HC	An experimental pain tolerance task	Painful vs. Non-painful thermal stimuli
8. Dowd and Barch, 2012	25 SZ20 HC	Pavlovian reward prediction task	Money cue vs. No money cue
9. Eich et al., 2014	18 SZ18 HC	An item-recognition task	Pre cue vs. Post cue; Lure vs. Control
10. Esslinger et al., 2012	27 SZ27 HC	Monetary reward & face-matching tasks	Monetary vs. Control; Famous vs. Non-famous stimuli
11. Gradin et al., 2013	15 SZ20 HC	Pavlovian reward learning task	Reward vs. No reward
12. Grillon et al., 2013	15 SZ15 HC	Refresh task	Refresh vs. Read
13. Harvey and Lepage, 2014	28 SZ26 HC	A social and nonsocial picture recognition memory task	Old social pictures vs. New social pictures; Old nonsocial vs. New nonsocial pictures
14. Hashimoto et al., 2014	17 SZ17 HC	One-back visual task	Biological motion (BM) vs. Static state (ST); Scrambled motion (SM) vs. ST; BM vs. SM
15. Li et al., 2012	12 SZ12 HC	Facial emotion processing task	Happy vs. Neutral; Fearful vs. Neutral
16. Kauppi et al., 2014	63 SZ118 HC	WM N-back task	2-back vs. 0-back
17. Lakis and Mendrek, 2013	37 SZ37 HC	An emotion processing task	Negative vs. Neutral; Positive vs. Neutral
18. Lee J. et al. (2014)	20 SZ26 HC	A 4-Dot object substitution masking task	Stimulus-onset asynchrony (SOA)1 vs. SOA234
19. Lee J. S. et al., 2014 2014	15 SZ16 HC	Facial expression task	Emotional vs. Meaningless
20. Lesh et al., 2013	43 SZ54 HC	Stroop and AX-CPT	I vs. C; B vs. A
21. Lindner et al., 2014	36 SZ40 HC	Facial processing task	Masked disgust vs. Neutral; Unmasked disgust vs. Neutral
22. Linnman et al., 2013	15 SZ13 HC	A classical conditioning paradigm	Conditioned stimulus (CS)+end vs. CS-end; Unconditioned stimulus (US) vs. CS-end
23. Liu et al., 2014	15 SZ15 HC	Referential task	Self vs. Other
24. Mashal et al., 2013	14 SZ14 HC	Metaphor comprehension task	Novel metaphors vs. Meaningless word pairs; Novel metaphors vs. Conventional metaphors; Novel metaphors vs. Literal expressions
25. Matsumoto et al., 2013	6 SZ6 HC	Rorschach inkblots speech	Between clause vs. Within clause pauses
26. Matsuo et al., 2013	46 SZ46 HC	Sternberg verbal WM task	High load vs. Low load
27. Natsubori et al., 2014	20 SZ20 HC	Visual lexical decision task	Non-words vs. Words
28. Niendam et al., 2014	35 SZ35 HC	Cue phase of AX-CPT	Cue B vs. Cue A
29. Pauly et al., 2013	13 SZ13 HC	Self evaluation task	Other vs. Lexical; Self vs. Lexical; Self vs. Other
30. Pedersen et al., 2012	15 SZ14 HC	"Moving shapes" paradigm	ToM vs. non ToM
31. Ragland et al., 2012	20 SZ19 HC	WM task	Relational (reorder trials) processing vs. Item-specific (rehearse trials) processing
32. Sapara et al., 2014	18 SZ20 HC	N-back WM task	1-back vs. 0-back; 2-back vs. 0-back; 2-back vs. 1-back
33. Shad et al., 2012	17 SZ15 HC	Self-awareness task	Self-directed sentence-stimuli vs. Other-directed sentence-stimuli within the self-referential (SR) cue epoch; Self-directed sentence-stimuli vs. Other-directed sentence-stimuli within the other-referential cue epoch
34. Smieskova et al., 2012	21 SZ20 HC	N-back WM task	2-back vs. 0-back
35. Straube et al., 2013	16 SZ16 HC	Gesture processing task	Metaphoric vs. Iconic
36. Subramaniam et al., 2014	30 SZ15 HC	N-back WM task	2-back vs. 0-back
37. Tully et al., 2014	23 SZ24 HC	Multi-source interference task	Negative vs. Neutral; Neutral interferente vs. Neutral control; Negative interferente vs. Negative control
38. van der Meer et al., 2013	20 SZ20 HC	An emotion regulation task	Reappraise vs. Attend negative
39. van der Meer et al., 2014	47 SZ21 HC	A self-reflection task	Self vs. Semantic; Other vs. Semantic
40. Vercammen et al., 2012	20 SZ23 HC	Verbal emotional go/ no-go task	Inhibit negative vs. Neutral; Inhibit positive vs. Neutral
41. Vercammen et al., 2013	18 SZ22 HC	An emotional go/no-go task	Inhibit negative vs. Inhibit neutral
42. Villalta-Gil et al., 2013	22 SZ31 HC	Facial emotion processing task	Emotions at 50% intensity vs. Neutral emotions; Fearful faces vs. Neutral emotions; Match emotion vs. match gender (neutral faces)
43. Villarreal et al., 2014	14 SZ14 HC	Social functioning tasks	Theory of mind task-eyes (EToM) vs. Test of Adaptive Behavior in Schizophrenia (TABS)

SZ, schizophrenic patients; HC, healthy controls; WM, working memory; ToM, Theory of Mind; CPT, Continuous Performance Task.

because they did not meet the eligibility criteria. A total of 43 studies were included in the meta-analysis. Of these, 14 involve a standardized cognitive task such as N-back, Sternberg, Stroop, or Continuous Performance Test. Basic cognition (such as executive functions, working memory, attention or verbal fluency) is examined in 24 papers; social cognition, in 11; and controlled processes, in the remaining eight studies. Sample size for the included studies ranges from a minimum of six participants for each group to a maximum of 118, with a total participation of 954 schizophrenia patients and 999 healthy volunteers (**Table 1**).

Patients showed hypoactivation in bilateral inferior frontal and superior occipital gyri, right supplementary motor area, left inferior parietal gyri, left cuneus, and red nucleus. Patients also exhibited hyperactivation or failure of deactivation in right postcentral and fusiform gyri (**Figure 1**, **Table 2**).

Jackknife analysis showed that differences between groups in bilateral inferior frontal gyri, right superior occipital gyrus, the right supplementary motor area, and the red nucleus were found in all combinations of studies, indicating a high replicability. Between-group differences in the right postcentral gyrus, the right fusiform gyrus, the left inferior parietal gyrus, the left cuneus, and the left superior occipital gyrus failed to appear in some combinations of studies.

Visual inspection of peak funnel plots did not reveal potential publication bias or other gross abnormalities. The Egger test was only marginally significant in the peak of the red nucleus (4, -26, -6; see Supplementary Material).

Findings are consistent with our team's previously published ALE meta-analysis on neuroimaging of time perception in schizophrenia. This previous work concluded that schizophrenic patients showed, in comparison to healthy controls, significantly lower activation of the right precentral gyrus [Brodmann Area (BA) 6], the superior (BA 9), and middle (BA 8 and 10) frontal gyrus, the left anterior cingulate (BA 32), the right parietal cortex (BA 39), the right putamen and the thalamus (see Supplementary Material; **Figure 1**, **Table 2**).

The results of the multimodal meta-analysis (**Figure 2B**) suggest bilateral overlapping of cortical and subcortical regions: particularly frontal areas (mainly right BA 6), as well as parietal regions and the basal ganglia. The participation of these regions, primarily in the right hemisphere, was reduced in schizophrenic patients relative to control subjects, not only by time perception tasks but also by an increase in the difficulty of non-temporal tasks.

Note that overlapping was only found in those brain regions that were deactivated or hypoactivated by cognitive difficulty. However, the brain regions, which were activated by cognitive difficulty, did not overlap with the map of the BOLD response to time perception.

Together with the overlapping cortical and subcortical regions during both task types, statistically significant activation was found to occur in a group of non-overlapping brain regions (**Figure 2A**): the right thalamus and the left anterior cingulate were specifically activated only in time perception tasks whereas, the bilateral superior occipital gyrus and the right fusiform gyrus were only activated during tasks requiring cognitive effort.

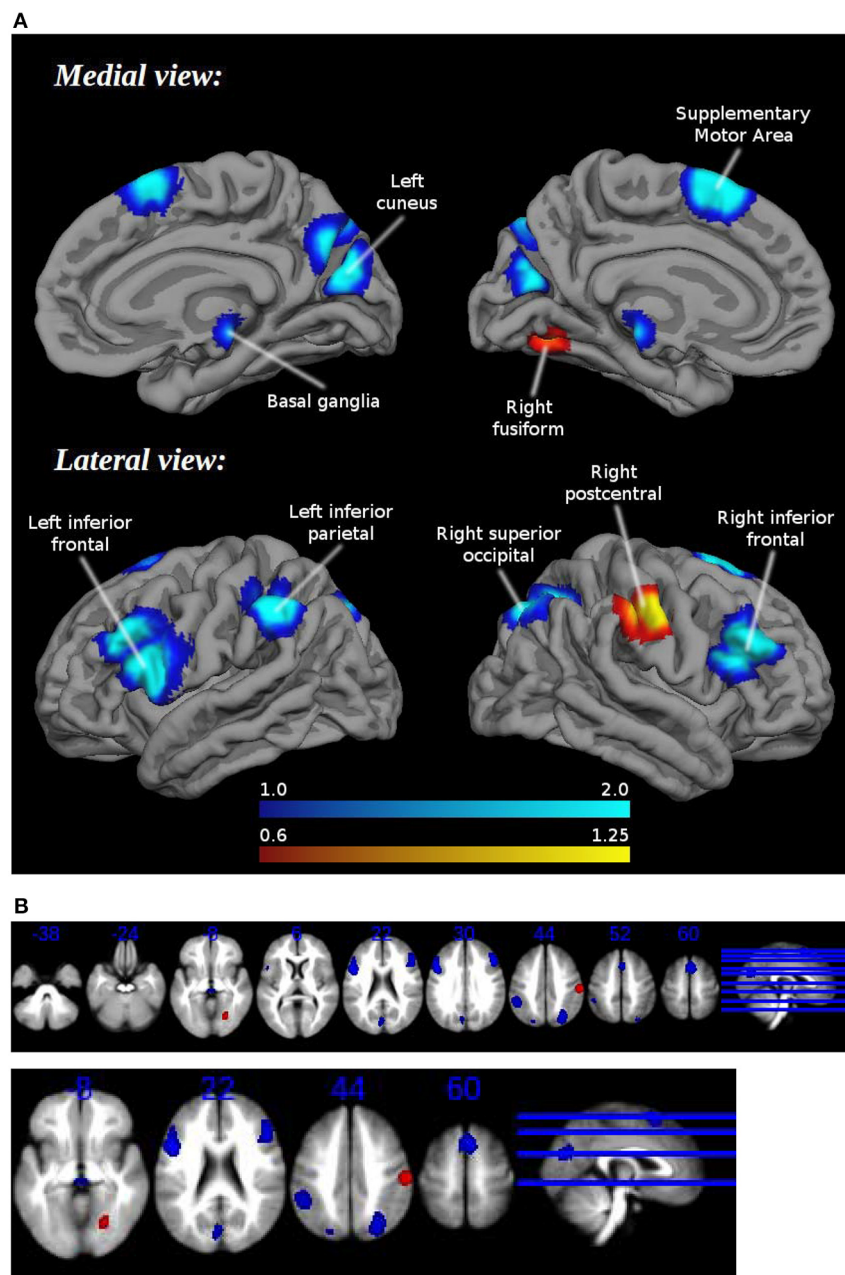
## DISCUSSION

Overall, our findings support the hypothesis that timing structures are activated by an increase in the difficulty of non-temporal cognitive tasks in schizophrenia. The findings are in broad agreement with a recent meta-analysis of functional neuroimaging studies in healthy volunteers (Radua et al., 2014a). Both meta-analyses suggest a partial overlap of cortical and subcortical brain regions engaged in time perception tasks with regions engaged in tasks requiring increased cognitive effort. Specifically, we found a pattern of fronto-parietal and basal ganglia activation common to timing and increased cognitive effort. In schizophrenia patients, the involvement of most of these overlapping cortical and subcortical areas, primarily in the right hemisphere, was reduced in comparison to that in healthy controls.

The involvement in common of some regions by both timing and non-temporal cognitive tasks can be interpreted to indicate that these two functions require similar cognitive abilities. During cognitive tasks with various levels of effort or control, some temporal processing is engaged. Thus, we hypothesize that certain brain regions (such as the insula) traditionally associated with timing are engaged during non-temporal cognitive tasks in response to increases in the level of difficulty. Furthermore, since timing tasks involve different cognitive processes (such as sustained attention, working memory, decision making, or preparation of motor responses), specific brain regions usually associated with these domains (such as the prefrontal cortex and fronto-parietal regions) are hypothesized to be engaged during these tasks.

Another recent meta-analytic study (Niendam et al., 2012) found evidence of a superordinate cognitive control network subserving diverse executive functions. This network involves dorsolateral prefrontal, anterior cingulate, and parietal cortices. The results of our study support the idea that the aforementioned network exists, but they also suggest that the network responds to changes in task demands. With regard to the regions involved, the current meta-analysis coincided in large measure with other studies but indicated that the medial frontal (SMA), temporal insula, and basal ganglia should be included as part of what we propose functions as a *temporal-cognitive control network*.

To date there are only a few published neuroimaging studies of timing in schizophrenia (e.g., Volz et al., 2001; Ojeda et al., 2002; Ortuño et al., 2005; Davalos et al., 2011). We hypothesize, in line with previous theory laid out by Andreasen (1999) and in line with the findings discussed below, that the observed timing impairment displayed in schizophrenia is mediated by a specific fronto-thalamo-striatal dysfunction. A recent functional neuroimaging study (Davalos et al., 2011) that examined the effects of task-difficulty in temporal processing in schizophrenia patients compared to healthy controls found, as we do here, that neuroanatomical regions known to be engaged in timing (SMA, the insula/operculum and striatum) showed signs of dysfunctionality in schizophrenia patients. The higher the level of task difficulty, the greater were found to be the differences in engagement of these



**FIGURE 1 | (A)** Brain regions engaged during tasks requiring cognitive control: differences between healthy controls and schizophrenia patients. **(B)** Brain regions engaged during tasks requiring cognitive control: differences between healthy controls and schizophrenia patients. Multislice.

regions between patients and controls. These findings, however, are not inconsistent with those of a fMRI study of an exclusively healthy population (Tregellas et al., 2006): the authors concluded that activation of certain regions (including SMA, insula/operculum and the striatum) during timing tasks is load-dependent.

What role does the SMA play in dysfunctional temporal processing in schizophrenia? The SMA has been proposed as a key structure during timing (Rao et al., 2001; Macar et al., 2002; Ferrandez et al., 2003; Tregellas et al., 2006) in the “pulse

accumulation” process (Macar et al., 2004), and in attending to an internal timeline against which timing comparisons can be made (Coull et al., 2004). Whilst the role of the SMA is traditionally seen to be purely motor-oriented, a recent review considers that it may be activated by demand for implementation of several cognitive tasks: mental arithmetic, spatial and non-spatial working memory, attention control, silent work production, and conceptual reasoning (Hanakawa et al., 2008). The implication of a dysfunctional SMA is consistent with the idea proposed by Rao et al. (2001) of an early cortical failure related to

**TABLE 2 | Brain regions engaged during tasks requiring cognitive control: differences between healthy controls and schizophrenia patients.**

Location	Peak			Cluster	
	MNI	Z	P	Voxels	Breakdown
Right postcentral gyrus, BA 3	60, -18, 42	1.251	0.000029802	279	Right postcentral gyrus (200), mostly BA 3 Right supramarginal gyrus (71), mostly BA 3/1
Right fusiform gyrus, BA 19	26, -66, -8	1.005	0.000230074	127	Right inferior network, inferior longitudinal fasciculus (52) Right fusiform gyrus (44), mostly BA 19 Right lingual gyrus (26), mostly BA 18
Left inferior frontal gyrus, opercular part	-46, 10, 26	-2.041	0.000001013	927	Left inferior frontal gyrus, opercular part (365), mostly BA 44 Left inferior frontal gyrus, triangular part (318), mostly BA 48 Left precentral gyrus (109), mostly BA 44 Left middle frontal gyrus (61), mostly BA 44 Left frontal aslant tract (35) Left frontal inferior longitudinal fasciculus (16) Corpus callosum (12) Left superior longitudinal fasciculus III (10)
Right inferior frontal gyrus, triangular part, BA 44	48, 26, 28	-1.955	0.000001729	642	Right inferior frontal gyrus, triangular part (331), mostly BA 45 Right middle frontal gyrus (215), mostly BA 45 Right inferior frontal gyrus, opercular part (68), mostly BA 44 Right frontal inferior longitudinal fasciculus (25)
Right superior occipital gyrus, BA 7	28, -68, 44	-1.536	0.000065565	474	Right superior occipital gyrus (192), mostly BA 7 Right angular gyrus (99), mostly BA 7 Right superior parietal gyrus (98), mostly BA 7 Corpus callosum (47) Right superior longitudinal fasciculus II (12)
Right supplementary motor area, BA 6	4, 14, 58	-1.710	0.000013709	462	Right supplementary motor area (259), mostly BA 6 Left supplementary motor area (196), mostly BA 6
Left inferior parietal (excluding supramarginal and angular) gyri, BA 40	-50, -42, 46	-1.392	0.000229657	351	Left inferior parietal (excluding supramarginal and angular) gyri (333), mostly BA 40 Left postcentral gyrus (11), mostly BA 2
Left cuneus cortex, BA 18	-4, -76, 24	-1.210	0.000986218	277	Left cuneus cortex (153), mostly BA 18 Left precuneus (39), mostly BA 7 Left calcarine fissure/surrounding cortex (37), mostly BA 18 Corpus callosum (18) Left median network, cingulum (15)
Basal ganglia	4, -26, -6	-1.078	0.002775669	57	
Left superior occipital gyrus, BA 19	-20, -76, 42	-1.179	0.001261711	53	Left superior parietal gyrus (31), mostly BA 7 Left superior occipital gyrus (22), mostly BA 19

Threshold: voxel  $P < 0.00500$ , peak  $SDM-Z > 1.000$ , cluster extent size  $\geq 10$  voxels. Breakdown regions with  $< 10$  voxels are not reported.

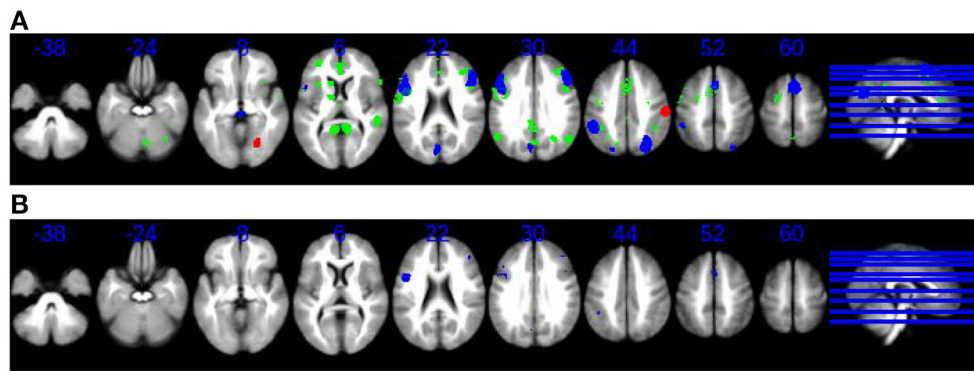
attention disturbances leading to temporal processing deficits in schizophrenia.

In agreement with the Radua et al. meta-analysis (Radua et al., 2014a), the current meta-analysis found the occipital cortex (BA 19) to be a region engaged by tasks requiring cognitive effort. This suggests that this region together with the claustrum is engaged not only in time perception but also in executive functioning.

Since the participation of most of the cortical and subcortical regions primarily in the right hemisphere is reduced relative to healthy subjects, this finding suggests that a pattern of disconnectivity of the timing circuit is a characteristic of the schizophrenia condition (Ortuño et al., 2011).

Owing to the wide overlapping between neural networks involved in high-level cognitive functions and temporal processing, timing performance could be a sensitive measure





**FIGURE 2 | (A)** Overlap and lack of overlap between brain regions engaged during time perception tasks and during tasks requiring cognitive control. Axial slices in neurological convention showing regions with statistically significant activation only during time perception tasks (SDM meta-analysis, green) and regions with statistically significant activation during tasks requiring cognitive control (SDM meta-analysis, blue, and red). Red for hyperactivations (patients > controls in difficult > easy) or for failures of deactivation (patients < controls in difficult < easy), and blue for hypoactivations (patients < controls in difficult > easy) or hyperdeactivations (patients > controls in difficult < easy). **(B)** Overlap and lack of overlap between brain regions engaged during time perception tasks and during tasks requiring cognitive control. Axial slices in neurological convention showing regions with statistically significant activation both during time perception tasks and during tasks requiring cognitive control (blue).

of cognitive functioning and a reliable indicator of impairment to the underlying neural substrate (Piras et al., 2014). In fact, temporal processing has been suggested as a “cognitive primitive,” a fundamental neuropsychological process that has a broad influence on cognition (Fuster et al., 2013).

The common networks that support modulation of effort during non-temporal cognitive tasks also support timing tasks. This finding somewhat belatedly provides backing to Aristotle’s philosophical concepts that timing is related to the perception of change and that time is ubiquitous. As time is omnipresent in the processes of nature, so must time be dealt with by all the higher human cognitive functions (Ortuño and Alústiza, 2014).

It should be noted that the studies we selected for our meta-analysis compared neural activation between two levels of difficulty of their respective experimental tasks. These studies, therefore, reflect how the brain responds to an increase in cognitive load, an increase in the effort required, or an increase in the intensity of what is demanded while the underlying nature of the cognitive function of the task remains essentially the same. The fact that all the studies involve this kind of change in cognitive effort is fundamental to the design of the study and, we believe, critical to the interpretation of our results.

Impaired performance in tests sensitive to different functions (involving the frontal, temporal, hippocampal, parietal, striatal, and cerebellar areas) delineates schizophrenia. In this disease, therefore, there is evidence of a generalized cognitive deficit affecting general neurobiological mechanisms (Gómez et al., 2014). While neuroscience studies indicate that timing-related symptoms are only primary to cognitive impairments and secondary to thought disorders, psychopathological and phenomenological studies strongly imply that disturbance in time perception is the core symptom in schizophrenia.

Difficulty in controlling the involvement of other cognitive domains in temporal processing execution contributes to the continued debate over the specificity of timing dysfunction.

The question is whether the dysfunction is associated with a disturbance in central temporal processes or whether it is attributable to a cognitive or biological dysfunction (Bonnot et al., 2011). It should be noted that the involvement of cognitive brain areas in the discrimination of short (50–500 ms) durations and with automated (pre-conscious) processes is less than that in the discrimination of longer durations and conscious processes.

Three conclusions can be drawn from this study. First, in schizophrenia, there is a widespread network of brain regions (frontal, parietal, and basal ganglia) engaged both in timing tasks and in tasks involving an increase in the cognitive effort demanded for execution of non time-related mental processes. Second, these cerebral circuits, which might be called a *temporal-cognitive control network*, sustain and are common to all mental processes and operations that involve increases (and possibly also decreases) in cognitive load. Lastly, response deficits in this network are highly load-dependent, which suggests that generalized timing deficits in schizophrenia may involve a broad network dysfunction. An important implication of our findings is that the link between a dysfunctional timing network and other impaired cognitive functions only becomes evident when there is comparison of a task performed at different levels of cognitive effort.

A focus on the processing of temporal information offers a way to understand the cognitive deficits of schizophrenia and how these deficits might contribute to a variety of psychiatric symptoms and have an adverse effect on the everyday activities of patients. In this sense, we suggest that a deficit in timing be tentatively considered as a trait marker of the schizophrenia cognitive profile.

Inferences about the dysfunctional overlap observed in the present study are limited by the lack of a way to make an objective assessment of the supposed internal clock. This difficulty has led to dependence in our study on tasks involving both a temporal component and other non time-specific cognitive domains.

It should be noted that the network overlap might be due to a task difficulty effect on neural activation in the time perception studies included.

A methodological note: as far as we know, this study is the first to use the technique of multi-source pre-processing. It is through this technique that the main value of the meta-analysis is established. The main implication thus derived is that, in schizophrenia, the link between a dysfunctional timing network and other impaired functions becomes evident with an increase in the demand for cognitive effort.

It would be interesting to examine whether *temporal-cognitive control network* regions can be attributed to specific cognitive domains accessed by different tasks. Additionally, future research could address the questions of whether timing distortions are a manifestation of, or a mechanism for, cognitive and behavioral symptoms, and whether the relationship applies not only in schizophrenia but also in psychosis in general.

## AUTHOR CONTRIBUTIONS

Each co-author contributed substantially to the manuscript, addressing different tasks. IA contributed to the conception and design of the work, as well as to the acquisition and interpretation of data and drafting the paper. JR and AA contributed to the

design of the work, data analysis and interpretation, and drafting and critical revision of the text in terms of intellectual content. MD contributed mainly in the acquisition of data for the work. FO contributed principally in the conception and design of the study. FO and EA gave final approval of the version to be published and agree to be accountable for all aspects of the work and to ensure that any questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

## FUNDING

Clínica Universidad de Navarra, Pamplona, Navarra, Spain. Instituto de Investigación Sanitaria de Navarra, Navarra, Spain.

## ACKNOWLEDGMENTS

We thank David Burdon for English proofreading.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2016.00192>

## REFERENCES

- Allman, M. J., and Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain* 135, 656–677. doi: 10.1093/brain/awr210
- Andreasen, N. M. (1999). A unitary model of schizophrenia Bleuler's "fragmented phrene" as schizencephaly. *Arch. Gen. Psychiatry* 56, 781–787. doi: 10.1001/archpsyc.56.9.781
- Andreasen, N. C., Nopoulos, P., O'Leary, D. S., Miller, D. D., Wassink, T., and Flaum, M. (1999). Defining the phenotype of Schizophrenia: cognitive dysmetria and its neural mechanisms. *Biol. Psychiatry* 46, 908–920. doi: 10.1016/S0006-3223(99)00152-3
- Anticevic, A., Repovs, G., and Barch, D. M. (2012). Emotion effects on attention, amygdala activation, and functional connectivity in schizophrenia. *Schizophr. Bull.* 38, 967–980. doi: 10.1093/schbul/sbq168
- Avsar, K. B., Weller, R. E., Cox, J. E., Reid, M. A., White, D. M., and Lahti, A. C. (2013). An fMRI investigation of delay discounting in patients with schizophrenia. *Brain Behav.* 3, 384–401. doi: 10.1002/brb3.135
- Basso, G., Nichelli, P., Wharton, C. M., Peterson, M., and Grafman, J. (2003). Distributed neural systems for temporal production: a functional MRI study. *Brain Res. Bull.* 59, 405–411. doi: 10.1016/S0361-9230(02)00941-3
- Bender, J., Reuter, B., Möllers, D., Kaufmann, C., Gallinat, J., and Kathmann, N. (2013). Neural correlates of impaired volitional action control in schizophrenia patients. *Psychophysiology* 50, 872–884. doi: 10.1111/psyp.12060
- Bjorkquist, O. A., and Herbener, E. S. (2013). Social perception in schizophrenia: evidence of temporo-occipital and prefrontal dysfunction. *Psychiatry Res.* 212, 175–182. doi: 10.1016/j.psychres.2012.12.002
- Bleich-Cohen, M., Hendler, T., Weizman, R., Faragian, S., Weizman, A., and Poyurovsky, M. (2014). Working memory dysfunction in schizophrenia patients with obsessive-compulsive symptoms: an fMRI study. *Eur. Psychiatry* 29, 160–166. doi: 10.1016/j.eurpsy.2013.02.004
- Bonnot, O., de Montalembert, M., Kermarrec, S., Botbol, M., Walter, M., and Coulon, N. (2011). Are impairments of time perception in schizophrenia a neglected phenomenon? *J. Physiol. Paris* 105, 164–169. doi: 10.1016/j.jphysparis.2011.07.006
- Buhusi, C. V., and Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6, 755–765. doi: 10.1038/nrn1764
- Ciullo, V., Spalletta, G., Caltagirone, C., Jorge, R. E., and Piras, F. (2015). Explicit time deficit in schizophrenia: systematic review and meta-analysis indicate it is primary and not domain specific. *Schizophr. Bull.* doi: 10.1093/schbul/sbv104. [Epub ahead of print].
- Coull, J. T., Vidal, F., Nazarian, B., and Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science* 303, 1506–1508. doi: 10.1126/science.1091573
- Das, P., Lagopoulos, J., Coulston, C. M., Henderson, A. F., and Malhi, G. S. (2012). Mentalizing impairment in schizophrenia: a functional MRI study. *Schizophr. Res.* 134, 158–164. doi: 10.1016/j.schres.2011.08.019
- Davalos, D. B., Kiseley, M. A., and Ross, R. G. (2003). Effects of interval duration on temporal processing in schizophrenia. *Brain Cogn.* 52, 295–301. doi: 10.1016/S0278-2626(03)00157-X
- Davalos, D. B., Rojas, D. C., and Tregellas, J. R. (2011). Temporal processing in schizophrenia: effects of task-difficulty on behavioral discrimination and neuronal responses. *Schizophr. Res.* 127, 123–130. doi: 10.1016/j.schres.2010.06.020
- de la Fuente-Sandoval, C., Favila, R., Gómez-Martín, D., León-Ortiz, P., and Graff-Guerrero, A. (2012). Neural response to experimental heat pain in stable patients with schizophrenia. *J. Psychiatr. Res.* 46, 128–134. doi: 10.1016/j.jpsychires.2011.09.008
- Dowd, E. C., and Barch, D. M. (2012). Pavlovian reward prediction and receipt in schizophrenia: relationship to anhedonia. *PLoS ONE* 7:e35622. doi: 10.1371/journal.pone.0035622
- Eich, T. S., Nee, D. E., Insel, C., Malapani, C., and Smith, E. E. (2014). Neural correlates of impaired cognitive control over working memory in schizophrenia. *Biol. Psychiatry* 76, 146–153. doi: 10.1016/j.biopsych.2013.09.032
- Esslinger, C., Englisch, S., Inta, D., Rausch, F., Schirmbeck, F., Mier, D., et al. (2012). Ventral striatal activation during attribution of stimulus saliency and reward anticipation is correlated in unmedicated first episode schizophrenia patients. *Schizophr. Res.* 140, 114–121. doi: 10.1016/j.schres.2012.06.025

- Ferrandez, A. M., Hugueville, L., Lehericy, S., Poline, J. B., Marsault, C., and Pouthas, V. (2003). Basal ganglia and supplementary motor area subsecond duration perception: an fMRI study. *Neuroimage* 19, 1532–1544. doi: 10.1016/S1053-8119(03)00159-9
- Franck, N., Posada, A., Pichon, S., and Haggard, P. (2005). Altered time of events in schizophrenia. *J. Nerv. Ment. Dis.* 193, 350–353. doi: 10.1097/01.nmd.0000161699.76032.09
- Fuster, S. M., Kisley, M. A., Davis, H. P., Diede, N. T., Campbell, A. M., and Davalos, D. B. (2013). Cognitive function predicts neural activity associated with pre-attentive temporal processing. *Neuropsychologia* 51, 211–219. doi: 10.1016/j.neuropsychologia.2012.09.017
- Gibbon, J., Church, R. M., and Meck, W. (1984). “Scalar timing in memory,” in *Annals of the New York Academy of Sciences, Vol 423*, eds J. Gibbon and L. Allan (New York, NY: New York Academy of Sciences), 52–77.
- Gómez, J., Marín-Méndez, J., Molero, P., Atakan, Z., and Ortuño, F. (2014). Time perception networks and cognition in schizophrenia: a review and a proposal. *Psychiatry Res.* 220, 737–744. doi: 10.1016/j.psychres.2014.07.048
- Gradin, V. B., Waiter, G., O'Connor, A., Romaniuk, L., Stickle, C., Matthews, K., et al. (2013). Salience network-midbrain dysconnectivity and blunted reward signals in schizophrenia. *Psychiatry Res.* 211, 104–111. doi: 10.1016/j.psychres.2012.06.003
- Grillon, M., Oppenheim, C., Varoquaux, G., Charbonneau, F., Devauchelle, A., Krebs, M., et al. (2013). Hyperfrontality and hypoconnectivity during refreshing in schizophrenia. *Psychiatry Res.* 211, 226–233. doi: 10.1016/j.psychres.2012.09.001
- Hanakawa, T., Dimyan, M. A., and Hallett, M. (2008). Motor planning, imagery, and execution in the distributed motor network: a time-course study with functional MRI. *Cereb. Cortex* 18, 2775–2788. doi: 10.1093/cercor/bhn036
- Harvey, P. O., and Lepage, M. (2014). Neural correlates of recognition memory of social information in people with schizophrenia. *J. Psychiatry Neurosci.* 39, 97–109.
- Hashimoto, N., Toyomaki, A., Hirai, M., Miyamoto, T., Narita, H., Okubo, R., et al. (2014). Absent activation in medial prefrontal cortex and temporoparietal junction but not superior temporal sulcus during the perception of biological motion in schizophrenia: a functional MRI study. *Neuropsychiatr. Dis. Treat.* 10, 2221–2230. doi: 10.2147/NDT.S70074
- Ivry, R. B., and Spencer, R. M. (2004). The neural representation of time. *Curr. Opin. Neurobiol.* 14, 225–232. doi: 10.1016/j.conb.2004.03.013
- Kauppi, K., Westlye, L. T., Tesli, M., Bettella, F., Brandt, C. L., Mattingdal, M., et al. (2014). Polygenic risk for schizophrenia associated with working memory-related prefrontal brain activation in patients with schizophrenia and healthy controls. *Schizophr. Bull.* 41, 736–743. doi: 10.1093/schbul/sbu152
- Lakis, N., and Mendrek, A. (2013). Individuals diagnosed with schizophrenia assign emotional importance to neutral stimuli: an fmri study. *ISRN Psychiatry* 2013:965428. doi: 10.1155/2013/965428
- Lee, J., Cohen, M. S., Engel, S. A., Glahn, D., Nuechterlein, K. H., Wynn, J. K., et al. (2014). Neural substrates of visual masking by object substitution in schizophrenia. *Hum. Brain Mapp.* 35, 4654–4662. doi: 10.1002/hbm.22501
- Lee, J. S., Chun, J. W., Yoon, S. J., Park, H. J., and Kim, J. J. (2014). Involvement of the mirror neuron system in blunted affect in schizophrenia. *Schizophr. Res.* 152, 268–274. doi: 10.1016/j.schres.2013.10.043
- Lesh, T. A., Westphal, A. J., Niendam, T. A., Yoon, J. H., Minzenberg, M. J., Ragland, J. D., et al. (2013). Proactive and reactive cognitive control and dorsolateral prefrontal cortex dysfunction in first episode schizophrenia. *Neuroimage* 2, 590–599. doi: 10.1016/j.nicl.2013.04.010
- Li, H. J., Chan, R. C., Gong, Q. Y., Liu, Y., Liu, S. M., Shum, D., et al. (2012). Facial emotion processing in patients with schizophrenia and their non-psychotic siblings: a functional magnetic resonance imaging study. *Schizophr. Res.* 134, 143–150. doi: 10.1016/j.schres.2011.10.019
- Lindner, C., Dannlowski, U., Walhöfer, K., Rödiger, M., Maisch, B., and Bauer, J. (2014). Social alienation in schizophrenia patients: association with insula responsiveness to facial expressions of disgust. *PLoS ONE* 9:e85014. doi: 10.1371/journal.pone.0085014
- Linnman, C., Coombs, G. III, Goff, D. C., and Holt, D. J. (2013). Lack of insula reactivity to aversive stimuli in schizophrenia. *Schizophr. Res.* 143, 150–157. doi: 10.1016/j.schres.2012.10.038
- Liu, J., Corbera, S., and Wexler, B. (2014). Neural activation abnormalities during self-referential processing in schizophrenia: an fMRI study. *Psychiatry Res.* 222, 165–171. doi: 10.1016/j.psychres.2014.04.003
- Macar, F., Anton, J. L., Bonnet, M., and Vidal, F. (2004). Timing functions of the supplementary motor area: an event-related fMRI study. *Brain Res. Cogn. Brain Res.* 21, 206–215. doi: 10.1016/j.cogbrainres.2004.01.005
- Macar, F., Lejeune, H., Bonnet, M., Ferrara, A., Pouthas, V., Vidal, F., et al. (2002). Activation of the supplementary motor area and of attentional networks during temporal processing. *Exp. Brain Res.* 142, 475–485. doi: 10.1007/s00221-001-0953-0
- Macar, F., and Vidal, F. (2009). Timing processes: an outline of behavioural and neural indices not systematically considered in timing models. *Can. J. Exp. Psychol.* 63, 227–239. doi: 10.1037/a0014457
- Mashal, N., Vishne, T., Laor, N., and Titone, D. (2013). Enhanced left frontal involvement during novel metaphor comprehension in schizophrenia: evidence from functional neuroimaging. *Brain Lang.* 124, 66–74. doi: 10.1016/j.bandl.2012.11.012
- Matsumoto, K., Kircher, T. T., Stokes, P. R., Brammer, M. J., Liddle, P. F., and Maguire, P. K. (2013). Frequency and neural correlates of pauses in patients with formal thought disorder. *Front. Psychiatry* 4:127. doi: 10.3389/fpsy.2013.00127
- Matsuo, K., Chen, S. H., Liu, C. M., Liu, C. C., Hwang, T. J., Hsieh, M. H., et al. (2013). Stable signatures of schizophrenia in the cortical-subcortical-cerebellar network using fMRI of verbal working memory. *Schizophr. Res.* 151, 133–140. doi: 10.1016/j.schres.2013.10.028
- Matell, M. S., and Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Brain Res. Cogn.* 21, 139–170. doi: 10.1016/j.cogbrainres.2004.06.012
- Natsubori, T., Hashimoto, R., Yahata, N., Inoue, H., Takano, Y., Iwashiro, N., et al. (2014). An fMRI study of visual lexical decision in patients with schizophrenia and clinical high-risk individuals. *Schizophr. Res.* 157, 218–224. doi: 10.1016/j.schres.2014.05.027
- Navon, D. (1978). Perception of misoriented words and letter strings. *Can. J. Psychol.* 32, 129–140. doi: 10.1037/h0081681
- Niendam, T. A., Laird, A. R., Ray, K. L., Dean, Y. M., Glahn, D. C., and Carter, C. S. (2012). Meta-analytic evidence for a superordinate cognitive control network subserving diverse executive functions. *Cogn. Affect. Behav. Neurosci.* 12, 241–268. doi: 10.3758/s13415-011-0083-5
- Niendam, T. A., Lesh, T. A., Yoon, J., Westphal, A. J., Hutchison, N., Ragland, J. D., et al. (2014). Impaired context processing as a potential marker of psychosis risk state. *Psychiatry Res.* 221, 13–20. doi: 10.1016/j.psychres.2013.09.001
- Ojeda, N., Ortuño, F., Arbizu, J., López, P., Martí-Climent, J. M., Penuelas, I., et al. (2002). Functional neuroanatomy of sustained attention in schizophrenia: contribution of parietal cortices. *Hum. Brain Mapp.* 17, 116–130. doi: 10.1002/hbm.10055
- Ortuño, F., and Alústiza, I. (2014). Aristotle got it right again! *Med. Hypotheses* 83, 509–510. doi: 10.1016/j.mehy.2014.07.014
- Ortuño, F., Guillén-Grima, F., López-García, P., Gómez, J., and Pla, J. (2011). Functional neural networks of time perception: Challenge and opportunity for schizophrenia research. *Schizophr. Res.* 125, 129–135. doi: 10.1016/j.schres.2010.10.003
- Ortuño, F. M., López, P., Ojeda, N., and Cervera, S. (2005). Dysfunctional supplementary motor area implication during attention and time estimation tasks in schizophrenia: a PET-O15 water study. *Neuroimage* 24, 575–579. doi: 10.1016/j.neuroimage.2004.09.010
- Pauly, K., Finkelmeyer, A., Schneider, F., and Habel, U. (2013). The neural correlates of positive self-evaluation and self-related memory. *Soc. Cogn. Affect. Neurosci.* 8, 878–886. doi: 10.1093/scan/nss086
- Pedersen, A., Koelkebeck, K., Brandt, M., Wee, M., Kueppers, K. A., Kugel, H., et al. (2012). Theory of mind in patients with schizophrenia: is mentalizing delayed? *Schizophr. Res.* 137, 224–229. doi: 10.1016/j.schres.2012.02.022
- Piras, F., Piras, F., Ciullo, V., Danese, E., Caltagirone, C., and Spalletta, G. (2014). Time dysperception perspective for acquired brain injury. *Front. Neurol.* 4:217. doi: 10.3389/fneur.2013.00217
- Radua, J., and Mataix-Cols, D. (2009). Voxel-wise meta-analysis of grey matter changes in obsessive-compulsive disorder. *Br. J. Psychiatry* 195, 393–402. doi: 10.1192/bjp.bp.108.055046
- Radua, J., and Mataix-Cols, D. (2010). Heterogeneity of coordinate-based meta-analyses of neuroimaging data: an example from studies

- in OCD Reply. *Br. J. Psychiatry* 197, 76–77. doi: 10.1192/bjp.197.1.76a
- Radua, J., and Mataix-Cols, D. (2012). Meta-analytic methods for neuroimaging data explained. *Biol. Mood Anxiety Disord.* 2, 1–11. doi: 10.1186/2045-5380-2-6
- Radua, J., Mataix-Cols, D., Phillips, M. J., El-Hage, W., Kronhaus, D. M., Cardoner, N., et al. (2011). A new meta-analytic method for neuroimaging studies that combines reported peak coordinates and statistical parametric maps. *Eur. Psychiatry* 27, 605–611. doi: 10.1016/j.eurpsy.2011.04.001
- Radua, J., Ojeda Del Pozo, N., Gómez, J., Guillén-Grima, F., and Ortuño, F. (2014a). Meta-analysis of functional neuroimaging studies indicates that an increase of cognitive difficulty during executive tasks engages brain regions associated with time perception. *Neuropsychologia* 58, 14–22. doi: 10.1016/j.neuropsychologia.2014.03.016
- Radua, J., Romeo, M., Mataix-Cols, D., and Fusar-Poli, P. (2013). A general approach for combining voxel-based meta-analysis conducted in different neuroimaging modalities. *Curr. Med. Chem.* 20, 462–466.
- Radua, J., Rubia, K., Canales-Rodríguez, E. J., Pomarol-Clotet, E., Fusar-Poli, P., and Mataix-Cols, D. (2014b). Anisotropic kernels for coordinate-based meta-analyses of neuroimaging studies. *Front. Psychiatry* 5:13. doi: 10.3389/fpsy.2014.00013
- Ragland, J. D., Blumenfeld, R. S., Ramsay, I. S., Yonelinas, A., Yoon, J., Solomon, M., et al. (2012). Neural correlates of relational and item-specific encoding during working and long-term memory in schizophrenia. *Neuroimage* 59, 1719–1726. doi: 10.1016/j.neuroimage.2011.08.055
- Rao, S. M., Mayer, A. R., and Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nat. Neurosci.* 4, 317–323. doi: 10.1038/85191
- Roy, M., Grondin, S., and Roy, M. (2012). Time perception disorders are related to working memory impairment in schizophrenia. *Psychiatry Res.* 200, 159–166. doi: 10.1016/j.psychres.2012.06.008
- Rubia, K., Alegria, A. A., Cubillo, A. I., Smith, A. B., Brammer, M. J., and Radua, J. (2014). Effects of stimulants on brain function in attention-deficit/hyperactivity disorder: a systematic review and meta-analysis. *Biol. Psychiatry* 76, 616–628. doi: 10.1016/j.biopsych.2013.10.016
- Sapara, A., Ffytche, D. H., Birchwood, M., Cooke, M. A., Fannon, D., Williams, S., et al. (2014). Preservation and compensation: the functional neuroanatomy of insight and working memory in schizophrenia. *Schizophr. Res.* 152, 201–209. doi: 10.1016/j.schres.2013.11.026
- Shad, M. U., Keshavan, M. S., Steinberg, J. L., Mihalakos, P., Thomas, B. P., Motes, M. A., et al. (2012). Neurobiology of self-awareness in schizophrenia: an fMRI study. *Schizophr. Res.* 138, 113–119. doi: 10.1016/j.schres.2012.03.016
- Smieskova, R., Allen, P., Simon, A., Aston, J., Bendfeldt, K., Drewe, J., et al. (2012). Different duration of at-risk mental state associated with neurofunctional abnormalities. A multimodal imaging study. *Hum. Brain Mapp.* 33, 2281–2294. doi: 10.1002/hbm.21360
- Straube, B., Green, A., Sass, K., Kirner-Veselinovic, A., and Kircher, T. (2013). Neural integration of speech and gesture in schizophrenia: evidence for differential processing of metaphoric gestures. *Hum. Brain Mapp.* 34, 1696–1712. doi: 10.1002/hbm.22015
- Subramaniam, K., Luks, T. L., Garrett, C., Chung, C., Fisher, M., Nagarajan, S., et al. (2014). Intensive cognitive training in schizophrenia enhances working memory and associated prefrontal cortical efficiency in a manner that drives long-term functional gains. *Neuroimage* 99, 281–292. doi: 10.1016/j.neuroimage.2014.05.057
- Tregellas, J. R., Davalos, D. B., and Rojas, D. C. (2006). Effect of task difficulty on the functional anatomy of temporal processing. *Neuroimage* 32, 307–315. doi: 10.1016/j.neuroimage.2006.02.036
- Tully, L. M., Lincoln, S. H., and Hooker, C. I. (2014). Lateral prefrontal cortex activity during cognitive control of emotion predicts response to social stress in schizophrenia. *Neuroimage* 6, 43–53. doi: 10.1016/j.nicl.2014.08.012
- van der Meer, L., de Vos, A. E., Stiekema, A. P., Pijnenborg, G. H., van Tol, M. J., Nolen, W. A., et al. (2013). Insight in schizophrenia: involvement of self-reflection networks? *Schizophr. Bull.* 39, 1288–1295. doi: 10.1093/schbul/sbs122
- van der Meer, L., Swart, M., van der Velde, J., and Pijnenborg, G. (2014). Neural correlates of emotion regulation in patients with schizophrenia and non-affected siblings. *PLoS ONE* 9:e99667. doi: 10.1371/journal.pone.0099667
- Vercammen, A., Morris, R., Green, M. J., Lenroot, R., Kulkarni, J., Carr, V. J., et al. (2012). Reduced neural activity of the prefrontal cognitive control circuitry during response inhibition to negative words in people with schizophrenia. *J. Psychiatry Neurosci.* 37, 379–388. doi: 10.1503/jpn.110088
- Vercammen, A., Skilleter, A. J., Lenroot, R., Catts, S. V., Weickert, C. S., and Weickert, T. W. (2013). Testosterone is inversely related to brain activity during emotional inhibition in schizophrenia. *PLoS ONE* 8:e77496. doi: 10.1371/journal.pone.0077496
- Villalta-Gil, V., Meléndez-Pérez, I., Russell, T., Surguladze, S., Radua, J., Fusté, M., et al. (2013). Functional similarity of facial emotion processing between people with a first episode of psychosis and healthy subjects. *Schizophr. Res.* 149, 35–41. doi: 10.1016/j.schres.2013.06.020
- Villareal, M. F., Drucaroff, L. J., Goldschmidt, M. G., de Achával, D., Costanzo, E. Y., and Castro, M. N. (2014). Pattern of brain activation during social cognitive tasks is related to social competence in siblings discordant for schizophrenia. *J. Psychiatr. Res.* 56, 120–129. doi: 10.1016/j.jpsychires.2014.05.011
- Vogele, K., and Kupke, C. (2007). Disturbances of time consciousness from a phenomenological and a neuroscientific perspective. *Schizophr. Bull.* 33, 157–165. doi: 10.1093/schbul/sbl056
- Volz, H. P., Nenadic, I., Gaser, C., Rammsayer, T., Häger, F., and Sauer, H. (2001). Time estimation in schizophrenia: an fMRI study at adjusted levels of difficulty. *Neuroreport* 12, 313–316. doi: 10.1097/00001756-200102120-00026
- von Steinbüchel, N., and Pöppel, E. (1993). Domains of rehabilitation: a theoretical perspective. *Behav. Brain Res.* 56, 1–10. doi: 10.1016/0166-4328(93)90017-K
- Ward, R. D., Kellendonk, C., Kandel, E. R., and Balsam, P. D. (2011). Timing as a window on cognition in schizophrenia. *Neuropharmacology* 62, 1175–1181. doi: 10.1016/j.neuropharm.2011.04.014
- Wiener, M., Lohoff, F. W., and Coslett, H. B. (2011). Double dissociation of dopamine genes and timing in humans. *J. Cogn. Neurosci.* 23, 2811–2821. doi: 10.1162/jocn.2011.21626
- Wiener, M., Turkeltaub, P., and Coslett, H. B. (2010). The image of time: a voxelwise metaanalysis. *Neuroimage* 49, 1728–1740. doi: 10.1016/j.neuroimage.2009.09.064

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Alústiza, Radua, Albajes-Eizaguirre, Domínguez, Aubá and Ortuño. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.





# Commentary: Effects of psilocybin on time perception and temporal control of behavior in humans

Katarina L. Shebloski<sup>1\*</sup> and James M. Broadway<sup>2</sup>

<sup>1</sup> Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA, USA, <sup>2</sup> Department of Neuroscience, University of New Mexico, Albuquerque, NM, USA

**Keywords:** subjective time perception, temporal processing, psilocybin, 5-HT<sub>2A</sub> receptor, schizophrenia, serotonin, altered states of consciousness

## A commentary on

### Effects of psilocybin on time perception and temporal control of behavior in humans

by Wittmann, M., Carter, O., Hasler, F., Cahn, B. R., Grimberg, U., Spring, P., et al. (2007). *J. Psychopharmacol.* 21, 50–64. doi: 10.1177/0269881106065859

## INTRODUCTION

We discuss Wittmann et al. (2007) “Effects of Psilocybin on Time Perception and Temporal Control of Behavior in Humans,” proposing that altered states of consciousness induced by pharmacological treatments and neurological disorders can reveal much about the circuitry underlying time perception in normal states of consciousness. Further research is needed to integrate these separate research domains.

The brain integrates partial sensory input with internal representations to construct the elaborate story we know as time (Hammond, 2012). Our ordinary experiences reveal the complicated game the mind can play with perceived time: the day drags when we are bored, yet slips through our fingers when distracted or amused. Despite varying phenomenological experiences of time, the appropriate integration of physical time with functional behavior requires a sufficiently accurate perception structured by the conceptual framework of past, present, and future (Eagleman et al., 2005). Working memory, attention, and executive control support this integrated construction (Fuchs, 2007; Marchetti, 2014). Effects on human time perception are observed when these cognitive systems are modulated by pharmacological treatments or psychiatric disorders (González-Maeso and Sealfon, 2009), suggesting the presence of a neurophysiological process that is intrinsic to temporal information processing (Rammsayer, 2008).

## PSILOCYBIN AND TEMPORAL PROCESSING IN NORMAL SUBJECTS

Serotonergic hallucinogens generally slow the perceived flow of time (Shanon, 2003). Pharmacological manipulations using psilocybin have shed light on mechanisms responsible for this distorted time experience. Wittmann et al. (2007) investigated time estimation under the influence of psilocybin. The study addressed the functional role of serotonergic 5-HT<sub>2A</sub> receptors in internal clock models (ICMs) in duration discrimination and temporal control of motor performance. The study revealed a decreased ability to accurately produce intervals longer than 3 s and synchronize finger-tapping to auditory beats separated by more than 2 s. This suggests that effects of psilocybin on temporal processing are specific to relatively long durations, attributable to

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College, USA

### Reviewed by:

Irene Alustiza,  
Clínica Universidad de Navarra, Spain  
Rainer Krähenmann,  
University of Zurich, Switzerland

### \*Correspondence:

Katarina L. Shebloski  
kshebloski1992@gmail.com

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 05 March 2016

**Accepted:** 03 May 2016

**Published:** 19 May 2016

### Citation:

Shebloski KL and Broadway JM  
(2016) Commentary: Effects of  
psilocybin on time perception and  
temporal control of behavior in  
humans. *Front. Psychol.* 7:736.  
doi: 10.3389/fpsyg.2016.00736

memory, and decision-making components of the ICM (Gibbon et al., 1984; Block and Zakay, 1996; Rammsayer, 2008; Allman and Meck, 2012), rather than to more basic pacemaker/accumulator mechanisms (Wittmann et al., 2007).

Comparable results are observed in Wackermann et al. (2008), assessing psilocybin duration reproductions of intervals between 1.5 and 5 s. The analyses rely on the “dual klepsydra” model (DKM), a contemporary alternative to the ICM. In the study, the DKM model is applied to Wittmann et al. (2007) data as well as new data. Results indicate temporal processing influenced by psilocybin are dose-dependent (Wackermann et al., 2008).

Temporal processing of longer durations is impaired in people with schizophrenia (Fuchs, 2005; Bonnot et al., 2011). However, recent meta-analyses suggest that timing deficits in schizophrenia generalize across sub- and supra-second intervals, as well as across perceptual and motor tasks; and are independent from more generalized cognitive impairments (Alústiza et al., 2016; Ciullo et al., 2016). The employment of psychoactive substances may be a useful approach to understanding temporal processing in both the ordinary brain and that which is affected by psychiatric disorders.

## 5-HT<sub>2A</sub> RECEPTORS AND TEMPORAL PROCESSING IN SCHIZOPHRENIA

Psychopharmacological research suggests that drugs such as psilocybin may serve as useful tools for understanding temporal serotonergic signaling mechanisms underlying psychosis, due to their capacity to cause distorted perception in normal subjects (Rammsayer, 2008; González-Maeso and Sealfon, 2009). Modulated 5-HT<sub>2A</sub> receptor agonists may induce clinical symptoms of schizophrenia such as hallucinations, delusion, psychomotor poverty, and distorted perception (Teixeira et al., 2013), including distorted time perception (Allman and Meck, 2012). Pharmaceutical alterations of 5-HT<sub>2A</sub>R activation have shown to assist NMDAR-dependent memory mechanisms (Zhang and Stackman, 2015), and demonstrate that altered time perception is a defining characteristic in schizophrenia due to cognitive changes from NMDA receptor antagonists (Ciullo et al., 2016). Additionally, dopamine-release manipulations cause motor and cognitive defects seen in schizophrenia (Raote et al., 2007), and impair duration discrimination in healthy subjects (Wittmann, 2009). Likewise, schizophrenia is associated with poor accumulation of signal durations derived from impairments in sensory integration (Allman and Meck, 2012; Teixeira et al., 2013). Sysoeva et al. (2010) found that genotypes characterized by higher 5-HT transmission exemplify a higher “loss rate” of duration representation, which may correlate to the very high 5-HT<sub>2A</sub> R occupancy in the prefrontal cortex of schizophrenic patients (Zhang and Stackman, 2015).

## REFERENCES

- Allman, M. J., and Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain*. 135, 656–677. doi: 10.1093/brain/awr210

Impairments in working memory, selective attention, and executive control, as seen in schizophrenia, lead to distorted sequencing and integration of past, present, and future into a personal narrative. Carter et al. (2005) demonstrate a reduction in attentional tracking abilities affected by psilocybin, and implicate 5-HT receptors in these processes through pretreatment of the 5-HT<sub>2A</sub> receptor antagonist ketanserin.

## PROPOSAL FOR INTEGRATIVE RESEARCH

5-HT<sub>2A</sub> receptor activity is associated with time distortion in both psychiatric disorders and hallucinogenic experiences. Manipulating antagonists/agonists provides an approach to utilizing psychoactive drugs as tools in research for understanding time perception in the ordinary brain. It would be fruitful to compare healthy subjects under the influence of psilocybin with patients with acute schizophrenia, utilizing a common paradigm as in Wittmann et al. (2007). However, Wittmann et al. (2007) excludes significant moderating factors of time estimation: attention and emotion (Droit-Volet and Meck, 2007). An fMRI test of acute treatment with psilocybin in healthy volunteers found decreased amygdala reactivity during emotion processing (Kraehenmann et al., 2015). Negative pictures led to an overestimation of duration, indicating greater attention allotted to emotional valence (Wittmann, 2009).

Neuroimaging techniques combined with psychophysical tests of time perception (for a review see Grondin, 2010), including manipulations to assess attentional and emotional factors, will illuminate neural activity responsible for temporal processing in schizophrenia and psychedelic perceptions. Comparing performance and brain activities in these altered states with those of untreated healthy subjects under the same experimental conditions will elucidate mechanisms underlying time perception.

## CONCLUSION

Slowing of perceived time is induced by psilocybin and schizophrenia; having a common basis in 5-HT<sub>2A</sub> receptor activities. Commonalities across pharmacological treatments and neurological disorders should be explored within a common experimental paradigm to better understand neurochemical processes mediating temporal processing in ordinary states.

## AUTHOR CONTRIBUTIONS

KS obtained comprehensive research and drafted original article, with collaboration of JB's efforts. JB provided substantial advisory of research materials and writing processes. Both authors edited and revised the article, insured the integrity of the product, and agreed on the final version for submission.

- Alústiza, I., Radula, J., Albajes-Eizagirre, A., Domíngue, M., Aubá, E., and Ortuño, F. (2016). Meta-analysis of functional neuroimaging and cognitive control studies in schizophrenia: Preliminary elucidation of a core dysifunctional timing network. *Front. Psychol.* 7:192. doi: 10.3389/fpsyg.2016.00192

- Block, R. A., and Zakay, D. (1996). "Models of psychological time revisited," in *Time and Mind*, ed H. Helfrich (Seattle, WA: Hogrefe and Huber), 171–195.
- Bonnot, O., Montalembert, M., Kermarrec, S., Botbol, M., Walter, M., and Coulon, N. (2011). Are impairments of time perception in schizophrenia a neglected phenomenon? *J. Physiol. Paris* 105, 164–169. doi: 10.1016/j.jphysparis.2011.07.006
- Carter, O. L., Burr, D. C., Pettigrew, J. D., Wallis, G. M., Hasler, F., and Vollenweider, F. X. (2005). Using psilocybin to investigate the relationship between attention, working memory, and the serotonin 1A and 2A receptors. *J. Cogn. Neurosci.* 17, 1497–1508. doi: 10.1162/089892905774597191
- Ciullo, V., Spalletta, G., Caltagirone, C., Jorge, R. E., and Piras, F. (2016). Explicit time deficit in schizophrenia: Systematic review and meta-analysis indicate it is primary and not domain specific. *Schizophrenia Bulletin* 42, 505–518. doi: 10.1093/schbul/sbv104
- Droit-Volet, S., and Meck, W. H. (2007). How emotions colour our perception of time. *Trends Cogn. Sci.* 11, 504–513. doi: 10.1016/j.tics.2007.09.008
- Eagleman, D. M., Tse, P. U., Buonomano, D., Janssen, P., Nobre, A. C., and Holcombe, A. O. (2005). How subjective time relates to neural time. *J. Neurosci.* 25, 45. doi: 10.1523/JNEUROSCI.3487-05.2005
- Fuchs, T. (2005). Corporealized and disembodied minds a phenomenological view of the body in melancholia and schizophrenia. *Philos. Psychiatr. Psychol.* 12, 95–107.
- Fuchs, T. (2007). The temporal structure of intentionality and its disturbance in schizophrenia. *Psychopathology* 40, 229–235. doi: 10.1159/000101365
- Gibbon, J., Church, R. M., and Meck, W. H. (1984). Scalar timing in memory. *Ann. N. Y. Acad. Sci.* 423, 52–77. doi: 10.1111/j.1749-6632.1984.tb23417.x
- González-Maeso, J., and Sealfon, S. C. (2009). Psychedelics and schizophrenia. *Trends Neurosci.* 32, 225–232. doi: 10.1016/j.tins.2008.12.005
- Gronin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Atten. Percept. Psychophys.* 72, 561–582. doi: 10.3758/APP.72.3.561
- Hammond, C. (2012). *Time Warped: Unlocking the Mysteries of Time Perception*. Toronto, ON: House of Anansi Press.
- Kraehenmann, R., Preller, K. H., Scheidegger, M., Pokorny, T., Bosch, G. O., Seifritz, E., et al. (2015). Psilocybin-induced decrease in amygdala reactivity correlates with enhanced positive mood in healthy volunteers. *Biol. Psychiatry* 78, 572–581. doi: 10.1016/j.biopsych.2014.04.010
- Marchetti, G. (2014). Attention and working memory: two basic mechanisms for constructing temporal experiences. *Front. Psychol.* 5:880. doi: 10.3389/fpsyg.2014.00880
- Rammsayer, T. H. (2008). *Neuropharmacological Approaches to Human Timing. Psychology of Time*. Bingley: Emerald Group Publishing Ltd.
- Raote, I., Bhattacharya, A., and Panicker, M. M. (2007). "Serotonin 2A (5-HT<sub>2A</sub>) receptor function: ligand-dependent mechanisms and pathways," in *Serotonin Receptors in Neurobiology*. Available online at: <http://www.ncbi.nlm.nih.gov/books/NBK1853/>
- Shanon, B. (2003). Altered states and the study of consciousness- the case of Ayahuasca. *J. Mind Behav.* 24, 125–154.
- Sysoeva, O. V., Tonevitsky, A. G., and Wackermann, J. (2010). Genetic determinants of time perception mediated by the serotonergic system. *PLoS ONE* 5:e12650. doi: 10.1371/journal.pone.0012650
- Teixeira, S., Machado, S., Paes, F., Velasques, B., Silva, J. G., Sanfim, A. L., et al. (2013). Time perception distortion in neuropsychiatric and neurological disorders. *CNS Neurol. Disord. Drug Targets* 12, 567–582. doi: 10.2174/18715273113129990080
- Wackermann, J., Wittmann, M., Hasler, F., and Vollenweider, F. X. (2008). Effects of varied doses of psilocybin on time interval reproduction in human subjects. *Neurosci. Lett.* 435, 51–55. doi: 10.1016/j.neulet.2008.02.006
- Wittmann, M. (2009). The inner experience of time. *Philos. Trans. R. Soc.* 364, 1955–1967. doi: 10.1098/rstb.2009.0003
- Wittmann, M., Carter, O., Hasler, F., Cahn, B. R., Grimberg, U., Spring, P., et al. (2007). Effects of psilocybin on time perception and temporal control of behavior in humans. *J. Psychopharmacol.* 21, 50–64. doi: 10.1177/0269881106065859
- Zhang, G., and Stackman, R. W. (2015). The role of serotonin 5-HT<sub>2A</sub> receptors in memory and cognition. *Front. Pharmacol.* 6:225. doi: 10.3389/fphar.2015.00225

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Shebloski and Broadway. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



# When and How-Long: A Unified Approach for Time Perception

Michail Maniadakis\* and Panos Trahalias

Computational Vision and Robotics Laboratory, Institute of Computer Science, Foundation for Research and Technology Hellas, Heraklion, Greece

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College, USA

### Reviewed by:

Julien Vitay,  
Chemnitz University of Technology,  
Germany

Tadeusz Wladyslaw Kononowicz,  
Université Paris-Sud, France

### \*Correspondence:

Michail Maniadakis  
mmaniada@ics.forth.gr

### Specialty section:

This article was submitted to  
Perception Science,  
a section of the journal  
Frontiers in Psychology

**Received:** 30 October 2015

**Accepted:** 16 March 2016

**Published:** 31 March 2016

### Citation:

Maniadakis M and Trahalias P (2016)  
When and How-Long: A Unified  
Approach for Time Perception.  
Front. Psychol. 7:466.  
doi: 10.3389/fpsyg.2016.00466

The representation of the environment assumes the encoding of four basic dimensions in the brain, that is the 3D space and time. The vital role of time for cognition is a topic that recently attracted increasing research interest. Surprisingly, the scientific community investigating mind-time interactions has mainly focused on interval timing, paying less attention on the encoding and processing of distant moments. The present work highlights two basic capacities that are necessary for developing temporal cognition in artificial systems. In particular, the seamless integration of agents in the environment assumes they are able to consider *when* events have occurred and *how-long* they have lasted. This information, although rather standard in humans, is largely missing from artificial cognitive systems. In this work we consider how a time perception model that is based on neural networks and the Striatal Beat Frequency (SBF) theory is extended in a way that besides the duration of events, facilitates the encoding of the time of occurrence in memory. The extended model is capable to support skills assumed in temporal cognition and answer time-related questions about the unfolded events.

**Keywords:** time perception and timing, temporal distance, past perception model, when, how long, computational modeling, temporal cognition

## INTRODUCTION

Our sense of time exhibits unique characteristics that distinguishes it from the typical group of human senses (sight, hearing, touch, smell, and taste). A crucial difference is that the sense of time is not associated with a specific sensory system in the brain. As it is noted in Bruss and Ruschendorf (2010), the perception of time seems different in nature from what we usually understand as perception. It seems to have its own ways and own laws. Since we cannot stop time, we cannot experience a moment twice. In the contrary, we can hear a sound, view a light, taste a food as many times as we want.

In an attempt to understand the unique characteristics of time perception, the recent years, a significant amount of research studies have been devoted on understanding the brain mechanisms that enable experiencing and processing time, with controversial theories attempting to explain experimental observations. Broadly speaking, there are two main approaches to describe how our brain represents duration (Ivry and Schlerf, 2008; Buetti, 2011). The first is the dedicated approach (also known as extrinsic, or centralized) that assumes an explicit metric of time. This is the oldest and most influential explanation on interval timing. The models included in this category employ mechanisms that are designed specifically to represent duration. Traditionally such models follow an information processing perspective in which pulses that are emitted regularly by a pacemaker are temporally stored in an accumulator, similar to



a clock (Gibbon et al., 1984; Droit-Volet et al., 2007). This has inspired the subsequent pacemaker approach that uses oscillations to represent clock ticks (Miall, 1989; Large, 2008). Following a broader consideration, the Striatal Beat Frequency (SBF) model assumes timing to be the results of the coincidental activation of basal ganglia neurons by cortical neural oscillators (Matell and Meck, 2004; Meck et al., 2008). Other dedicated models assume monotonous increasing or decreasing processes to encode elapsed time (Staddon and Higa, 1999; Simen et al., 2011). The second approach includes intrinsic explanations (also known as distributed) that describe time as a general and inherent property of neural dynamics (Dragoi et al., 2003; Karmarkar and Buonomano, 2007). According to this approach, time is intrinsically encoded in the activity of general purpose networks of neurons. Thus, rather than using a time-dedicated neural circuit, time coexists with the representation and processing of other external stimuli. Recent models combine intrinsic and dedicated representations into active oscillations that do not only produce “ticks” but additionally adjust their characteristics to perceive, measure, and process time in order to facilitate the accomplishment of a variety of temporal tasks (Maniadakis and Trahanias, 2015). Similarly, models assuming oscillations with adaptive pulse rates extend the classic pacemaker-accumulator model to accomplish timescale invariance in interval timing (Simen et al., 2013).

The aforementioned models focus on estimating the duration of events (i.e., *how-long*), without typically paying much attention on the time of occurrence of events (i.e., *when*), as an important temporal information. The combined consideration of these two temporal aspects is vital for understanding the evolved phenomena in the environment in a rich and meaningful way. While interval timing is typically related to short-term time perception, considering *when* events have occurred is mostly related to the perception of mid and distant past. It is now believed in the timing community that the short-term duration perception mechanisms in the brain are different than those involved in the long-term, past perception (Aschoff, 1985; Rammsayer, 1999; Lewis and Miall, 2003).

However, given that the present is included in the entire timeline linking the past and the future, it is reasonable to assume a connection between the short- and long-term time perception. Along this line, the present work investigates the possibility that a universal time source may support both aspects of time perception. This is the focus of the present study which explores possible means for combining *when* and *how-long* in a single cognitive system. This does not aim to argue that the two mechanisms coincide or overlap. The subsystems of short- and long-term time perception are kept separate but it is possible that they share common timing inputs and in that sense we are interested to explore their possible bridging. It is noted that in order to explore the long- and short-term aspects of time perception, the implemented models must consider both the moments experienced during the occurrence of events and the moments passing without being associated to the given event. These two time periods exhibit very different characteristics as we will discuss in the following sections.

The present work adopts a memory encoding perspective to explore the possible mechanisms supporting *how-long* and *when* temporal cognition. Interestingly, besides providing an explanation on how the two times related cognitive capacities may be linked, the present work accomplishes a crucial milestone for introducing time perception in artificial systems enabling the later to consider the inherent temporal dimension of human-machine symbiotic interaction.

The composite model is developed following an incremental procedure. We start by implementing a neural network model that is capable to estimate and memorize the duration of simple tone-events. The model is implemented using a “black box” artificial coevolutionary procedure that tunes system components and enforces their cooperation. Subsequently, we consider the possibility of extending the model with the ability of keeping track the time of occurrence of the underlying events. We explore whether the previously implemented mechanism of time flow perception that is used for interval timing can be also employed for encoding when events have occurred. Moreover, we explore the option that past perception may use temporal distance measures as indicators of past times. Our experiments show that a single time source can facilitate encoding of both the duration and the time of occurrence of events.

## ARTIFICIAL EVOLUTION OF INTERVAL TIMING MODEL

To develop a brain inspired duration perception system, we borrow ideas from the Striatal Beat Frequency (SBF) model (Matell and Meck, 2004; Meck et al., 2008) that is one of the most widely referenced paradigm explaining interval timing in the brain. The model assumes that durations are coded by the coincidental activation of a large number of cortical neurons projecting onto spiny neurons in the striatum that respond to timing patterns. The present work explores a very simple version of the SBF model using only a small number of input oscillatory signals. The goal of this simplified model is not to compete against the original SBF model, but rather to suggest a new direction for using interval timing models.

### Modeling

We employ the coevolutionary neural network framework that has been described in detail in (Maniadakis and Trahanias, 2008, 2009) to develop a modular neural network system for interval timing. In the past, we have used the same technology to develop cognitive models for artificial agents, which have been capable of time-informed behavior switching (Maniadakis et al., 2009) and multi-context duration processing (Maniadakis and Trahanias, 2015).

The structure of the neural network model is shown in **Figure 1**. In short Continuous Time Recurrent Neural Networks (CTRNNs) are used as modules to develop a composite cognitive system. CTRNNs represent knowledge in terms of internal neurodynamic attractors and it is therefore

particularly appropriate for implementing cognitive capacity that is inherently continuous, similar to time perception. The neurons implementing CTRNN components are governed by the standard leaky integrator equation:

$$\frac{dy_i}{dt} = \frac{1}{\tau} \left( -\gamma_i + \sum_{k=1}^R w_{ik}^s I_k + \sum_{m=1}^N w_{im}^p A_m \right) \quad (1)$$

where  $\gamma_i$  is the state (cell potential) of the  $i$ -th neuron. All neurons in a network share the same time constant  $\tau = 0.25$  in order to avoid explicit differentiation in the functionality of CTRNN parts. The state of each neuron is updated according to external sensory input  $I$  weighted by  $w_s$ , and the activity of presynaptic neurons  $A$  weighted by  $w_p$ . After estimating the state of neurons based on the above equation, the activation of the  $i$ -th neuron is calculated by the non-linear sigmoid function according to:

$$A_i = \frac{1}{1 + e^{-(\gamma_i - \theta_i)}} \quad (2)$$

where  $\theta_i$  is the activation bias applied on the  $i$ -th neuron. The model considered in the present study assumes 16 neurons for the building blocks tSen1, tSen2, and 2 neurons for the blocks implementing t-Duration1, ..., t-Duration6. A hierarchical coevolutionary procedure is used as a mechanism for tuning CTRNN modules, specifying synaptic weights and activation bias of neurons.

Following the assumption of fusing cortical neural oscillators for implementing sense of time, we use four oscillatory signals at different frequencies as inputs to the model. The use of such a small number of oscillatory inputs keeps manageable the complexity of the model providing at the same time the opportunity to obtain insight in the dynamics self-organized internally in the CTRNN. The oscillatory signals used in the current study are as follows:

$$\begin{aligned} \text{Inp}_1 &= \sin(4t + k_1) + u(-0.05, 0.05) \\ \text{Inp}_2 &= \sin(t + k_2) + u(-0.05, 0.05) \\ \text{Inp}_3 &= \sin(0.25t + k_3) + u(-0.05, 0.05) \\ \text{Inp}_4 &= \sin(0.1t + k_4) + u(-0.05, 0.05) \end{aligned} \quad (3)$$

Parameters  $k_1, k_2, k_3, k_4 \in [0, \pi]$ , implement random time shifts initialized at the beginning of every experimental session (i.e., different values are assumed for each evolutionary run, see below). Additive noise implemented as a uniform distribution in the range  $[-0.05, 0.05]$  aims to improve generalization of the internal representation of time and thus enable robust and accurate duration estimation.

Each temporal moment processed by the model is associated to one simulation step. Interestingly, assuming that one simulation step corresponds to 5–10 ms, the input signals described in Equation (3) can be associated to the known frequencies of cortical neural oscillations, from the 1–4 Hz of the delta band, up to the 30–70 Hz for the gamma band. It is noted that, for years, it has been hard to identify a single frequency band dominating temporal processing, (Treisman, 1984; Wiener

and Kanai, 2016). However, modern approaches assume that the combination of bands might be the key for explaining sense of time (for a discussion, see Kononowicz and van Wassenhove, 2016 in the present Research Topic). Such an assumption provides added value to our model, which combines oscillations at very different frequencies to develop sense of time. However, the input signals considered in the present study were not originally designed with cortical oscillation bands in mind, and thus we would like to avoid building further on this assumption. Besides targeting interval timing in the range of a few seconds, the model does not assume an explicit correspondence between simulation steps and the known metrics of physical time (e.g., ms, or sec). The main goal of the present work has been the development of a brain-inspired time perception system for robotic agents engaged in long-term symbiotic interaction with humans.

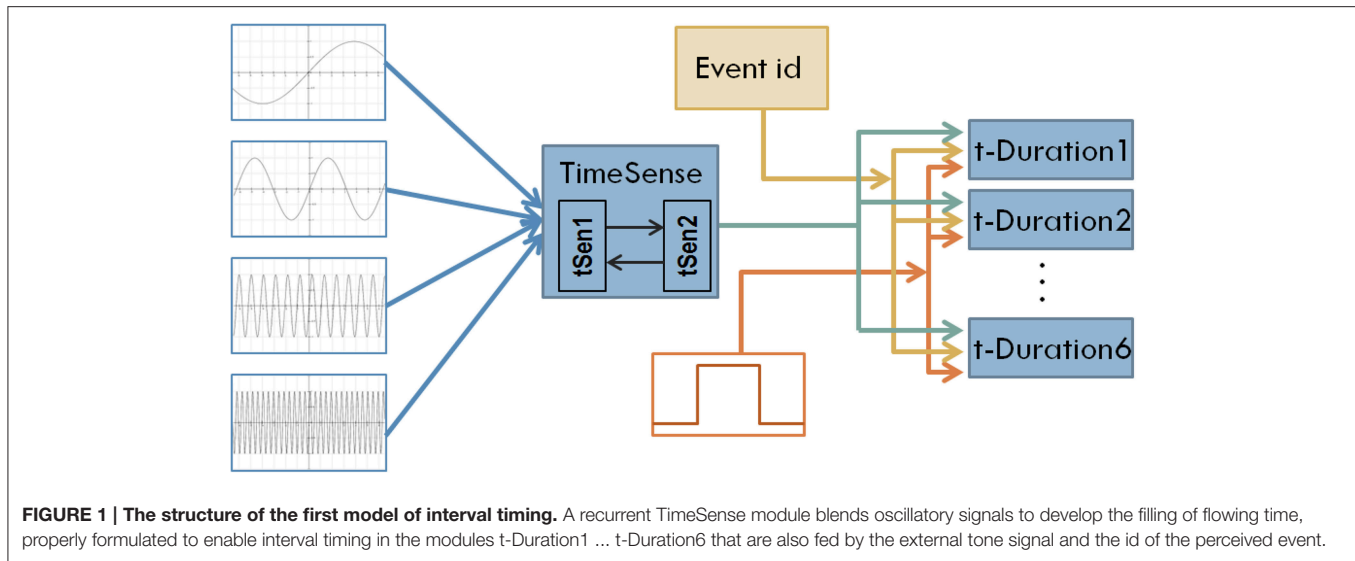
Turning back to **Figure 1**, oscillatory inputs project into a composite TimeSense module consisting of two recurrently connected sub-modules. The TimeSense module aims at gradually transforming oscillatory inputs to a composite time flow representation that is adequate for interval timing. To facilitate the applicability of the model in robotic applications, we use working memory to store the temporal properties of a small number of recently experienced events. In the current implementation we explore scenarios assuming the random occurrence of six events (the capacity of working memory) in a session of 1000 simulation steps. We employ 6 different duration estimation modules each one devoted to the perception of one tone-event. The duration estimation modules receive a binary tone input that represents the occurrence of events. Tones have randomly specified lengths that represent the duration of events. A binary signal representing the unique ID of the event enables differentiating the measured interval lengths. The actual duration of events is randomly specified every time a NN model is loaded and tested. We enforce a minimum distance of 100 moments between consecutive events.

## Parametric Tuning

The training of the model is achieved using Hierarchical Cooperative CoEvolution as described in (Maniadakis and Trahanias, 2008, 2009). By using this “black box” coevolutionary scheme we are able to consider the specialized characteristics of each component in the model and additionally enforce their synergetic functionality to accomplish the desired overall performance for the composite system.

We assume a brain-like encoding of interval timing. More specifically, a ramp-like encoding of time has been identified in the brain of monkeys (Leon and Shadlen, 2003; Maimon and Assad, 2006; Mita et al., 2009) for durations up to a few hundreds of milliseconds. The proposed model abstracts these findings by implementing a similar ramping mechanism for short-term interval timing, aiming mainly to support robotic applications.

Error-based functions are used to evaluate the performance of each event-specific module tDur1, ..., tDur6. In particular, the desired output of the module associated to the  $i$ -th event starting at time  $st_i$  and finishing at time  $et_i$ , having a maximum duration



$M$  (i.e.,  $e_i - st_i < M$ ) equals to:

$$D_i(t) = \begin{cases} 0, & t < st_i \\ (t - st_i) / M, & st_i \leq t < e_i \\ (e_i - st_i) / M, & e_i \leq t \end{cases} \quad (4)$$

In the current study we investigate events with maximum duration  $M = 50$  moments. The function that measures the success of the  $i$ -th temporal duration module is:

$$EDur_i = \sum_t (out_i(t) - D_i(t))^2 \quad (5)$$

This is the key component of the fitness function  $ff_i$  that drives the evolution of the corresponding module accomplishing parameter tuning:

$$ff_i = \frac{1000}{EDur_i} \quad (6)$$

Higher values of  $ff_i$  indicate better performance of the  $i$ -th duration module.

To accomplish parametric tuning for the neural network modules representing the components t-Sen1 and t-Sen2, which have a supportive role for all duration estimation modules, we employ a mixture of the afore mentioned fitness functions, described by:

$$ff = \prod_i ff_i \quad (7)$$

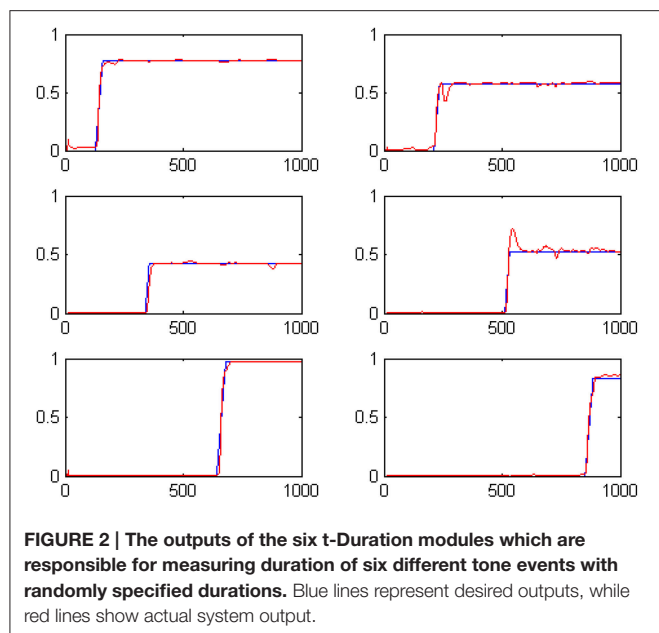
The hierarchical cooperative coevolutionary procedure (Maniadakis and Trahanias, 2008, 2009) accomplishes parametric tuning and optimization of component modules, enforcing their collaborative performance toward a successful composite model. We use one population of 1000 artificial chromosomes for each CTRNN module considered in the model. Each chromosome, encodes a different configuration

of the module. We combine candidate module configurations to develop full configurations of the complete system, which are tested on the duration estimation task described above. The 20% of the best performing chromosomes in each population are selected for reproduction following single point crossover. Mutation is applied on new chromosomes with a probability 2% for each encoded parameter. Mutation is implemented as additive noise in the range  $[-10\%, 10\%]$  relative to the previous value of the parameter.

## Results

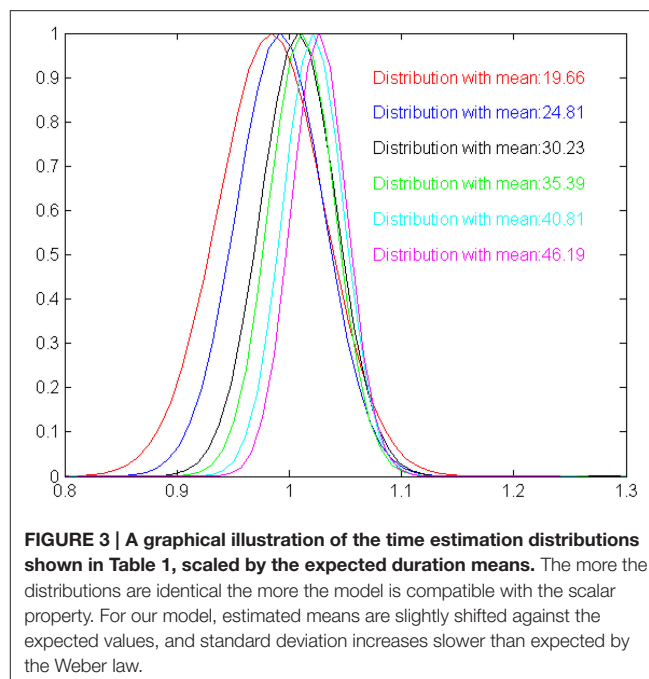
We have evolved the above described coevolutionary scheme for 500 generations, producing a successfully tuned CTRNN model for interval timing. An indicative set of results for six randomly specified tone events is shown in **Figure 2**. The fact that numerous event durations can be simultaneously preserved in the system is a valuable addition to interval timing models that enables further processing of the memorized durations. In particular, it has been straight forward to use a Multi-layer Perceptrons (MLPs) to develop decision making systems capable of comparing any two of the memorized durations to accomplish duration comparison tasks similar to those studied in (Droit-Volet et al., 2010).

Besides extensively testing the model with randomly specified interval times up to  $M$  simulation steps, we explore whether the output of the model exhibits the scalar characteristics that are typical observed in biological timing mechanisms (Lejeune and Wearden, 2006). Scalar timing implies that (i) measurements should vary linearly and near-accurately as time increases and (ii) the variance of perceptual mechanism increases as the duration of time also increases. To get an estimate of the scalar characteristics of the model, we have studied its ability to correctly estimate durations of 20, 25, 30, 35, 40, and 45 moments (without this limiting the model to perform successfully for in-between durations). For each one of the six durations considered here, we perform 50 statistically independent runs, feeding the model with



randomly initialized oscillatory inputs. The mean and standard deviation for each one of the durations considered are shown in **Table 1**. Clearly, the average of the estimated intervals remains close to the true time in all cases, satisfying mean accuracy. The variance increases as the model experiences longer intervals, however, in a rate that is slower to the increase of the mean. The scalar property assumes a constant coefficient of variation (the ratio of the standard deviation to the mean), which is not true for our model. This is depicted more clearly in **Figure 3**, where relevant output distributions are scaled by the expected duration value. Even if the model is not fully compatible with the scalar property, **Table 1** shows that the output of the model is sufficiently accurate for making the model usable in robotic systems. Nevertheless, it is worth emphasizing that, currently, the two main characteristics of the scalar property have been self-organized without any explicit instructions by the modeler. Therefore, it seems valid to assume that our model can be easily rendered fully compatible to the scalar property, by introducing a constraint for a constant coefficient of variation in the fitness function of the evolutionary design procedure.

The notably small variations in time estimations shown in **Table 1** (we remind it summarizes the results of 50 randomly initialized runs of the model) indicate that the implemented model is particularly tolerant to the noise added in the oscillatory input. To further assess model robustness, we have explored the performance of the model against different levels of input noise. Results are summarized in **Table 2**. The model shows to perform satisfactorily for input noise up to the range  $[-0.07, 0.07]$ . More noise than that significantly affects the estimation of durations for specific events. In particular, noise in the range  $[-0.09, 0.09]$  often results into a single mismeasured event, noise in the range  $[-0.11, 0.11]$  results into more than two mismeasured events (on average 2.3), and noise in the range  $[-0.13, 0.13]$  results, into nearly random duration measurements. Practically, the increase of noise affects the performance of the TimeSense modules which



in turn introduces disturbances (i.e., occasional picks) in the corresponding ramping activities therefore destroying accurate interval timing.

Interestingly, in the case that the noise is added to the input signal for a relative short time (e.g.,  $<10$  simulation steps) the performance of the model remains largely unaffected, even for noise in the range  $[-0.13, 0.13]$ . This is explained by the use of leaky integrator neurons which smooth out the strong but temporally-short noise, enabling the model to quickly recover into the normal mode of operation.

While previous SBF models have been particularly sensitive to sensory noise (Matell and Meck, 2004; Gu et al., 2015), the model implemented in the current work exhibits more robust performance, therefore enabling interval timing in noisy environments. This is a particularly desirable feature that is developed for free in the model due to the noise included in the oscillatory sensory inputs and the randomness introduced in the experimental setup. This is mainly because we do not artificially describe coincidental activation of oscillatory inputs, but we let the neural network self-organize the fusion of inputs. Fitness assignment favors the more robust neural networks which filter out noise and estimate durations that are closer to the target. Therefore, the evolutionary procedure produces solutions that are gradually more tolerant to noise. However, it is worth emphasizing that sensory noise has been shown to facilitate time scale invariance in the case of a large number of input neural oscillators (Oprisan and Buhusi, 2014).

## EXTEND THE MODEL TO ADDRESS “WHEN”

The model described above has been able to accomplish accurate interval timing in a series of randomly initialized binary events.



**TABLE 1 | Studying the scalar properties of the model.**

Actual time	20	25	30	35	40	45
Estimated time—Mean	19.66	24.81	30.23	35.39	40.81	46.19
Estimated time—STD	0.87	0.93	1.04	1.09	1.17	1.21

The model does not assume a direct relationship between simulation time and physical time. The number of simulation steps is used as indicator of the elapsed time.

**TABLE 2 | Model performance against different levels of input noise.**

Noise range	[−0.03, 0.03]	[−0.05,0.05]	[−0.07,0.07]	[−0.09,0.09]	[−0.11,0.11]	[−0.13,0.13]
Estimated time average error	0.47	0.53	0.72	6.38	10.13	18.67

With this timing mechanism at hand it is particularly interesting to explore, whether we can achieve other temporal cognition skills beyond interval timing. In the current study we explore if it is possible to use the previously developed timing mechanism as a base for encoding information related to the time of occurrence of events, that is to represent time moments in the distant past. While estimating the duration of an event requires the active perception of the external stimulus, keeping track of *when* that event occurred assumes perceiving time that is not anymore related to the underlying event, filtering also out any other external input that may appear in the meanwhile. This is an important qualitative difference that distinguishes *when* and *how-long* perception.

There are two alternative options for encoding when an event has occurred in the past. The first assumes a coordinate system centered on “now,” e.g., “John was here one hour ago.” Following this approach the center of the coordinate system is non-static but it is moved together with the flow of time, causing a continuous increase in the time elapsed from the occurrence of the event until now (i.e., in a while, the above statement will change to “John was here two hours ago” and so on). The other alternative assumes a timeline centered on a predefined moment that is assumed to represent the zero-point and all time moments are perceived relative to that particular zero-point. For example, most western cultures assume as zero-point the birth of Jesus Christ and thus dates are typically measured as distances from this point, e.g., “I met John on February 10, 2015”.

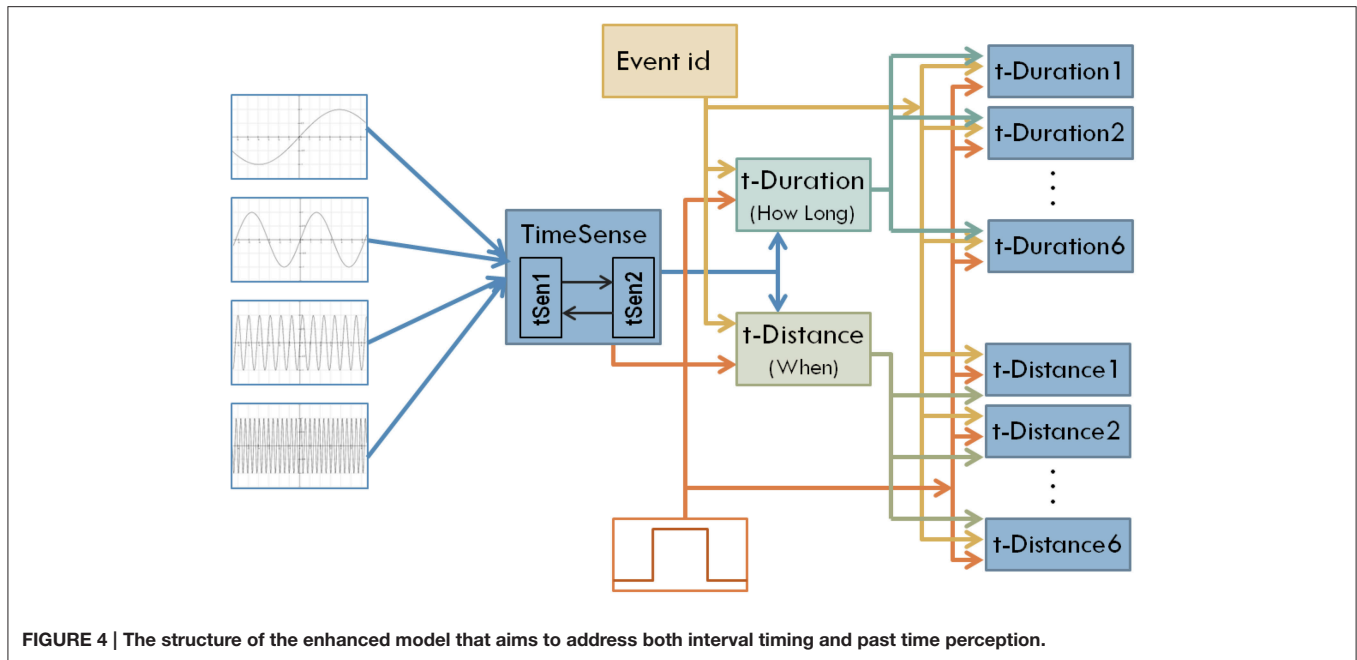
Human adults can equally perceive both alternative options. However, it seems more likely that the development of the past perception for young children starts centered on “now”. This is because even if infants are capable to perceive time very early in their life (Droit-Volet, 2011), the conceptual development of an objective zero point develops not earlier than the middle childhood (Friedman, 2005). The now-centered perception of time is further supported by developmental studies showing a decline in the accuracy of children responses with increasing distances to the past (Friedman, 1998) and the fact that children have autobiographical memories before they learn how to use clocks and calendars (Campbell, 1997). Finally, from a numerical point of view, young children seem to slowly develop the concept of ordinal relationship between small values which gradually develops to the understanding of the broader number line

(Gallistel and Gelman, 1992; Rouder and Geary, 2014). The above suggest that a first, basic approach for representing when events have occurred should be implemented relative to “now” rather than relative to a fixed point in time. The latter option may be developed at a following stage as a higher level capacity that processes encoded events.

Interestingly, the now-centered representation of the timeline suggests that the duration perception mechanisms may have a key role in the representation of past times. To elaborate further on this assumption, we borrow from the past perception literature (Arzy et al., 2009; Wyer et al., 2010) the term “temporal distance,” which describes the temporal properties of past events in relation to the present. We implement the computational analogous of temporal distance in our model, and we investigate the possibility of using this measure as a representation of when events have occurred in the past.

In particular, we extend the model discussed in Section Artificial Evolution of Interval Timing Model to additionally incorporate the capacity of memorizing the times of events’ occurrence based on the assumption of encoding temporal distances to the present. In that sense, the composite model will work in two different time scales (i) up to 50 simulation steps for the *how-long* mode and (ii) up to 1000 simulation steps for the *when* mode. The revised model will be capable of using a single sense of time to derive both the duration of events and their time of occurrence.

Interestingly the coevolutionary framework used in the current work is particularly appropriate for the incremental modification and enhancement of modular neural network models (Maniadakis and Trahanias, 2009). To incorporate distant time perception, a set of neural network components is integrated into the model as shown in **Figure 4**. Two central components aim to transform general purpose sense of time to a form that is appropriate for measuring duration (t-Duration module) and temporal distance (t-Distance module). Similar to the earlier version, we use dedicated modules t-Duration1, t-Duration2 ... t-Duration6 to memorize durations and modules t-Distance1, t-Distance2 ... t-Distance6 to memorize temporal distances for the six tone-events considered in the current experimental setup. The CTRNN-based implementation of the modules assumes 16 neurons for the building blocks tSen1, tSen2, t-Duration, t-Distance, and 2 neurons for the blocks



**FIGURE 4 |** The structure of the enhanced model that aims to address both interval timing and past time perception.

implementing t-Duration1, ..., t-Duration6, and t-Distance1, ..., t-Distance6.

A key issue for implementing temporal distances regards the representation of time in the long-term. There is classical debate on psychophysics asking whether the humans perceive the time-line in a linear or a logarithmic basis. Without any restriction<sup>1</sup> the present work adopts the assumption of a logarithmic representation of distant time which is supported by recent experimental data (Arzy et al., 2009; Glicksohn and Leshem, 2011) and is in line with modern numerical cognition theories (Nieder and Miller, 2003). Cognitive models assuming logarithmic and other non-linear forms of time perception have also appeared in the literature (Staddon and Higa, 1999; van Rijn et al., 2014).

Following the logarithmic representation, the temporal-distance  $TD$  between current time  $t$  and the time  $st_i$  that the  $i$ -th event started, is encoded as:

$$TD_i(t) = \begin{cases} 0, & t \leq st_i \\ \log(\frac{t}{st_i}), & st_i < t \end{cases} \quad (8)$$

We use  $TD_i(t)$  as the target of the  $i$ -th t-Distance module. Therefore, to evaluate the performance of the module encoding temporal distance of the  $i$ -th event we use an error-based measure that is:

$$EDist_i = \sum_t (out_i(t) - TD_i(t))^2 \quad (9)$$

<sup>1</sup>We have followed both modeling assumptions in our work and we have successfully implemented distant time models assuming either a linear or a logarithmic representation of time. The current paper demonstrates only the logarithmic approach but it is straightforward to adapt the evolutionary procedure with the assumption of a linear time.

This is used to define the fitness function that drives the evolution of the corresponding  $i$ -th t-Distance module. In particular, the modules t-Distance1, t-Distance2, ... t-Distance6 and all relevant incoming links are evolved according to the fitness function:

$$ff_{dist,i} = (1000/EDist_i) \quad (10)$$

Similar to the early setup of the coevolutionary procedure the modules t-Duration1, t-Duration2 ... t-Duration6, and all incoming links are evolved according to the fitness function:

$$ff_{dur,i} = (1000/EDur_i) \quad (11)$$

The module specific fitness functions are properly mixed to develop composite fitness functions that drive the evolution of the supportive modules. More specifically, the fitness function of the module t-Distance considers the performance of all six t-Distance $i$  modules:

$$ff_{dist} = \prod_i ff_{dist,i} \quad (12)$$

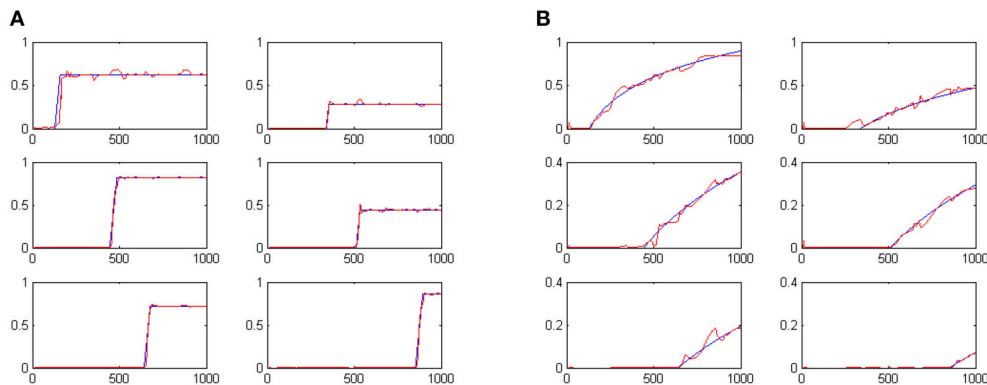
Similarly, the fitness function of the t-Duration module considers the performance of all six t-Duration $i$  modules:

$$ff_{dur} = \prod_i ff_{dur,i} \quad (13)$$

Finally the root components of the system t-Sen1 and t-Sen2 that implement time sense are evolved according to both the temporal distance and the temporal duration criteria, resulting into the fitness function:

$$ff_{global} = ff_{dur} * ff_{dist} \quad (14)$$

The hierarchical coevolutionary procedure accomplishes parametrical tuning of all system components taking into



**FIGURE 5 | System outputs after the perception of six random tone events.** Panel (A) shows the estimated duration of events. Panel (B) shows perceived temporal distance to the present as a means of representing when events have occurred.

account their special features as well as the successful functionality of the composite time processing system. The hierarchical and synthetic structure of the fitness functions enforces the coevolutionary scheme to improve collaboration between the component neural networks. As a result, the coevolutionary procedure can successfully converge to partial solutions that synthesize a composite system capable of memorizing the duration and the time of occurrence of events.

## Results

Following the coevolutionary procedure described above, the cognitive system described in Section Artificial Evolution of Interval Timing Model is advanced to address both the *when* and the *how-long* aspects of events. The configurations of previously existing CTRNN modules have been reloaded and evolved further, together with the configurations of the newly introduced components. The extended cognitive system has been evolved for 300 epochs producing a composite cognitive system that can successfully process temporal information. Sample results of the system outputs when memorizing 6 randomly initiated tone events are shown in **Figures 5A,B**. The plots show in blue the desired output and in red the actual output of the system. For example the two plots shown in the first column, second line of **Figures 5A,B** encode the fact that a tone event of duration 42 (note:  $42/50 = 0.84$ ) has occurred at a past time that is 557 moments back from the present (note:  $\log(1000/443)=0.353$ ).

The development of temporal processing internally in the model is shown in **Figures 6A–D**. The four plots show neural activity in the t-Sen1, t-Sen2, t-Duration, and T-Distance modules for the whole period of perceiving the 6 events. In the first stage of processing (**Figure 6A**), neural activity is mainly directed by the input oscillatory signals. Subsequently (**Figure 6B**) oscillations are mixed to produce a complex temporally structured neural activity. The first event occurs approximately at the moment 150. It seems that this event triggers a more structured oscillation fusion in t-Sen2 resulting in neural activity that looks like oscillation multiplexing. While the present model was not implemented on the basis of integrating oscillations that correspond to the known brain rhythms (delta

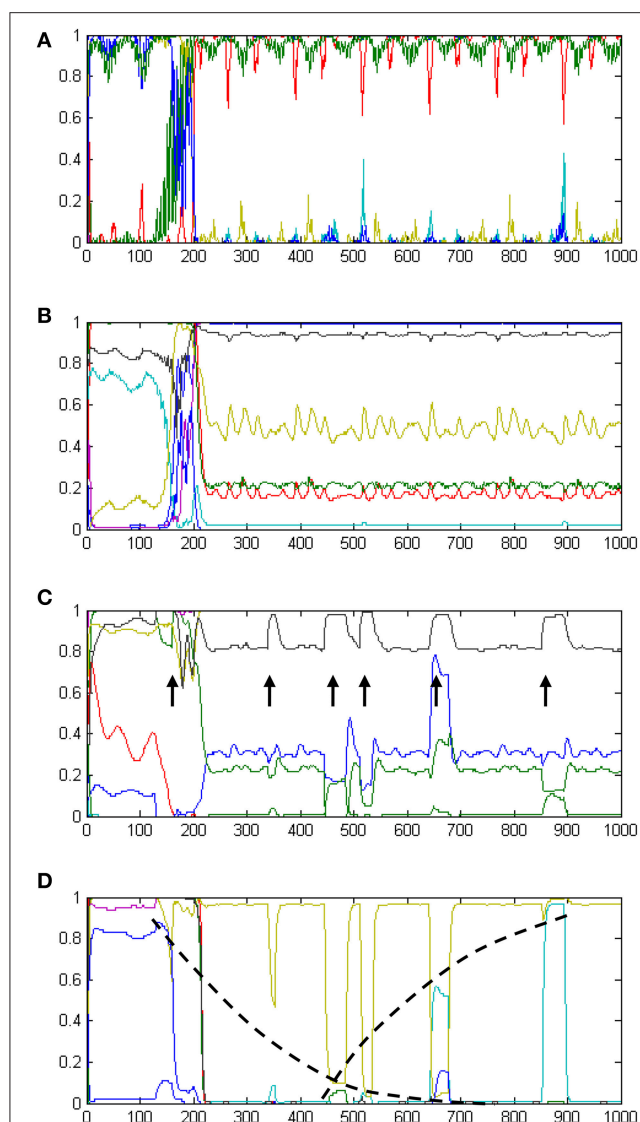
band to gamma band), our results show that the combination of input signals at different frequencies may significantly contribute in the sense of time as suggested also in (Kononowicz and van Wassenhove, 2016).

At the third stage, processing separates to interval timing and temporal distance to the past. Neural activation in the t-Duration module is presented in **Figure 6C**. As it is shown in the plot, the length of the events appearing (approximately) at times 340, 440, 510, 650, and 860 is correlated to the width of the peak disturbances (marked with arrows), as shown in the respective plot. The final stage of processing is the one shown in **Figure 5A**, demonstrating the correct estimate of interval timing. Interestingly, longer durations correspond to flat peaks that take longer to smooth out, while shorter durations have no time to develop flat activities

Neural activation in the t-Distance module is shown in **Figure 6D**. The plot shows a gradual increase in the amplitude of activation disturbances as more and more events gradually occur. The dotted lines drawn on top of the neural activities shown in yellow and cyan reveal two non-linear measures to be kept internally in the model. The mixture of these two self-organized measures is adequate for measuring temporal distances to the present as it appears by the relevant outputs of the model in **Figure 5B**.

## DISCUSSION

We have presented a neural network model that is capable of measuring short time intervals assuming linear ramp activity and keep track of past times based on the logarithmic representation of temporal distances. The model is implemented following a semi-automated procedure that assumes parameterized CTRNN modules attuned with the help of coevolutionary optimization. The tuning of model parameters is accomplished in an offline mode, similar to the supervised learning approach followed in other timing neural network models (Laje and Buonomano, 2013). Interestingly, evolutionary methods can be nicely combined with on-line adaptation procedures to facilitate



**FIGURE 6 | A summary of the internal dynamics (neural activations) developed in the model over time.** Panel (A) shows neural activation in tSen1, the first receiving component of the recurrent TimeSense module. Panel (B) shows neural activation in tSen2, the second output component of the TimeSense module. Panel (C) shows neural activation in the t-Duration module which supports interval timing. Arrows indicate the times of event experiencing. The width of peaks is analogous to the duration of events, therefore enabling accurate duration estimation as shown in Figure 5A. Panel (D) shows neural activation in the t-Distance module which supports past perception. Neural activities shown in yellow and cyan implement internal time keeping of the elapsed time as illustrated by the log-shaped dotted black lines following their peaks.

life-long learning (Maniadakis and Trahanias, 2008) and thus enable modifying the range of processed durations.

The neuro-evolutionary framework considered in the present study provides increased flexibility in designing the internal mechanisms of the model, accomplishing to easily bridge oscillatory input and ramping activity in a single model. While the two mechanisms have been frequently

considered contradictory in the literature, the use of oscillations with gradually adapted characteristics provides the basis for implementing effective interval timing mechanisms (Simen et al., 2013) and has been used for accomplishing multiple interval timing tasks (Maniadakis and Trahanias, 2015).

In contrast to previous works proposing timing models that have been rather minimally integrated with other cognitive functions (Gibbon et al., 1984; Staddon and Higa, 1999; Dragoi et al., 2003; Droit-Volet et al., 2007), the incremental NN modeling approach greatly facilitates the implementation of complex time-aware cognitive systems that will enable robotic systems to further exploit temporal cognition. The present work considers the strong coupling of time perception and short-term memory as suggested in (Gu et al., 2015). Other relevant works have considered spatiotemporal patterns related to motor behaviors (Laje and Buonomano, 2013). The use of spiking recurrent neural networks for timing has been shown to be particularly sensitive to noise (Banerjee et al., 2008). Relevant computational models shown that, especially for SBF, different types of noise may differentially affect the encoding and recall of timing intervals (Oprisan and Buhusi, 2014). Despite enforcing noise tolerance through learning (Laje and Buonomano, 2013), our study shows that the use of rate coding neurons may significantly facilitate model robustness.

The main contribution of the present study in comparison to the state of the art regards the use of past distance measures as a means of encoding the time of occurrence of experienced events. Our results show that a single timing source can be used as a basis for implementing cognitive systems capable of encoding *when* events occurred and *how-long* they have lasted. The proposed model suggests it is possible to bridge both short- and long-time keeping mechanisms that in the literature have been so far considered largely independent (Aschoff, 1985; Rammsayer, 1999; Lewis and Miall, 2003).

We note that the SBF-like characteristics assumed in the current implementation are not restrictive for bridging when and how long. Apart from the specific timing mechanism assumed by SBF, the proposed modeling approach could be nicely combined with other representations of time, such as (Miall, 1989; Staddon and Higa, 1999; Karmarkar and Buonomano, 2007; Large, 2008; Simen et al., 2011). However, even if nearly all timing models could equally support interval timing and past-distance measuring, using a single timing mechanism for both *when* and *how-long* can hardly comply with the brain studies explicitly distinguishing the two systems. Along this line, the current model assumes separate subsystems dedicated to the estimation of short-term durations and long-term temporal distances, providing the means to sufficiently address qualitative differences between them. This is accomplished by assuming different forms of temporal information to be readout by TimeSense neurons, which are subsequently processed assuming different mechanisms and processes.

The encoding of estimated times in memory highlights two very interesting problems that a time-aware cognitive system must concern in order to be functional in naturalistic conditions. The first problem regards how *how-long* and *when* should be represented in memory. In the former case the duration is



necessary to gradually increase as long as the event is experienced by the agent and stop at a specific value that will be encoded in memory, representing the static (never changing again) duration of the event. The latter case assumes a counting mechanism that increases together with the evolution of the event but continues increasing after the end of the event, resulting into a dynamic (non-static) representation of past times relative to the present. The distinction between static and dynamic time representations gets even more complicated by considering the second problem, which regards how a cognitive system links specific events with specific temporal characteristics successfully keeping track of their values while other events may also occur. In our implementation, the use of a dedicated Event-id module (see **Figure 4**) enables the correct association between events and times, filtering out irrelevant external stimuli.

Overall, the following points summarize the differences between the *how-long* and *when* modes of operation in the model:

- *How-long* and *when* assume respectively a linear and a logarithmic representation of time, therefore accomplishing to measure durations of different scales (up to 50 simulation steps for *how-long* vs. up to 1000 simulation steps for *when*).
- The *how-long* mechanism aims at counting the time filled with the occurrence of an event, while the *when* mechanism counts time that is not any more in direct link to the given event.
- *How-long* results into the final encoding of a static duration value in memory, while *when* assumes an ever-changing, dynamic representation of past times relative to the moving present.

Currently, the model exhibits two limitations which, at the same time, offer two important strands for future work. The first regards the representation of far distant times which ordinary models address by assuming processes that can increase without limit (Miall, 1989; Matell and Meck, 2004; Large, 2008; Simen et al., 2011). Despite the fact that such unbounded processes can hardly provide a realistic explanation of time perception (Staddon and Higa, 1999), they do not address multiscale time perception that is innate for humans. Interestingly, the newly introduced DDM (Simen et al., 2013) model which uses adapting pulse rates to measure time intervals could provide a means for implementing multi-scale time perception, assuming the future implementation of a time abstraction mechanism (i.e., I am only aware that I moved to a new city 6 months ago, but I do not know how many seconds or minutes have passed since then). In the present work, the use of sigmoid activation functions in the output neurons of the model does not fully comply with the representation of far distant times. Sigmoid functions produce outputs in the range  $[0, 1]$ , therefore they are not appropriate for approximating logarithmic times greater than one. To compensate this limitation we plan to implement multi-scale time perception in the cognitive system, similar to (Staddon and Higa, 1999). Each time scale will be implemented as a logarithmic function with a basis of a second, a minute, an hour and so on (i.e.,  $\log_{\text{sec}}$ ,  $\log_{\text{min}}$ ,  $\log_{\text{hour}}$ , etc.). An event that approximates the maximum sigmoid value of one in a given scale will “jump” to the next scale, starting from a relatively low value

which will gradually increase to one being ready for a new “jump” and so on.

The second direction for advancing the model regards the perception of not only past, but also future times. This important addition will pave the way for investigating long-term planning, self-projection to the future, imagination and other high level cognitive skills which are currently unattainable in artificial systems. Similar to past perception, we plan to implement future time perception following the assumption of logarithmic multi-scale times. Future perception will look like past perception, horizontally flipped with respect to zero-time that represents “now”. The composite model will be able to perceive future (expected) events approaching the present, be part of reality (occur) and then moved to the past (memorized).

The embodiment of the model into a robotic system and its practical application in real life has revealed some particularly challenging issues for artificial temporal cognition. So far we assume that experienced events are assigned ids in a periodic manner, i.e., in the form 1,2,3,4,5,6,1,2,3,4... and so on, and thus their temporal characteristics are circularly encoded in the relevant output modules in short term memory. As new events are experienced by the agent, previous events should be either deleted or transferred to long-term memory. The details of this mechanism remains an open research issue, however by mixing elapsed time and the attention devoted to the event we have been able to implement rough criteria that facilitate decision making with respect to the handling of past events. Currently we use a simple Data Base system to encode past events in LTM, but we are also investigating neural representations that will enable abstracting and encoding events in the form of episodes.

## CONCLUSIONS

Our perception and consideration of time, is key in determining how we behave and in the decisions we make. Besides the increasing research interest that is recently devoted on temporal cognition there not much studies linking the *how-long* and *when* aspects of perceived events. Both of these aspects are fundamental for the rich and meaningful perception of the environment. The present work considers a memory representation perspective to link short- and long-term time perception, accomplished by using a single timing source to perceive both event-specific and event-irrelevant times.

The broader vision of our research aims at time-aware artificial autonomous systems. The particularly promising results of the current work suggest that the proposed timing model can be the basis for implementing artificial systems that successfully interact with humans for the collaborative accomplishment of short- and mid-term goals.

## ACKNOWLEDGMENTS

This work has been partially supported by the EU FET grant (GA: 641100) TIMESTORM - Mind and Time: Investigation of the Temporal Traits of Human-Machine Convergence.

## REFERENCES

- Arzy, S., Adi-Japha, E., and Blanke, O. (2009). The mental time line: an analogue of the mental number line in the mapping of life events. *Conscious. Cogn.* 18, 781–785. doi: 10.1016/j.concog.2009.05.007
- Aschoff, J. (1985). On the perception of time during prolonged temporal isolation. *Hum. Neurobiol.* 4, 41–52.
- Banerjee, A., Seriès, P., and Pouget, A. (2008). Dynamical constraints on using precise spike timing to compute in recurrent cortical networks. *Neural Comput.* 20, 974–993. doi: 10.1162/neco.2008.05-06-206
- Bruss, F. T., and Ruschendorf, L. (2010). On the perception of time. *Gerontology* 56, 361–370. doi: 10.1159/000272315
- Bueti, D. (2011). The sensory representation of time. *Front. Integr. Neurosci.* 5:34. doi: 10.3389/fnint.2011.00034
- Campbell, J. (1997). The structure of time in autobiographical memory. *Eur. J. Philos.* 5, 105–118. doi: 10.1111/1468-0378.00031
- Dragoi, V., Staddon, J. E., Palmer, R. G., and Buhusi, C. V. (2003). Interval timing as an emergent learning property. *Psychol. Rev.* 110, 126–144. doi: 10.1037/0033-295X.110.1.126
- Droit-Volet, S., Meck, W. H., and Penney, T. (2007). Sensory modality and time perception in children and adults. *Behav. Proc.* 74, 244–250. doi: 10.1016/j.beproc.2006.09.012
- Droit-Volet, S., Mermillod, M., Cocenas-Silva, R., and Gil, S. (2010). The effect of expectancy of a threatening event on time perception in human adults. *Emotion* 10, 908–914. doi: 10.1037/a0020258
- Droit-Volet, S. (2011). *Child and Time. Multidisciplinary Aspects of Time and Time Perception*, Vol. 6789. Athens: Lecture Notes in Computer Science.
- Friedman, W. (1998). The effects of elapsed time and retrieval on young children's judgments of the temporal distances of past events. *Cogn. Dev.* 13, 335–367. doi: 10.1016/S0885-2014(98)90015-6
- Friedman, W. (2005). Developmental and cognitive perspectives on humans' sense of the times of past and future events. *Learn. Motiv.* 36, 145–158. doi: 10.1016/j.lmot.2005.02.005
- Gallistel, C. R., and Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition* 44, 43–74. doi: 10.1016/0010-0277(92)90050-R
- Gibbon, J., Church, R., and Meck, W. (1984). "Scalar timing in memory," in *Timing and Time Perception*, eds J. Gibbon and L. G. Allan (New York, NY: Academy of Sciences), 52–77.
- Glicksohn, J., and Leshem, R. (2011). *Reproduction of Duration: How Should I Count the Ways? Multidisciplinary Aspects of Time and Time Perception*, Vol. 6789. Athens: Lecture Notes in Computer Science.
- Gu, B.-M., van Rijn, H., and Meck, W. (2015). Oscillatory multiplexing of neural population codes for interval timing and working memory. *Neurosci. Biobehav. Rev.* 48, 160–161. doi: 10.1016/j.neubiorev.2014.10.008
- Ivry, R. B., and Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends Cogn. Sci.* 12, 273–280. doi: 10.1016/j.tics.2008.04.002
- Karmarkar, U. R., and Buonomano, D. V. (2007). Timing in the absence of clocks: encoding time in neural network states. *Neuron* 53, 427–438. doi: 10.1016/j.neuron.2007.01.006
- Kononowicz, T., and van Wassenhove, V. (2016). In search for oscillatory traces of the internal clock. *Front. Psychol.* 7:224. doi: 10.3389/fpsyg.2016.00224
- Laje, R., and Buonomano, D. (2013). Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nat. Neurosci.* 16, 925–933. doi: 10.1038/nn.3405
- Large, E. (2008). "Resonating to musical rhythm: Theory and experiment," in *Psychology of Time*, ed S. Grondin (Bingley: Emerald Publishing Group, Ltd.), 189–232.
- Lejeune, H., and Wearden, J. (2006). Scalar properties in animal timing: conformity and violations. *Q. J. Exp. Psychol.* 59, 1875–1908. doi: 10.1080/17470210600784649
- Leon, M. I., and Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron* 3, 317–327. doi: 10.1016/S0896-6273(03)00185-5
- Lewis, P. A., and Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Curr. Opin. Neurobiol.* 1, 250–255. doi: 10.1016/S0959-4388(03)00036-9
- Maimon, G., and Assad, J. A. (2006). A cognitive signal for the proactive timing of action in macaque LIP. *Nat. Neurosci.* 9, 948–955. doi: 10.1038/nn1716
- Maniadakis, M., and Trahanias, P. (2008). Hierarchical coevolution of cooperating agents acting in the brain-arena. *Adap. Behav.* 16, 221–245. doi: 10.1177/1059712308087285
- Maniadakis, M., and Trahanias, P. (2009). Agent-based brain modelling for artificial organisms by means of hierarchical cooperative coevolution. *Artif. Life* 15, 293–336. doi: 10.1162/artl.2009.Trahanias.007
- Maniadakis, M., and Trahanias, P. (2015). Integrated intrinsic and dedicated representations of time: a computational study involving robotic agents. *Timing Time Percept.* 3, 1–23. doi: 10.1163/22134468-03002052
- Maniadakis, M., Trahanias, P., and Tani, J. (2009). Explorations on artificial time perception. *Neural Netw.* 22, 509–517. doi: 10.1016/j.neunet.2009.06.045
- Matell, M. S., and Meck, W. H. (2004). Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cogn. Brain Res.* 21, 139–170. doi: 10.1016/j.cogbrainres.2004.06.012
- Meck, W. H., Penney, T. B., and Pouthas, V. (2008). Cortico-striatal representation of time in animals and humans. *Curr. Opin. Neurobiol.* 18, 145–152. doi: 10.1016/j.conb.2008.08.002
- Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Comput.* 1, 359–371. doi: 10.1162/neco.1989.1.3.359
- Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., and Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nat. Neurosci.* 12, 502–507. doi: 10.1038/nn.2272
- Nieder, A., and Miller, E. (2003). Coding of Cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* 37, 149–157. doi: 10.1016/S0896-6273(02)01144-3
- Oprisan, A. A., and Buhusi, C. V. (2014). What is all the noise about in interval timing? *Philos. Trans. R. Soc. B Biol. Sci.* 369, 1–12. doi: 10.1098/rstb.2012.0459
- Rammesayer, T. H. (1999). Neuropharmacological evidence for different timing mechanisms in humans. *Q. J. Exp. Psychol. B* 52, 273–286.
- Rouder, J. N., and Geary, D. C. (2014). Children's cognitive representation of the mathematical number line. *Dev. Sci.* 17, 525–536. doi: 10.1111/desc.12166
- 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. doi: 10.1523/JNEUROSCI.3121-10.2011
- 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. doi: 10.1163/22134468-00002018
- Staddon, J. E., and Higa, J. J. (1999). Time and memory: towards a pacemaker-free theory of interval timing. *J. Exp. Anal. Behav.* 71, 215–251. doi: 10.1901/jeab.1999.71-215
- Treisman, M. (1984). Temporal rhythms and cerebral rhythms. *Ann. N.Y. Acad. Sci.* 423, 542–565. doi: 10.1111/j.1749-6632.1984.tb23458.x
- van Rijn, H., Gu, B.-M., and Meck, W. H. (2014). Dedicated clock/timing-circuit theories of interval timing and timed behavior. *Adv. Exp. Med. Biol.* 829, 75–99. doi: 10.1007/978-1-4939-1782-2\_5
- Wiener, M., and Kanai, R. (2016). Frequency tuning for temporal perception and prediction. *Curr. Opin. Behav. Sci.* 8, 1–6. doi: 10.1016/j.cobeha.2016.01.001
- Wyer, N. A., Perfect, T. J., and Pahl, S. (2010). Temporal distance and person memory: thinking about the future changes memory for the past. *Pers. Soc. Psychol. Bull.* 36, 805–816. doi: 10.1177/0146167210370965

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Maniadakis and Trahanias. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

