



# **TEMPORAL STRUCTURE OF NEURAL PROCESSES COUPLING SENSORY, MOTOR AND COGNITIVE FUNCTIONS OF THE BRAIN**

EDITED BY: Daya Shankar Gupta, Arpan Banerjee, Dipanjan Roy and  
Federica Piras

PUBLISHED IN: Frontiers in Computational Neuroscience and  
Frontiers in Neuroscience





# frontiers

## Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88966-150-3

DOI 10.3389/978-2-88966-150-3

## 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)

# TEMPORAL STRUCTURE OF NEURAL PROCESSES COUPLING SENSORY, MOTOR AND COGNITIVE FUNCTIONS OF THE BRAIN

Topic Editors:

**Daya Shankar Gupta**, Camden County College, United States

**Arpan Banerjee**, National Brain Research Centre (NBRC), India

**Dipanjan Roy**, National Brain Research Centre (NBRC), India

**Federica Piras**, Santa Lucia Foundation (IRCCS), Italy

**Citation:** Gupta, D. S., Banerjee, A., Roy, D., Piras, F., eds. (2020). Temporal Structure of Neural Processes Coupling Sensory, Motor and Cognitive Functions of the Brain. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-150-3

# Table of Contents

- 05 Editorial: Temporal Structure of Neural Processes Coupling Sensory, Motor and Cognitive Functions of the Brain**  
Daya Shankar Gupta, Arpan Banerjee, Dipanjan Roy and Federica Piras
- 09 Dynamic Causal Modeling of the Cortical Responses to Wrist Perturbations**  
Yuan Yang, Bekir Guliyev and Alfred C. Schouten
- 17 Decoding English Alphabet Letters Using EEG Phase Information**  
YiYan Wang, Pingxiao Wang and Yuguo Yu
- 27 Temporal Dissociation of Neocortical and Hippocampal Contributions to Mental Time Travel Using Intracranial Recordings in Humans**  
Roey Schurr, Mor Nitzan, Ruth Eliahou, Laurent Spinelli, Margitta Seeck, Olaf Blanke and Shahar Arzy
- 39 The Amount of Time Dilation for Visual Flickers Corresponds to the Amount of Neural Entrainments Measured by EEG**  
Yuki Hashimoto and Yuko Yotsumoto
- 48 Temporal Reference, Attentional Modulation, and Crossmodal Assimilation**  
Yingqi Wan and Lihan Chen
- 59 Timing Deficits in ADHD: Insights From the Neuroscience of Musical Rhythm**  
Jessica L. Slater and Matthew C. Tate
- 67 An Oscillatory Neural Autoencoder Based on Frequency Modulation and Multiplexing**  
Karthik Soman, Vignesh Muralidharan and V. Srinivasa Chakravarthy
- 82 Sensorimotor Synchronization With Auditory and Visual Modalities: Behavioral and Neural Differences**  
Daniel C. Comstock, Michael J. Hove and Ramesh Balasubramaniam
- 90 Effect of Stimulus Contrast and Visual Attention on Spike-Gamma Phase Relationship in Macaque Primary Visual Cortex**  
Aritra Das and Supratim Ray
- 102 Gender-Dependent Changes in Time Production Following Quadrato Motor Training in Dyslexic and Normal Readers**  
Tal Dotan Ben-Soussan and Joseph Glicksohn
- 112 Metastable States of Multiscale Brain Networks are Keys to Crack the Timing Problem**  
Tommaso Gili, Valentina Ciullo and Gianfranco Spalletta
- 120 Slower is Higher: Threshold Modulation of Cortical Activity in Voluntary Control of Breathing Initiation**  
Pierre Pouget, Etienne Allard, Tymothee Poitou, Mathieu Raux, Nicolas Wattiez and Thomas Similowski
- 130 Role of Oscillations in Auditory Temporal Processing: A General Model for Temporal Processing of Sensory Information in the Brain?**  
Andreas Bahmer and Daya Shankar Gupta



- 142** *The Temporal Dynamic Relationship Between Attention and Crowding: Electrophysiological Evidence From an Event-Related Potential Study*  
Chunhua Peng, Chunmei Hu and Youguo Chen
- 153** *Why Do Durations in Musical Rhythms Conform to Small Integer Ratios?*  
Andrea Ravignani, Bill Thompson, Massimo Lumaca and Manon Grube
- 162** *The Musical Structure of Time in the Brain: Repetition, Rhythm, and Harmony in fMRI During Rest and Passive Movie Viewing*  
Dan Lloyd



# Editorial: Temporal Structure of Neural Processes Coupling Sensory, Motor and Cognitive Functions of the Brain

Daya Shankar Gupta<sup>1\*</sup>, Arpan Banerjee<sup>2</sup>, Dipanjan Roy<sup>2</sup> and Federica Piras<sup>3</sup>

<sup>1</sup> Biology Department, Camden County College, Blackwood, NJ, United States, <sup>2</sup> Cognitive Brain Dynamics Lab, National Brain Research Center, Gurugram, India, <sup>3</sup> Laboratory of Neuropsychiatry, Department of Clinical and Behavioral Neurology, IRCCS Santa Lucia Foundation, Rome, Italy

**Keywords:** binding, perception, multi-scale timing, temporal coupling, mutual information, time-dimension in the brain

## Editorial on the Research Topic

## Temporal Structure of Neural Processes Coupling Sensory, Motor and Cognitive Functions of the Brain

## INTRODUCTION

Temporal structure of cognitive and sensory processing holds the key to understanding complex neural mechanisms involved in higher order brain functions like perception of time. A hypothesis of embodied cognition posits that cognitive processes are deeply rooted in the interactions with the external world (Wilson et al., 2002; Anderson et al., 2012). These interactions of the brain with the external world depend on the accurate representation of the time-dimension in neural circuits (Gupta, 2014). For example, one cannot catch a flying ball unless the timing of the movements matches the speed of the ball. Many real world situations depend on the mapping between the neural and physical representation of time, which is maintained at different hierarchical levels. Hierarchical processing, consistent with multiple time scales, is manifested during goal-driven tasks, such as interval timing, duration judgement, and movement coordination. Contributions to this Research Topic elucidate how key aspects of the time-dimension such as the temporal binding of neural events play important roles in various cognitive processes, which include perception, mental time travel, and speech production. Additionally, the multi-scale representation of such processes from the micro to meso scales—from single neurons to a population of neurons to field potentials and macroscopic scales of EEG - is, discussed.

## CONTRIBUTIONS

Hashimoto and Yotsumoto studied an oscillator-based model of time perception by recording EEG data during interval timing tasks. They observed that the duration reproduction of a visual stimulus, flickering at 10 Hz, was 1.22 times longer than constantly illuminated stimulus. The EEG data further revealed an event-related potential (ERP), phase-locked to the flicker, fluctuating at 10 Hz, suggesting an increase in the certainty about the physical time-dimension in the neural circuits of the brain, which can be analyzed as mutual information. Authors also found desynchronization of spontaneous neural oscillations during the flicker observation period,

## OPEN ACCESS

### Edited and reviewed by:

Si Wu,  
Peking University, China

### \*Correspondence:

Daya Shankar Gupta  
dayagup@gmail.com

**Received:** 05 July 2020

**Accepted:** 08 July 2020

**Published:** 15 September 2020

### Citation:

Gupta DS, Banerjee A, Roy D and Piras F (2020) Editorial: Temporal Structure of Neural Processes Coupling Sensory, Motor and Cognitive Functions of the Brain. *Front. Comput. Neurosci.* 14:73. doi: 10.3389/fncom.2020.00073

which is consistent with the presence of desynchronization during information processing (Kumar et al., 2020). Furthermore, during time reproduction, there was an increase of the spontaneous alpha oscillation amplitude. This is consistent with the synchronization of distributed oscillators by neural oscillations in the calibration of local circuits during time reproduction tasks (Gupta, 2014). Thus, the phase-locked 10 Hz ERP induced by the flickering stimulus is likely to improve the accuracy of neural timing mechanisms in the brain.

Arguably, an increase in the error in timing tasks will lead to an underestimation of time intervals. This is based on the likely effect of evolutionary pressures, whenever there is less certainty about the physical time-dimension in neural circuits we determine a decrease in interval timing accuracy. Indeed, the underestimation of time intervals, when accuracy of timing mechanisms decreases due to the allocation of neuronal resources/attention to other environmental aspects, will confer a survival advantage, for example, protecting oneself from an incoming projectile or catching fruit dropping at a certain speed from the tree. The same yet to be determined mechanism, which results in the underestimation of time due to a decrease in accuracy, could cause time-interval overestimation if there is an improvement in accuracy. Thus, the stimulus-induced 10 Hz neural entrainment, which increases the certainty of the time-dimension in neural circuits, would result in the overestimation of the time interval.

Ben-Soussan and Glicksohn studied the effect of 1 month of Quadrato Motor Training (QMT), a type of motor training, on time production in dyslexic patients. In contrast to the control subjects, the dyslexic subjects produced longer intervals. Authors argue that longer time interval production is due to increased attention. QMT activates many parts of the brain simultaneously, such as areas that are responsible for motor control and language-related functions. Thus, QMT training will result in an increased ability to simultaneously keep many networks active, which may increase neuronal resources for attention. Development of additional neuronal resources could also improve the calibration of the clock mechanisms (Gupta, 2014). This would increase accuracy of internal neural clocks, which as hypothesized, could be responsible for the over-production of time intervals.

The magnitude of time-intervals can also affect perception. This is suggested by the findings of Wan and Chen, who showed using the Ternus display in a forced choice task that prior exposure to longer mean (or last) auditory intervals elicited more reports of group motion, whereas the shorter mean (or last) auditory interval gave rise to more dominant perception of element motion. Although longer intervals, greater than 50 ms, promoted group motion in general, the longer auditory intervals, which are also a more efficient form of the time-dimension input (Comstock et al.), can produce their effect by increasing the certainty of longer time intervals in the neural circuits processing the Ternus display, determining a greater report of group motion. Likewise, the exposure to shorter auditory intervals prior to the forced choice task, which would increase the certainty of shorter time intervals, results in a greater report of element motion. The role of increased certainty is consistent with theory by Gupta and

Bahmer (2019), who proposed that perception is the outcome of an increase in mutual information, as well as surprisal.

Slater and Tate highlight the significant overlap between neural systems involved in processing rhythm and those implicated in Attention Deficit Hyperactivity Disorder (ADHD). Authors link the impaired attentional control seen in ADHD to their rhythm-related deficits. They assert that the same neural bases—from the brain circuitry to dopamine signaling—that support the processing of musical rhythm are implicated in ADHD. Additionally, they present computational models of rhythm perception, based on the entrainment of multiple neural oscillators (Large and Palmer, 2002; Slater and Tate). The multiple-oscillators model can also provide a basis to represent the time-dimension, as well as the transfer of timing information from one modality to another in the brain (Gupta, 2014), for example, from auditory to visual, as suggested by the cross-modal interaction reported by Wan and Chen.

Ravignani et al. presented a mathematical model to explore small integer-ratio bias in rhythm perception and production. This small integer bias is likely to be also reflected in the representation of the time-dimension in neural oscillators, which would be responsible for the input of time-intervals that are small integer ratios. This could explain better perception for integer-ratio stimuli over more complex metrical patterns (Large and Kolen, 1994), which is in line with the role of temporal duration in perceptual functions as mentioned above (Wan and Chen).

In subjects doing mental time travel tasks, Schurr et al. recorded via electrodes inserted in the hippocampus and the lateral temporal cortex (LTC). Recordings revealed early modulatory activity between ~100 and 300 ms in the left LTC, followed by later activity in the left hippocampus between ~400 and 600 ms, which were independent, as shown by electrode classification. The authors suggest that this represents a division of labor. Additionally, separation by about 100 ms between two modulations, that are not correlated, also suggests that the activities in the left LTC and the hippocampus are temporally coupled, which could serve as a basis for the overall experience of mental time travel. It should be noted that, in addition to a synchronous occurrence, two or more neuronal events may be temporally coupled by a non-zero duration separating them. Moreover, many temporally coupled neural events may not be correlated, unless they are also causally related, for example, by an external stimulus or brain oscillation (Gupta and Bahmer, 2019). As suggested by this study, temporal coupling of neural events, without correlations, could contribute to brain cognitive functions, including the perception of time durations of various scales.

In a review article, Bahmer and Gupta have argued that temporal coupling, especially coincidental activation of neural circuits, is responsible for perceptual functions of the brain. Simulations showed that few chopper neurons, employing coincidental activation, can generate different inter-stimulus intervals, which can contribute to pitch perception at the cortical level, especially in a difference of 0.2% in pitch by sensitive listeners.

Pouget et al. have shown using a task of voluntary breathing control, that intentional initiation of breathing occurred

when premotor EEG recorded potential reached a threshold. Furthermore, the reaction time in the voluntary initiation of breathing was correlated to the amplitude of the threshold, which varied stochastically. Hence, the initiation of breathing occurred according to a stochastically distributed reaction time. The stochastic distribution of breathing initiation times is consistent with the role of surprisal information in information processing during speech production (Gupta and Bahmer, 2019). Gupta and Bahmer (2019) have argued that surprisal information combined with the increase in mutual information plays an important role in the mental representation of perceptual objects, such as the specific contents of speech.

Wang et al. have shown that English alphabet letters (vowels and the letter t) can be decoded from phase and power of EEG oscillations over the occipital and temporal regions. However, in comparison to the phase, the power of oscillation was less effective in decoding. It should be noted that an increase in oscillation power indicates that more oscillations are in the same phase, which would reduce the effectiveness of multiple individual spikes in representing information. In this study, synchronization appears, at least to some extent, to represent information about the alphabets.

Soman et al. have proposed an autoencoder model based on an oscillatory neural network, which is tested using real life situations and synthetic EEG signals. Autoencoders encode signals, perform dimensionality reduction, and decode the signals. As the authors point out, in the human brain dimensionality reduction occurs when  $\sim 125$  million photoreceptors converge to  $\sim 1$  million neurons in the lateral geniculate nucleus and then the visual information spreads to the primary visual areas. The proposed autoencoder uses biologically plausible features of neuronal oscillatory systems, such as phase synchronization, frequency tuning, and convergence. For example, convergence of inputs in the encoder, combined with an inhibitory network, maximizes the variance, which extracts useful features of the input, while resulting in dimensionality reduction. The study of autoencoders as a computational model can shed light on the functions of various relay centers throughout the nervous system. Moreover, in the relay centers, autoencoder-like functionality, such as synchronization, can incorporate time-dimension into information processing.

Lloyd et al. adapted network analysis from graph theory to reveal structures in time (rather than in space) in fMRI image series in healthy subjects at rest, or passively viewing a movie, from the human connectome project. In their analysis, each whole brain image is a temporal node, i.e., a “moment.” Collections of correlated moments or nodes across a time-interval, comprising time points where patterns of global brain activity are similar—referred to as themes—were significantly detected. The authors also found rhythms and harmonies in the patterns of themes, which were broadly similar in two different experimental conditions. They hypothesized that the detected rhythms and their harmonic relationship suggests that harmonic signaling might be adaptive from a computational point of view. Further analysis of themes revealed that sequences of 6 or 7 s were most often rhythmic. Rhythmic sequences of 6 to

7 s length are unlikely to play a direct role in the information processing underlying direct perception and action. However, it would be interesting to see if these rhythmic sequences play a role in higher mental functions, such as planning and mental time travel.

Gili et al. argue that metastable dynamics underlie the interactions between parts of the brain necessary for its dynamic functioning associated to time perception. Metastability can be understood as an energy landscape for an ensemble of possible states, which define the phase space of the brain system. These possible states tend to achieve local and global minima. Unlike synchronization, which constrains synchronized parts of the brain, metastable states are characterized by a tendency for interacting parts of the brain to be independent. The authors argue that metastable states underlie time perception in multimodal sensory processing when different parts of brain may be independently processing different sensory modalities or serving motor functions.

Yang et al. studied how wrist stretch (perturbation) modulated the effective connectivity for the early and late periods between multiple brain areas, which included the primary somatosensory cortex, primary motor cortex, premotor cortex, supplementary motor area, and the posterior parietal cortex on both hemispheres. Dynamic causal modeling was applied to analyze the connectivity between different areas and its modulations when a constant torque was applied in the presence of the external perturbation. There were greater modulations in the late phase 100–350 ms post-perturbation compared to 20–100 ms post-perturbation. This work highlighted interactions between motor and sensory areas during movements, which would reflect the interaction of the brain with the four-dimensional physical world. An increase in connectivity, which is an estimate of mutual information, would play an important role in the temporal processing of information underlying the perceptual functions of the brain.

Das and Ray used spike-LFP coherence to test if the phase coding by gamma rhythm varies with stimuli contrast or attention in V1. Here the authors use phase coding (PC) to test a specific hypothesis: whether stimuli contrast or attentional load can change the position of the spike relative to the phase of gamma frequencies in LFP. To be phase coded, spikes resulting in stronger activation of pyramidal cells appear earlier in gamma cycles as they overcome the inhibition of pyramidal cells earlier. Interestingly, they report only a weak effect of attention on the spike-field phase relationship of PC in V1, which contrasts with the findings of Fries et al. (2007) and Fries (2015). On the macroscopic scale, however, Peng et al. argue that visual crowding, which is mainly attributed to processing in early visual areas, can be modulated by top-down attention.

## CONCLUSION

In our endeavor to understand time-dimension as a bridge to integrate multi-scale observations of behavior and brain information processing, the contributions in this Research Topic have revealed certain key aspects of time-dimension.

These include the possible role of temporal coupling by non-zero intervals and the effects of an increase in mutual information in neural circuits on perception and cognitive functions, given different aspects of physical time-dimension. Important future goals at this juncture should include (a) a study of temporal coupling between unrelated neural events and (b) an increase in mutual information in the brain, given

the temporal characteristics of external events, such as speed and rhythmicity.

## AUTHOR CONTRIBUTIONS

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

## REFERENCES

- Anderson, M. L., Richardson, M. J., and Chemero, A. (2012). Eroding the boundaries of cognition: implications of embodiment(1). *Top. Cogn. Sci.* 4, 717–730. doi: 10.1111/j.1756-8765.2012.01211.x
- Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034
- Fries, P., Nikolic, D., and Singer, W. (2007). The gamma cycle. *Trends Neurosci.* 30, 309–316. doi: 10.1016/j.tins.2007.05.005
- 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 Bahmer, A. (2019). Increase in mutual information during interaction with the environment contributes to perception. *Entropy* 21:365. doi: 10.3390/e21040365
- Kumar, V. G., Dutta, S., Talwar, S., Roy, D., and Banerjee, A. (2020). Biophysical mechanisms governing large-scale brain network dynamics underlying individual-specific variability of perception. *Eur. J. Neurosci.* doi: 10.1111/ejn.14747. [Epub ahead of print].
- Large, E. W., and Kolen, J. F. (1994). Resonance and the perception of musical meter. *Connect. Sci.* 6, 177–208. doi: 10.1080/09540099408915723
- Large, E. W., and Palmer, C. (2002). Perceiving temporal regularity in music. *Cogn. Sci.* 26, 1–37. doi: 10.1016/S0364-0213(01)00057-X
- Wilson, M. (2002). Six views of embodied cognition. *Psychon. Bull. Rev.* 9, 625–636. doi: 10.3758/bf03196322

**Conflict of Interest:** 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 © 2020 Gupta, Banerjee, Roy and Piras. 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) and the copyright owner(s) 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.



# Dynamic Causal Modeling of the Cortical Responses to Wrist Perturbations

Yuan Yang<sup>1,2\*</sup>, Bekir Guliyev<sup>1†</sup> and Alfred C. Schouten<sup>1,3</sup>

<sup>1</sup> Neuromuscular Control Laboratory, Department of Biomechanical Engineering, Delft University of Technology, Delft, Netherlands, <sup>2</sup> Department of Physical Therapy and Human Movement Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States, <sup>3</sup> Department of Biomechanical Engineering, MIRA Institute for Biomedical Technology and Technical Medicine, University of Twente, Enschede, Netherlands

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

Fabien Dal Maso,  
Université de Montréal, Canada  
Hristos Courellis,  
University of California, San Diego,  
United States

### \*Correspondence:

Yuan Yang  
y.yang-2@tudelft.nl;  
yuan.yang@northwestern.edu

<sup>†</sup>These authors have contributed  
equally to this work.

### Specialty section:

This article was submitted to  
Brain Imaging Methods,  
a section of the journal  
Frontiers in Neuroscience

**Received:** 07 June 2017

**Accepted:** 01 September 2017

**Published:** 13 September 2017

### Citation:

Yang Y, Guliyev B and Schouten AC  
(2017) Dynamic Causal Modeling of  
the Cortical Responses to Wrist  
Perturbations.  
*Front. Neurosci.* 11:518.  
doi: 10.3389/fnins.2017.00518

Mechanical perturbations applied to the wrist joint typically evoke a stereotypical sequence of cortical and muscle responses. The early cortical responses (<100 ms) are thought to be involved in the “rapid” transcortical reaction to the perturbation while the late cortical responses (>100 ms) are related to the “slow” transcortical reaction. Although previous studies indicated that both responses involve the primary motor cortex, it remains unclear if both responses are engaged by the same effective connectivity in the cortical network. To answer this question, we investigated the effective connectivity cortical network after a “ramp-and-hold” mechanical perturbation, in both the early (<100 ms) and late (>100 ms) periods, using dynamic causal modeling. Ramp-and-hold perturbations were applied to the wrist joint while the subject maintained an isometric wrist flexion. Cortical activity was recorded using a 128-channel electroencephalogram (EEG). We investigated how the perturbation modulated the effective connectivity for the early and late periods. Bayesian model comparisons suggested that different effective connectivity networks are engaged in these two periods. For the early period, we found that only a few cortico-cortical connections were modulated, while more complicated connectivity was identified in the cortical network during the late period with multiple modulated cortico-cortical connections. The limited early cortical network likely allows for a rapid muscle response without involving high-level cognitive processes, while the complexity of the late network may facilitate coordinated responses.

**Keywords:** sensory feedback, stretch response, dynamic causal modeling, sensorimotor network, EEG, effective connectivity

## INTRODUCTION

Bodily movement is one of the main ways how humans interact with the physical world (Schwartz, 2016). Movement can be generated by voluntary and reflex driven actions. Muscle stretch during active motor task (e.g., maintain an isotonic wrist flexion) results in a sequence of cortical and muscle responses, involving the central nervous system and the periphery.

In the periphery, the immediate muscle responses to stretch are known as stretch reflexes. Many studies investigated muscle responses to stretch using electromyography (EMG) after ramp-and-hold mechanical perturbations. For lower arm muscles, they typically reported a short-latency stretch response (20–50 ms post-perturbation onset) followed by a long-latency stretch response



(50–120 ms) and later voluntary reactions (>120 ms) (Scott, 2002; Pruszynski et al., 2011). The short-latency stretch response depends on the stretch velocity and involves a spinal network (Houk et al., 1981). The time delays in the afferent pathway from the periphery to the brain (20–30 ms) (MacKinnon et al., 2000) and the efferent pathway from the brain to the periphery (~20 ms) (Perenboom et al., 2015) would not allow for a transcortical pathway in the short-latency stretch response. Several experimental studies indicated cortical contributions to the long-latency stretch response. Recordings from corticomotoneuronal cells in Macaque monkeys showed a cortical effect on the long-latency stretch response (Cheney and Fetz, 1984). Subthreshold transcranial magnetic stimulation (TMS) over the contralateral motor cortex can modulate the long-latency stretch response but not the short-latency stretch response (Perenboom et al., 2015). Recent studies indicate the long-latency stretch response is not as simple as a “reflex” and at least could partly involve a voluntary feedback control component (Pruszynski et al., 2011; Pruszynski and Scott, 2012). Thus, we avoid the term “reflex” and “voluntary” in this paper and use “rapid” and “slow” transcortical muscle reactions to roughly distinguish the cortical-involved muscle reactions before 120 ms (i.e., long-latency stretch response) and after 120 ms (i.e., “standard” voluntary reaction) post-perturbation. Similar terminology has been previously used in a review from Pruszynski and Scott (2012).

In the central nervous system, cortical responses to muscle stretch have been investigated by previous studies using the event-related potential (ERP) (Abbruzzese et al., 1985; Campfens et al., 2015). The latencies and topographies of the stretch-evoked ERP reflect the time courses of cortical activity and most active areas in response to the muscle stretch. Both early (<100 ms post-perturbation onset) and late (>100 ms) ERP components were reported around the contralateral motor cortex (Campfens et al., 2015). Considering the efferent motor conduction delay (~20 ms), the early cortical response is thought to be related to the rapid transcortical muscle reaction to the perturbation (<~120 ms) while the late cortical response may be related to the slow transcortical muscle reaction (>~120 ms) (MacKinnon et al., 2000; Pruszynski and Scott, 2012). ERP results indicate that the primary motor cortex may contribute to both rapid and slow transcortical muscle reactions; however, exact cortical pathways are yet to investigate. The full cortical network for motor control is thought to involve multiple brain areas, including primary somatosensory cortex (S1), primary motor cortex (M1), premotor cortex (PM), supplementary motor area (SMA), and posterior parietal cortex (PPC) (Scott, 2004; Szameitat et al., 2012). These regions constitute the cortical sensorimotor network, which is a distributed and adaptable network that orchestrates the overall human motor behavior (Scott, 2004; Shibasaki, 2012).

In this study, we used dynamic causal modeling (DCM) to model the effective connectivity in the cortical network modulated by muscle stretch. Effective cortical connectivity refers to the strength of the causal influences between multiple cortical areas, which can be modulated by external perturbations (Friston, 2011). A few studies suggested that the rapid and slow transcortical muscle reactions are engaged by similar neural

circuitries in the brain (Pruszynski et al., 2011; Pruszynski and Scott, 2012). However, we hypothesize that the early response engages effective cortical connectivity in a less complex network to accelerate the muscle response with a shorter delay, i.e., rapid transcortical muscle reactions; while the slow transcortical muscle reaction is governed by a more complex cortical network in the late cortical response.

To valid our hypothesis, we estimated effective connectivity among the cortical areas involved in sensorimotor control of the wrist in response to a perturbation. Previous studies considered only M1, SMA, and PM as “key motor regions” for upper limb movement (Grefkes et al., 2008; Chen et al., 2010). In line with review papers on feedback based motor control (Scott, 2002, 2004), we added S1 and PPC to our possible functional cortical network models, since these two areas are closely related to feedback-based motor control. S1 is the brain area receiving the peripheral somatosensory input, while the PPC is known as a sensory association area which is essential to integrate different sensory inputs. We investigated effective cortical connectivity in the early period within 100 ms post-perturbation onset in comparison to the late period between 100 and 350 ms post-perturbation onset to check if the rapid and slow transcortical muscle reactions involve similar cortical areas and signal propagation pathways. Considering the afferent sensory transmission time delay (~20 ms) (Abbruzzese et al., 1985; Campfens et al., 2015), we used 20–100 ms as the period to investigate the early cortical network.

## MATERIALS AND METHODS

### Subjects and Ethical Statement

Seven healthy right-handed volunteers (one female) aged 23–28 years old participated in the experiment. This study was carried out in accordance with the recommendations of Human Subject Research guidelines, the Human Research Ethics Committee of the Delft University of Technology with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. All subjects signed informed consent before the experiment and received a small financial compensation for their participations. The protocol was approved by the Human Research Ethics Committee of the Delft University of Technology.

### Experimental Protocol

Subjects sat next to a wrist manipulator (Wristalyzer, Moog Inc., the Netherlands), which is an actuated rotating device with a single degree of freedom to exert flexion and extension perturbations to the wrist joint. The lower arm of the subject was strapped in the armrest, while the subject was closely touching the handle of the wrist manipulator (fixed with velcro). Subjects were instructed to relax their fingers and only use the wrist to do the task. The axis of wrist manipulator rotation was aligned with the axis of wrist rotation. Wrist torque was measured by a force transducer within the handle of the wrist manipulator.

The protocol contained 30 trials. Each trial started with auditory cue “beep” and a fixation in the center of the screen with a random period of 1.5–2 s. After this random period,

visual feedback was provided with an arrow in a circle. The angle of the arrow is proportional to the (low-pass filtered 1 Hz) torque applied by subjects. Subjects were instructed to push with a constant flexion torque (1.0 Nm) to the handle with their right wrist (keeping the arrow pointing upwards) using the visual feedback. Each trial contains 20 flexion and 20 extension ramp perturbations. Note that the visual feedback was low-pass filtered to avoid fast visual corrections to the perturbations. The wrist manipulator applied angular ramp perturbation to stretch the wrist muscles when the subject maintained the constant flexion torque (with std. <5%) for a random period of 1.5–2 s, and then stopped (and held) at the new position until the next perturbation. A ramp duration of 40 ms was used with the velocity 1.5 rad/s; giving a ramp amplitude of 0.06 rad. This duration is below the expected saturation level of long-latency EMG response and allows for both inhibition and facilitation of the long-latency stretch response (Lee and Tatton, 1982; Meskers et al., 2009; Perenboom et al., 2015). During the ramp, subjects were instructed to maintain the same level of force. Since the subjects were required to maintain a flexion torque, only the data from the extension ramp perturbations (stretching the wrist flexors) were included for analysis.

Electroencephalogram (EEG) was recorded using a 128-channel cap (5/10 systems, WaveGuard cap, ANT Neuro, The Netherlands) with Al/AgCl electrodes. EMG signals were measured from the flexor and extensor carpi radialis muscles of the right forearm using bipolar derivations 2 cm inter-electrode distance. EEG signals were recorded at by a bio-signal amplifier (Refa System, TMSi, The Netherlands), which acquired data at a sampling frequency of 2,048 Hz. The amplifier contains an antialiasing low-pass filter with the cut-off frequency of 552 Hz.

## Data Preprocessing

The continuous EEG signals were filtered by a 0.5–100 Hz zero-phase shift band-pass filter using EEGLAB (Delorme and Makeig, 2004) to remove possible high-frequency noise and slow trends in the data (e.g., blood pressure, heartbeat, breathing). A notch filter was used to reject the 50 Hz line power noise. Afterwards, EEG were segmented into 570 ms epochs with 220 ms pre-stimulus baseline plus 350 ms post-stimulus recording. The epochs contaminated by the artifacts (e.g., eye blinks/movements and EMG artifacts) were removed by visual inspection. In the data, we did not see visible artifacts due to the transient perturbations. On average 118 epochs were removed per participant, leaving  $472 \pm 53$  epochs per participants for analysis. Then the ERPs were derived by grand averaging the remaining epochs using the period of 220–20 ms before stimulus onset as the baseline. These extracted ERPs corresponding to the neural activity in the cortical regions of interest are used to quantify effective connectivity between those regions via DCM.

## Dynamic Causal Modeling

DCM was applied to analyse the effective cortical connectivity. Although various methods are available for analyzing effective cortical connectivity, most of them focus on the linear connectivity, such as partial directed coherence (Kaminski and Blinowska, 1991; Porcaro et al., 2013) and directed transfer

function (Babiloni et al., 2005). Previous studies have reported non-linear neuronal coupling in human stretch responses (Yang et al., 2016b) and voluntary motor control (Chen et al., 2010; Yang et al., 2016a) of lower arm muscles. Different from linear connectivity methods, DCM is a non-linear identification approach to reveal how external inputs cause changes in the coupling of neural populations in the effective connectivity network (Friston et al., 2003; Goulinden et al., 2014).

We used the standard DCM for ERP (David et al., 2006) as implemented in Statistical Parametric Mapping toolbox (SPM12, Wellcome Trust Centre for Neuroimaging, London, UK) to model effective connectivity among distributed cortical sources within the sensorimotor network. The analysis was performed for two different periods 20–100 and 100–350 ms after the perturbation onset.

DCM estimates effective connectivity in a network of reconstructed cortical sources. DCM is a neurobiologically constrained source reconstruction scheme including both spatial forward modeling and model inversion (David et al., 2006). For the spatial forward model, DCM uses similar leadfields as other source reconstruction methods (Kiebel et al., 2006). Beyond other source reconstruction methods, DCM combines the spatial forward model with a biologically informed temporal forward model to estimate the connectivity between sources (Friston et al., 2003; David et al., 2006).

In this paper, the leadfield of each source is modeled by a single equivalent current dipole (Kiebel et al., 2006). DCM analysis requires users to specify the prior locations (in mm in MNI coordinates) of each source in the cortical network for building the spatial forward model (David et al., 2006). Based on the review from Scott (2002) (Scott, 2002), we selected eight key regions in the cortical sensorimotor network: left and right primary somatosensory cortex (S1), left and right primary motor cortex (M1), left and right bilateral premotor cortex (PM), supplementary motor area (SMA), and posterior parietal cortex (PPC). The MNI coordinates of hand/wrist regions in these eight cortical areas were informed by previous fMRI studies (Szameitat et al., 2012; Vlaar et al., 2016) and provided in **Table 1**. Based on these eight cortical sources (see **Figure 1**), we specified six different connectivity models as shown in **Figure 2**. In all of the models, the S1 is the source receiving the external input. The model space was created using two model attributes: (1) whether the connectivity is partially or fully modulated by the stimulus, and (2) whether interhemispheric connectivity is left lateralized (since the perturbation is given to the right wrist) or symmetric.

The network model is inverted using a Bayesian approach described by Friston (2002), where a fix-form Laplace approximation is used to estimate probability distributions of parameters. This is under the Gaussian assumption, which enables computation of the likelihood from the prediction error. We then used Bayesian model comparison to identify the best model, based on approximation to the log-evidence obtained in the model inversion (Friston and Penny, 2011). In this study, we did not find an identical optimal model for all individuals. According to the practical recommendations provided by Stephan et al. (2010), the group-level analysis



was performed to find the best cross-subject model, where the pooled log-evidence for each model ( $m_i$ ) across subjects ( $y_1, \dots, y_7$ ) is defined as  $\ln p(y_1, \dots, y_7|m_i)$ . Assuming that the data for different subjects are independent, we then have  $\ln p(y_1, \dots, y_7|m_i) = \ln p(y_1|m_i) + \ln p(y_2|m_i) + \dots + \ln p(y_7|m_i)$  (Penny et al., 2004). This pooled log-evidence indicates how well a particular model explains multiple datasets. To compare model evidences on group level, we used random-effect group Bayesian model selection (BMS). Classical random-effect analysis detects whether model evidence is consistent across subjects. In contrast, the group-BMS approach identifies the proportion of subjects, which is best described in terms of the model evidence, i.e., the posterior probability that each model is more frequent than others (Rigoux et al., 2014). The log group Bayes factor ( $\ln BF_{i,j}$ ) between models is computed from pooled log-evidences, i.e.,  $\ln BF_{i,j} = \ln p(y_1, \dots, y_7|m_i) - \ln p(y_1, \dots, y_7|m_j)$ , to indicate that how much model  $i$  is superior to model  $j$  for the whole data set. The value of  $\ln BF_{i,j}$  between 20 and 150 indicates a strong evidence (according to 95% confidence level) in favor of model  $i$  than model  $j$ , while  $\ln BF_{i,j}$  larger than 150 indicates a very strong evidence (99% confidence level) (Penny et al., 2004).

**TABLE 1 |** MNI coordinates (mm) of eight sources in the cortical sensorimotor network: left (L) and right (R) primary somatosensory cortex (S1), left and right primary motor cortex (M1), left and right bilateral premotor cortex (PM), and supplementary motor area (SMA), posterior parietal cortex (PPC).

Sources	MNI coordinates (mm)
S1 L	−26 −40 68
S1 R	26 −40 68
M1 L	−33 −28 70
M1 R	33 −28 70
PM L	−54 −2 46
PM R	54 −2 46
SMA	−4 −10 64
PPC	−4 −46 68

After identifying the best cross-subject model (with the highest pooled log-evidence), we obtained the mean posterior estimates of all effective connectivity parameters for each subject and each period. These parameters represent the relative connectivity strengths between the two sources. The inferences on these parameters reflect the input (i.e., the muscle stretch) modulated changes in the effective connectivity. By investigating these inferences, we can identify the activities of which cortical areas are modulated by the muscle stretch and how they influence other cortical areas.

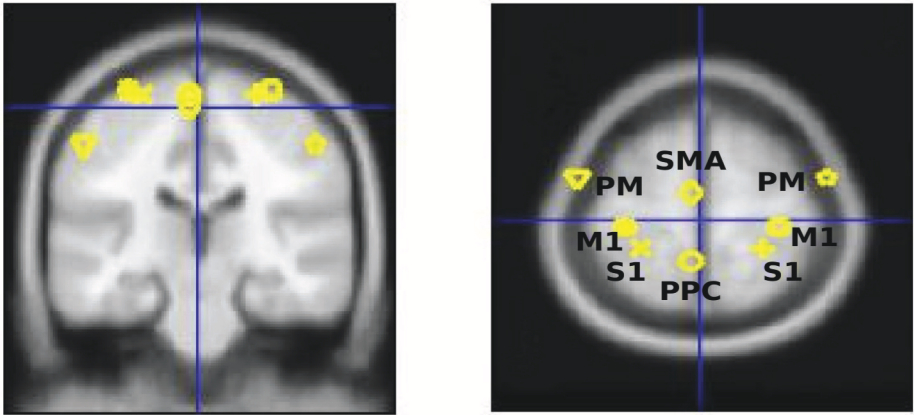
We averaged the connectivity strengths over subjects using Bayesian parameter averaging to get the mean estimate for each directional cortical interaction. We used one-sample  $t$ -test (two-tailed) to identify significant changes in the effective connectivity ( $p < 0.05$ , adjusted by false discovery rate estimation) in the best cross-subject model to get the perturbation-modulated effective connectivity for each period.

## RESULTS

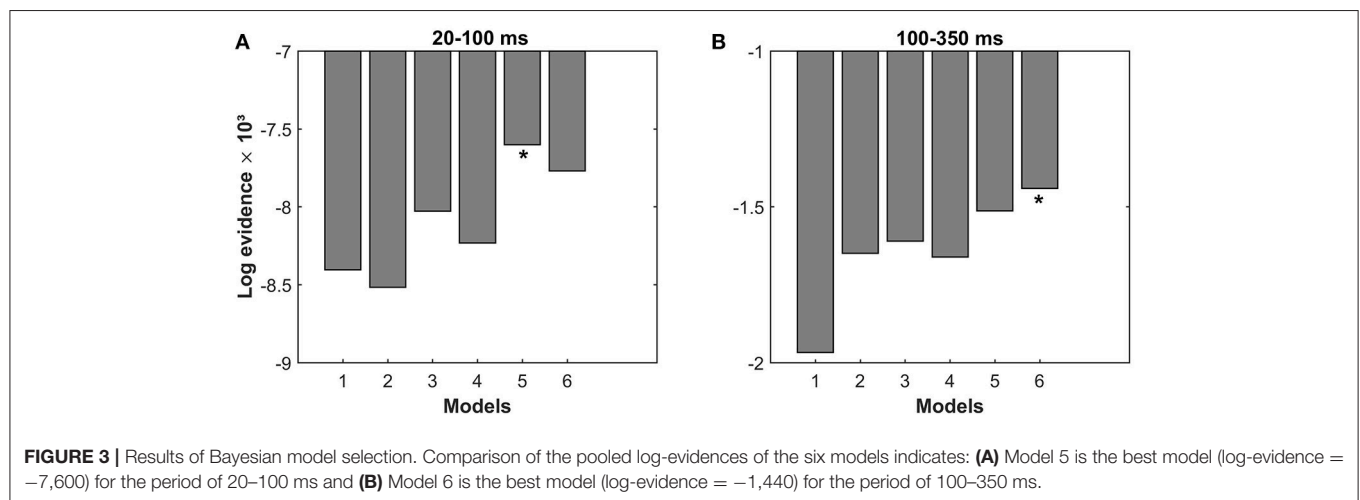
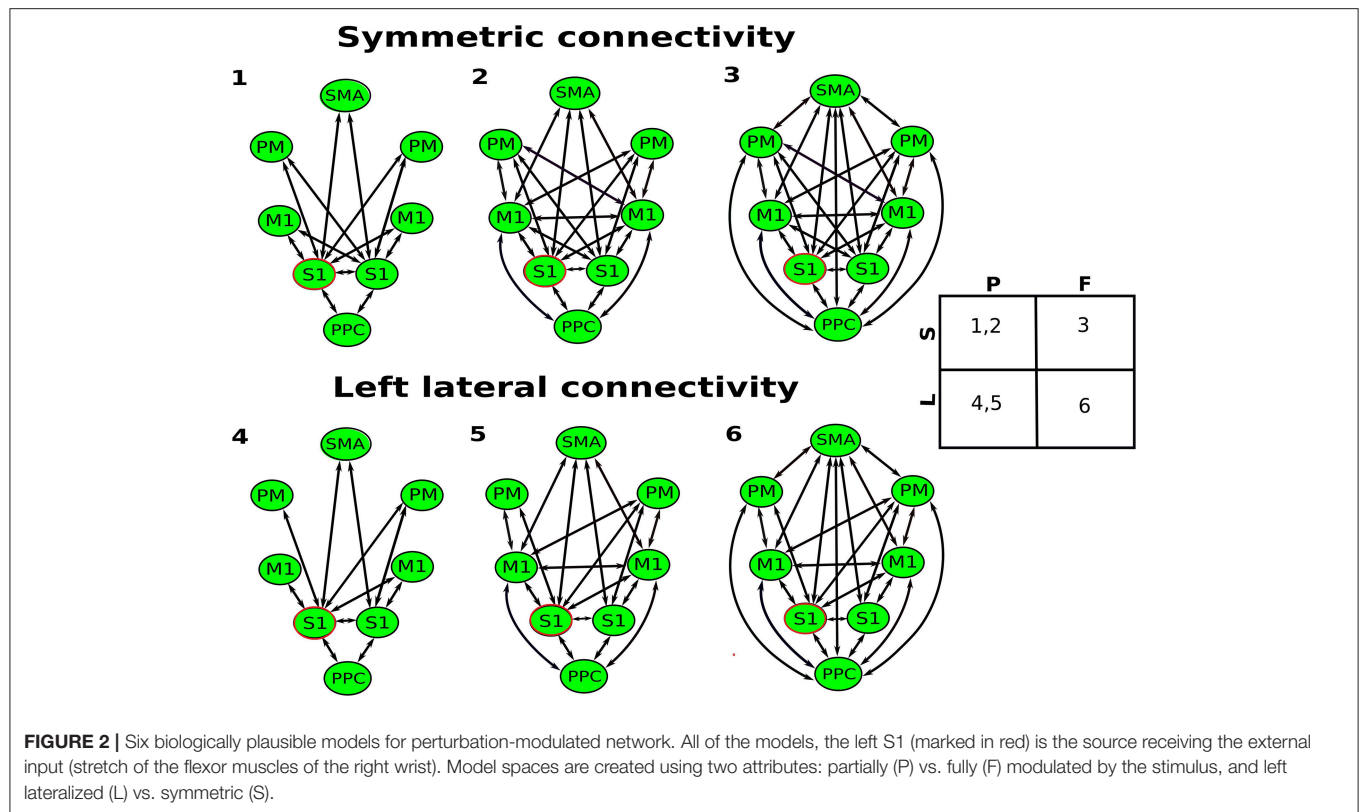
### Bayesian Model Selection

The effective connectivity in the cortical network after stretching the flexor muscles of the right wrist was modeled with DCM. We compared different Bayesian model families shown in Figure 2. Family-level Bayesian model comparison show that the left lateralized (L) models fit the data better than the symmetric (S) models for both periods ( $\ln BF_{L,S} = 1,350$  for 20–100 ms, and  $\ln BF_{L,S} = 611$  for 100–350 ms). The partial (P) modulated models fit the data better than the fully (F) modulated models for the period of 20–100 ms ( $\ln BF_{P,F} = 1,706$ ), while the fully modulated models provide substantially better fit for the period of 100–350 ms ( $\ln BF_{F,P} = 3,670$ ).

Figure 3 shows the pooled log-evidences for different models. For the period of 20–100 ms, the model comparison shows the strongest evidence for Model 5, which is a partially modulated left lateral model, with a Bayes Factor ( $\ln BF_{5,6}$ ) of 168 over the second-best model (Model 6). For the period of 100–350 ms,



**FIGURE 1 |** Eight selected cortical regions: left and right primary somatosensory cortex (S1), left and right primary motor cortex (M1), left and right bilateral premotor cortex (PM), and supplementary motor area (SMA), posterior parietal cortex (PPC).



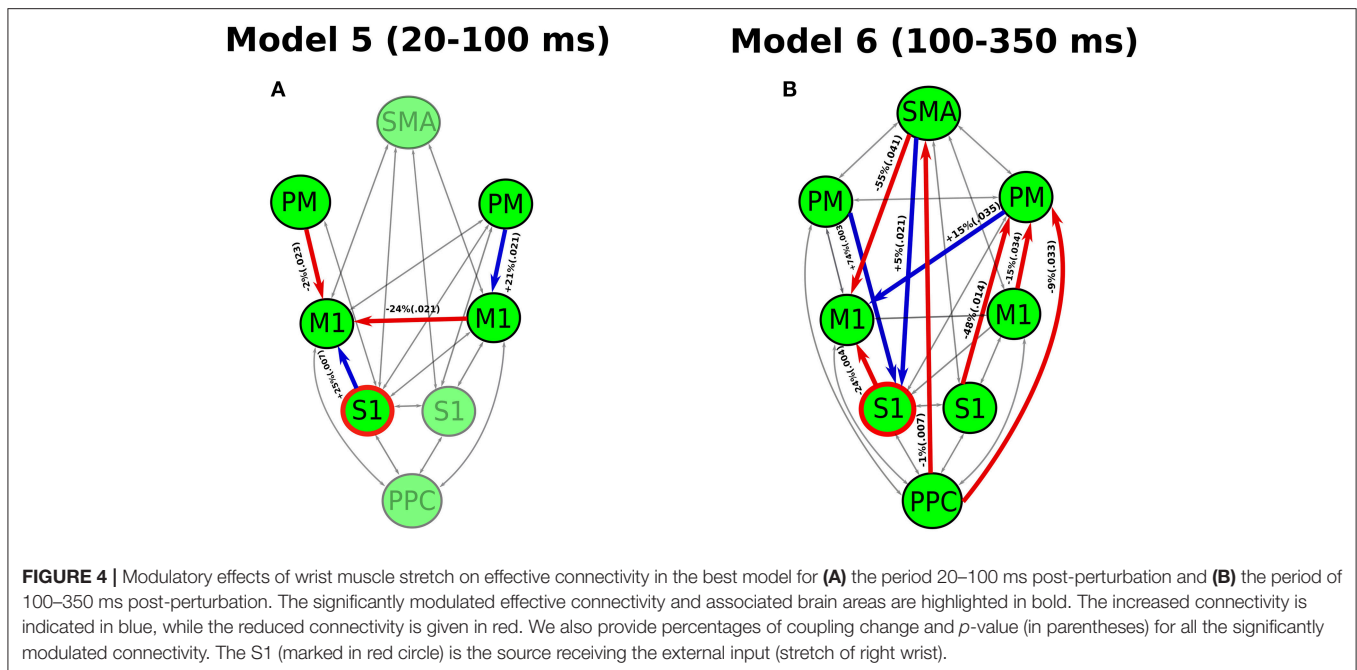
the strongest evidence is present for Model 6, which is a fully modulated left lateral model, with a with a Bayes Factor ( $\ln BF_{6,5}$ ) of 72 over the second-best model (Model 5).

### Inference on Coupling Parameters

The analysis of coupling parameters under the best cross-subject models reveals the significant modulations of effective connectivity by the perturbation for the period of 20–100 ms (Model 5) and 100–350 ms (Model 6), respectively (see **Figure 4**). During the period of 20–100 ms, the significant modulations only occur in the connectivity between M1 and a few cortical areas. In

the left hemisphere (contralateral side to the perturbation), we detected a decrease in the effective connectivity from PM to M1 while an increase from S1 to M1. In the right hemisphere, only an increase of effective connectivity is shown from PM to M1. The cross-hemisphere interaction shows a reduced connectivity from right M1 to left M1. The left M1, which comprises the upper motoneurons of the right wrist muscles, appears a “sink” for all modulated connectivity pathways in this period.

During the period of 100–350 ms, a larger number of connections among more cortical areas is modulated by the perturbation. Different from the period of 20–100 ms, the



connectivity with SMA, PPC and right S1 are also modulated. Specially, we found that a reduced connectivity pathway started from PPC through SMA to left M1. Additionally, there are three reduced connectivity pathways starting from PPC, left S1 and M1 all passing through right PM and arriving at left M1.

## DISCUSSION

In this study, we investigated the response of the effective connectivity in the cortical network to stretch of the flexor muscles of the right wrist. We built model spaces with left lateralized (i.e., contralateral to the perturbed wrist) and symmetric models for comparison. We did not include right lateralized models, since all subjects are right-handed and the task is performed with the right wrist. DCM suggested strong evidence that contralateral (left) lateralized models were superior to the symmetric models for both rapid (20–100 ms) and slow (100–350 ms) periods. These results are in line with previous studies reporting contralateral hemisphere dominance of the cortical response to wrist perturbations (Campfens et al., 2015) and during motor control (Chen et al., 2010; Yang et al., 2016b).

### DCM for the Early Cortical Response to Muscle Stretch

During the early period of 20–100 ms, the partially modulated models (of the effective connectivity in the cortical network) fit the data better than the fully modulated models, showing a relatively simpler network compared to the period of 100–350 ms. In the best model (Model 5), only a few connections among several key cortical areas are significantly modulated during the early period (see **Figure 4A**). This likely facilitates

a rapid motor reaction to the perturbation without involving high level cognitive processes. Previous studies have found direct monosynaptic connections between the S1 and M1, which allows fast signal propagations between S1 and M1 (Rocco-Donovan et al., 2011). Here, we detected an increased connectivity from S1 to M1 in the contralateral hemisphere. This enhanced S1-M1 connectivity may lead to a quick sensory-motor processing in response to the unpredicted change (caused by the perturbation) in the sensory periphery.

A reduced connectivity from PM to M1 is shown at contralateral hemisphere in the early period. The PM is thought to be associated with predictions of sensory consequences of voluntary movements (Christensen et al., 2007). In the experiment, the subjects were required to maintain an isotonic wrist flexor torque before the perturbation. Thus, this voluntary control was accompanied with both the efferent motor command and an “efference copy” of this information (Wolpert and Flanagan, 2001). The communication between the M1 and PM is likely related to the cortical process of the efference copy to mediate movement predictions. This process may be inhibited due to the unpredicted change of sensory input, showing a decrease of effective connectivity from the PM to M1.

Additionally, a decreased effective connectivity is also shown from ipsilateral M1 to contralateral M1. The interhemispheric interaction of M1 has been reported by TMS and EEG studies during forearm muscle movement control (Ferber et al., 1992; Bönstrup et al., 2016). This interhemispheric inhibition is thought to be related to the activity of inhibitory GABA-ergic interneurons (Daskalakis et al., 2002) to prevent the interference from the opposite hemisphere (e.g., mirror movement) during movement control (Mayston et al., 1999). Thus, this inhibitory effect may facilitate the cortical response to the unpredictable perturbation without the interruption of ipsilateral M1. All the

information of modulation eventually flows into the contralateral M1 which allows the early cortical activity to be transmitted to the motor units through the monosynaptic corticospinal connection (Nielsen, 2016). This is the fastest cortical pathway contributing to the muscle stretch response, which likely lead to the rapid transcortical muscle response.

## Effective Connectivity during the Late Cortical Response to Muscle Stretch

In the late period of 100–350 ms, there are more cortical areas and connections are modulated (see **Figure 4B**). In particular, the connection between PPC and SMA is modulated, indicating that these cortical areas may play important roles in the late cortical responses to muscle stretch. The PPC is thought to be involved in the multisensory integration and coordinate transformations from sensory inputs to motor outputs during feedback-based movement control (Andersen and Buneo, 2002). The SMA is crucial for linking cognition to motor action (Nachev et al., 2008). The modulation of PPC-SMA connectivity indicates a high-level cognitive process for the slow, voluntary response, which is not shown for the early period. The reduced connectivity from PPC to SMA likely indicates a negative feedback in sensorimotor control loop. This negative feedback may play a role in correcting the motor actions based on the integrated sensory information. Besides, multiple pathways ending at the contralateral M1 are modulated in this period, indicating rich communications between different cortical areas. The complexity of this network in late period likely delays the voluntary motor output to facilitate the coordinated (slow) muscle responses.

## REFERENCES

- Abbruzzese, G., Berardelli, A., Rothwell, J. C., Day, B., and Marsden, C. (1985). Cerebral potentials and electromyographic responses evoked by stretch of wrist muscles in man. *Exp. Brain Res.* 58, 544–551. doi: 10.1007/BF00235870
- Andersen, R. A., and Buneo, C. A. (2002). Intentional maps in posterior parietal cortex. *Annu. Rev. Neurosci.* 25, 189–220. doi: 10.1146/annurev.neuro.25.112701.142922
- Babiloni, F., Cincotti, F., Babiloni, C., Carducci, F., Mattia, D., Astolfi, L., et al. (2005). Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function. *Neuroimage* 24, 118–131. doi: 10.1016/j.neuroimage.2004.09.036
- Bönstrup, M., Schulz, R., Feldheim, J., Hummel, F. C., and Gerloff, C. (2016). Dynamic causal modelling of EEG and fMRI to characterize network architectures in a simple motor task. *Neuroimage* 124, 498–508. doi: 10.1016/j.neuroimage.2015.08.052
- Campfens, S. F., Meskers, C. G., Schouten, A. C., Van Putten, M. J., and Van Der Kooy, H. (2015). Stretch evoked potentials in healthy subjects and after stroke: a potential measure for proprioceptive sensorimotor function. *IEEE Trans. Neural Syst. Rehabil. Eng.* 23, 643–654. doi: 10.1109/TNSRE.2015.2388692
- Chen, C.-C., Kilner, J. M., Friston, K. J., Kiebel, S. J., Jolly, R. K., and Ward, N. S. (2010). Nonlinear coupling in the human motor system. *J. Neurosci.* 30, 8393–8399. doi: 10.1523/JNEUROSCI.1194-09.2010
- Cheney, P., and Fetz, E. (1984). Corticomotoneuronal cells contribute to long-latency stretch reflexes in the rhesus monkey. *J. Physiol.* 349:249. doi: 10.1113/jphysiol.1984.sp015155
- Christensen, M. S., Lundbye-Jensen, J., Geertsen, S. S., Petersen, T. H., Paulson, O. B., and Nielsen, J. B. (2007). Premotor cortex modulates somatosensory cortex

## CONCLUSION

Muscle stretch modulates different effective cortico-cortical connections during early (before 100 ms post-perturbation) and late (after 100 ms) periods of cortical responses. Only a few effective cortico-cortical connections are modulated in the early period, while more cortical areas are involved in the late period with more effective connections modulated. The limited early cortical network likely allows for a rapid muscle response without involving high-level cognitive processes. The complexity of the late network may delay the voluntary motor output from the cortex, so as to facilitate the coordinated responses in the “standard” voluntary reaction to muscle stretch.

## AUTHOR CONTRIBUTIONS

YY and AS contributed in problem identification. BG conducted the data analysis under the supervision of YY and AS. YY and BG drafted the manuscript. YY and AS edited the manuscript.

## ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Research Council under the ERC advanced grant agreement n° 291339 (4D-EEG project). The authors would like to thank L. Eleftheriou for collecting the experimental dataset. The authors would like to thank all members of 4D-EEG consortium at Delft University of Technology, Northwestern University and VU University Medical Center Amsterdam for the useful discussions.

- during voluntary movements without proprioceptive feedback. *Nat. Neurosci.* 10, 417–419. doi: 10.1038/nrn1873
- Daskalakis, Z. J., Christensen, B. K., Fitzgerald, P. B., Roshan, L., and Chen, R. (2002). The mechanisms of interhemispheric inhibition in the human motor cortex. *J. Physiol.* 543, 317–326. doi: 10.1113/jphysiol.2002.017673
- David, O., Kiebel, S. J., Harrison, L. M., Mattout, J., Kilner, J. M., and Friston, K. J. (2006). Dynamic causal modeling of evoked responses in EEG and MEG. *Neuroimage* 30, 1255–1272. doi: 10.1016/j.neuroimage.2005.10.045
- 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
- Ferbert, A., Priori, A., Rothwell, J., Day, B., Colebatch, J., and Marsden, C. (1992). Interhemispheric inhibition of the human motor cortex. *J. Physiol.* 453, 525–546. doi: 10.1113/jphysiol.1992.sp019243
- Friston, K. J. (2002). Bayesian estimation of dynamical systems: an application to fMRI. *Neuroimage* 16, 513–530. doi: 10.1006/nimg.2001.1044
- Friston, K. J. (2011). Functional and effective connectivity: a review. *Brain Connect.* 1, 13–36. doi: 10.1089/brain.2011.0008
- Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *Neuroimage* 19, 1273–1302. doi: 10.1016/S1053-8119(03)00202-7
- Friston, K., and Penny, W. (2011). Post-hoc Bayesian model selection. *Neuroimage* 56, 2089–2099. doi: 10.1016/j.neuroimage.2011.03.062
- Goulden, N., Khusnulina, A., Davis, N. J., Bracewell, R. M., Bokde, A. L., McNulty, J. P., et al. (2014). The salience network is responsible for switching between the default mode network and the central executive network: replication from DCM. *Neuroimage* 99, 180–190. doi: 10.1016/j.neuroimage.2014.05.052
- Grefkes, C., Eickhoff, S. B., Nowak, D. A., Dafotakis, M., and Fink, G. R. (2008). Dynamic intra- and interhemispheric interactions during unilateral and



- bilateral hand movements assessed with fMRI and DCM. *Neuroimage* 41, 1382–1394. doi: 10.1016/j.neuroimage.2008.03.048
- Houk, J. C., Rymer, W. Z., and Crago, P. E. (1981). Dependence of dynamic response of spindle receptors on muscle length and velocity. *J. Neurophysiol.* 46, 143–166.
- Kaminski, M., and Blinowska, K. J. (1991). A new method of the description of the information flow in the brain structures. *Biol. Cybern.* 65, 203–210. doi: 10.1007/BF00198091
- Kiebel, S. J., David, O., and Friston, K. J. (2006). Dynamic causal modelling of evoked responses in EEG/MEG with lead field parameterization. *Neuroimage* 30, 1273–1284. doi: 10.1016/j.neuroimage.2005.12.055
- Lee, R., and Tatton, W. (1982). Long latency reflexes to imposed displacements of the human wrist: dependence on duration of movement. *Exp. Brain Res.* 45, 207–216. doi: 10.1007/BF00235780
- MacKinnon, C. D., Verrier, M. C., and Tatton, W. G. (2000). Motor cortical potentials precede long-latency EMG activity evoked by imposed displacements of the human wrist. *Exp. Brain Res.* 131, 477–490. doi: 10.1007/s002219900317
- Mayston, M. J., Harrison, L. M., and Stephens, J. A. (1999). A neurophysiological study of mirror movements in adults and children. *Ann. Neurol.* 45, 583–594. doi: 10.1002/1531-8249(199905)45:5<583::AID-ANA6>3.0.CO;2-W
- Meskers, C. G., Schouten, A. C., De Groot, J. H., De Vlugt, E., Van Hilten, B. J., Van Der Helm, F. C., et al. (2009). Muscle weakness and lack of reflex gain adaptation predominate during post-stroke posture control of the wrist. *J. Neuroeng. Rehabil.* 6:29. doi: 10.1186/1743-0003-6-29
- Nachev, P., Kennard, C., and Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nat. Rev. Neurosci.* 9, 856–869. doi: 10.1038/nrn2478
- Nielsen, J. B. (2016). Human spinal motor control. *Annu. Rev. Neurosci.* 39, 81–101. doi: 10.1146/annurev-neuro-070815-013913
- Penny, W. D., Stephan, K. E., Mechelli, A., and Friston, K. J. (2004). Comparing dynamic causal models. *Neuroimage* 22, 1157–1172. doi: 10.1016/j.neuroimage.2004.03.026
- Perenboom, M., Van De Ruit, M., De Groot, J., Schouten, A., and Meskers, C. (2015). Evidence for sustained cortical involvement in peripheral stretch reflex during the full long latency reflex period. *Neurosci. Lett.* 584, 214–218. doi: 10.1016/j.neulet.2014.10.034
- Porcaro, C., Coppola, G., Pierelli, F., Seri, S., Di Lorenzo, G., Tomasevic, L., et al. (2013). Multiple frequency functional connectivity in the hand somatosensory network: an EEG study. *Clin. Neurophysiol.* 124, 1216–1224. doi: 10.1016/j.clinph.2012.12.004
- Pruszynski, J. A., and Scott, S. H. (2012). Optimal feedback control and the long-latency stretch response. *Exp. Brain Res.* 218, 341–359. doi: 10.1007/s00221-012-3041-8
- Pruszynski, J. A., Kurtzer, I., and Scott, S. H. (2011). The long-latency reflex is composed of at least two functionally independent processes. *J. Neurophysiol.* 106, 449–459. doi: 10.1152/jn.01052.2010
- Rigoux, L., Stephan, K. E., Friston, K. J., and Daunizeau, J. (2014). Bayesian model selection for group studies—revisited. *Neuroimage* 84, 971–985. doi: 10.1016/j.neuroimage.2013.08.065
- Rocco-Donovan, M., Ramos, R. L., Giraldo, S., and Brumberg, J. C. (2011). Characteristics of synaptic connections between rodent primary somatosensory and motor cortices. *Somatosens. Mot. Res.* 28, 63–72. doi: 10.3109/08990220.2011.606660
- Schwartz, A. B. (2016). Movement: how the brain communicates with the world. *Cell* 164, 1122–1135. doi: 10.1016/j.cell.2016.02.038
- Scott, S. H. (2002). Optimal strategies for movement: success with variability. *Nat. Neurosci.* 5, 1110–1111. doi: 10.1038/nn1102-1110
- Scott, S. H. (2004). Optimal feedback control and the neural basis of volitional motor control. *Nat. Rev. Neurosci.* 5, 532–546. doi: 10.1038/nrn1427
- Shibasaki, H. (2012). Cortical activities associated with voluntary movements and involuntary movements. *Clin. Neurophysiol.* 123, 229–243. doi: 10.1016/j.clinph.2011.07.042
- Stephan, K. E., Penny, W. D., Moran, R. J., Den Ouden, H. E., Daunizeau, J., and Friston, K. J. (2010). Ten simple rules for dynamic causal modeling. *Neuroimage* 49, 3099–3109. doi: 10.1016/j.neuroimage.2009.11.015
- Szameitat, A. J., Shen, S., Conforto, A., and Sterr, A. (2012). Cortical activation during executed, imagined, observed, and passive wrist movements in healthy volunteers and stroke patients. *Neuroimage* 62, 266–280. doi: 10.1016/j.neuroimage.2012.05.009
- Vlaar, M. P., Mugge, W., Groot, P. F., Sharifi, S., Bour, L. J., Van Der Helm, F. C., et al. (2016). Targeted brain activation using an MR-compatible wrist torque measurement device and isometric motor tasks during functional magnetic resonance imaging. *Magn. Reson. Imaging* 34, 795–802. doi: 10.1016/j.mri.2016.02.002
- Wolpert, D. M., and Flanagan, J. R. (2001). Motor prediction. *Curr. Biol.* 11, R729–R732. doi: 10.1016/S0960-9822(01)00432-8
- Yang, Y., Solis-Escalante, T., Van De Ruit, M., Van Der Helm, F. C., and Schouten, A. C. (2016a). Nonlinear coupling between cortical oscillations and muscle activity during isotonic wrist flexion. *Front. Comput. Neurosci.* 10:126. doi: 10.3389/fncom.2016.00126
- Yang, Y., Solis-Escalante, T., Yao, J., Van Der Helm, F. C. T., Dewald, J. P. A., and Schouten, A. C. (2016b). Nonlinear connectivity in the human stretch reflex assessed by cross-frequency phase coupling. *Int. J. Neural Syst.* 26:1650043. doi: 10.1142/S012906571650043X

**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 Yang, Guliyev and Schouten. 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.



# Decoding English Alphabet Letters Using EEG Phase Information

YiYan Wang<sup>1,2</sup>, Pingxiao Wang<sup>2</sup> and Yuguo Yu<sup>1\*</sup>

<sup>1</sup> State Key Laboratory of Medical Neurobiology, School of Life Science and the Collaborative Innovation Center for Brain Science, Center for Computational Systems Biology, Institutes of Brain Science, Fudan University, Shanghai, China, <sup>2</sup> Institute of Modern Physics, Fudan University, Shanghai, China

## OPEN ACCESS

### Edited by:

Dipankar Roy,  
National Brain Research Centre  
(NBRC), India

### Reviewed by:

Julian Keil,  
Christian-Albrechts-Universität zu Kiel,  
Germany  
Keisuke Kawasaki,  
Niigata University, Japan

### \*Correspondence:

Yuguo Yu  
yuyuguo@fudan.edu.cn

### Specialty section:

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

**Received:** 17 August 2017

**Accepted:** 25 January 2018

**Published:** 07 February 2018

### Citation:

Wang Y, Wang P and Yu Y (2018)  
Decoding English Alphabet Letters  
Using EEG Phase Information.  
*Front. Neurosci.* 12:62.  
doi: 10.3389/fnins.2018.00062

Increasing evidence indicates that the phase pattern and power of the low frequency oscillations of brain electroencephalograms (EEG) contain significant information during the human cognition of sensory signals such as auditory and visual stimuli. Here, we investigate whether and how the letters of the alphabet can be directly decoded from EEG phase and power data. In addition, we investigate how different band oscillations contribute to the classification and determine the critical time periods. An English letter recognition task was assigned, and statistical analyses were conducted to decode the EEG signal corresponding to each letter visualized on a computer screen. We applied support vector machine (SVM) with gradient descent method to learn the potential features for classification. It was observed that the EEG phase signals have a higher decoding accuracy than the oscillation power information. Low-frequency theta and alpha oscillations have phase information with higher accuracy than do other bands. The decoding performance was best when the analysis period began from 180 to 380 ms after stimulus presentation, especially in the lateral occipital and posterior temporal scalp regions (PO7 and PO8). These results may provide a new approach for brain-computer interface techniques (BCI) and may deepen our understanding of EEG oscillations in cognition.

**Keywords:** brain-computer interface, support vector machine (SVM), human brain, theta-band oscillation, visual cortex

## INTRODUCTION

The past decade has witnessed great developments in brain-computer interfaces (BCIs), aiming to help severely physically impaired patients interact with the external world through tasks such as typing letters of the English alphabet on a computer for communication. Studies have applied stimulus-evoked brain electroencephalogram (EEG) or electrocorticography (ECoG) signals, especially event-related potentials (ERPs) with P300 responses (Zhang et al., 2013) and steady-state visually evoked potentials (SSVEP) (Won et al., 2014; Nezamfar et al., 2016), to discriminate stimulus characteristics such as letters. There is increasing evidence that the frequency-related phase pattern and power of neural oscillations may code significant sensory information relevant to human perception of the external world, especially in low-frequency bands (Luo and Poeppel, 2007; Schyns et al., 2011; Wang et al., 2012; ten Oever and Sack, 2015). Luo et al. (Luo and Poeppel, 2007) demonstrated that the phase pattern of theta-band (5–8 Hz) activities from the human auditory cortex contains information used to discriminate spoken sentence signals. Their findings indicated a approximately 200 ms time window (approximately 5 Hz within the theta rhythm) that may be

critical for discrete perceptive processes. Subsequent phase-decoding studies in audio perception have observed that a similar oscillation frequency range (3~7 Hz) is dominant in spoken sentence recognition (Luo and Poeppel, 2007; Howard and Poeppel, 2010; Wang et al., 2012; Ng et al., 2013; ten Oever and Sack, 2015). Ng et al. (2013) demonstrated that stimuli can be discriminated by the firing rates and phase patterns but not by the oscillation amplitude. Another recent study presented evidence that syllables with varying visual-to-auditory delays are preferably processed at different oscillatory phases (ten Oever and Sack, 2015). Wang et al. (2012) employed the scalp tangential electric field and the surface Laplacian operator around the auditory cortical area to improve the recognition rate of English phonemes. They built a complicated bootstrap-based method that achieved 53% accuracy for all eight phonemes and showed that phase sequences performed better. Also revealed that changes in the amplitude (Worden et al., 2000; van Dijk et al., 2008) and phase (Vanrullen et al., 2011) of ongoing alpha activities (9–12 Hz) several hundred milliseconds before a stimulus can modulate the visual discrimination level. In fact, more recent evidence suggests that decreased alpha power may be tightly correlated to the increase in the visual baseline excitability level, which may serve to improve task performance (Lange et al., 2013; Iemi et al., 2017).

The above studies suggest the importance of the frequency, phase, and amplitude of slow oscillatory activities in object representation and categorization (Fries et al., 2007; Schyns et al., 2011). For example, the oscillatory power of various frequency bands may serve to modulate sensory excitability and attention (Klimesch, 1999; Engel et al., 2001; van Dijk et al., 2008), while oscillatory phase patterns across theta and gamma bands may be engaged in information processing, visual attention and working memory (Lisman and Idiart, 1995; Siegel et al., 2009; Heusser et al., 2016).

In this study, we examined the possibility of employing EEG phase and power signals to discriminate input stimulus for a brain-computer interface (BCI) approach. We chose the English alphabet as the visual stimulus because it is a “model” stimulus in BCI research. Based on the above experimental studies (Luo and Poeppel, 2007; van Dijk et al., 2008; Busch et al., 2009; Canolty and Knight, 2010; Schyns et al., 2011; VanRullen and Macdonald, 2012; Wang et al., 2012; ten Oever and Sack, 2015; Watrous et al., 2015; Heusser et al., 2016; Tomassini et al., 2017), which presented evidence on how the oscillatory parameters (phase, power, and frequency) may code visual and auditory information, we hypothesize that information from the visual presentation of different letters in the English alphabet may be encoded in EEG low-frequency phase patterns. Phase decoding and statistical machine-learning analysis may be a novel method, in addition to the traditional ERP method, for discriminating visualized letters. This may be of great benefit for the development of BCI techniques. In addition, it is believed that visual information first flows through the primary visual cortex and then up to higher levels such as V3/4 TEO and TE, which is called the ventral pathway in object recognition tasks (Tanaka, 1996; Krüger et al., 2013). The ventral pathway was thought to be particularly important for reading, including word

and letter recognition (Price and Devlin, 2011). Therefore, we questioned whether there was a classification accuracy difference between the scalp occipital and scalp tempo-occipital regions. To examine the above issues, a simple BCI protocol was designed in which subjects watched randomly selected letters on a computer monitor. EEG data were collected from each subject, and an analysis was applied to determine whether visual letter stimuli could be discriminated based on the EEG phase pattern and power amplitude.

## MATERIALS AND METHODS

### Subjects

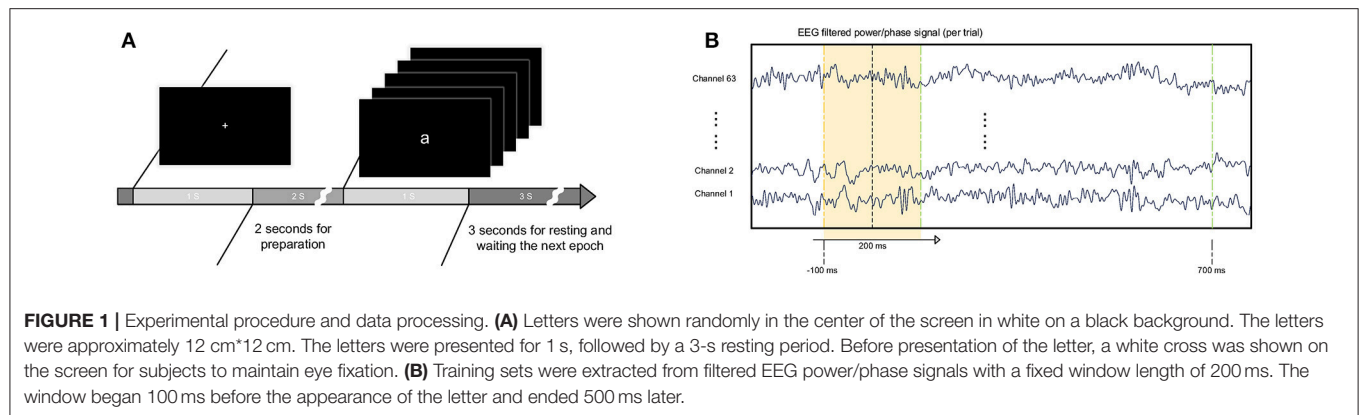
Fourteen right-handed students from Shanghai Fudan University were recruited by providing monetary compensation. Right-handedness was determined using the Edinburgh handedness inventory (Oldfield, 1971). All subjects (nice males and five females, mean age 25.4, range: 21–32) had normal color vision, corrected visual acuity and no history of neurological or psychiatric problems. This study was approved and supervised by the Ethics Committee of the School of Life Sciences at Fudan University (No. 290). All participants signed written informed consent.

### EEG Recordings and Experimental Design

The EEG data were recorded with a 500 Hz sampling rate in a sound-proof room using a 64-channel actiCHamp Brain Products recording system (Brain Products GmbH, Inc., Munich, Germany) relative to a Cz reference signal. The ground electrode was placed on the Fz electrode. The impedance levels were maintained below 10 kohm.

The stimuli were presented using a pre-programmed e-prime protocol. Five lowercase letters, “a,” “e,” “l,” “o,” and “t” were chosen as the letters to be visually presented on the computer screen. The letter “t” was chosen to exclude pronunciation peculiarity because the remaining four letters were vowels. The letters were in white Times New Roman font and presented on an approximately 12 cm\*12 cm black background, in a field of view (FOV) of 6.88 degrees. The subjects sat one meter away from a 23-inch screen. The screen was adjusted as high as the height of the seated subject so that the subjects could keep their eyes horizontal. The subjects were directed to focus on the screen and not to move their heads. When a letter was presented, the subjects were directed to read it silently without mouth action. This was intended to keep the subject focused and to avoid any myoelectric artifacts. The participants were instructed to minimize eye movements during the visual presentation and to fixate on the center.

**Figure 1** presents the experimental protocol. In each trial, a randomly displayed letter appeared on the screen for 1 s and was followed by a 3-s blank interval. Before the appearance of the letter, the subjects were directed to focus their eyes on a white cross on the screen for 1 s. In the study, the subject watched five letters appear individually randomly for 450 trials. The 450 trials were divided into three blocks, with each block containing 150 trials. At the beginning of each block, an instruction was presented on the screen, and the program was paused until the



subject pressed the “enter” button to continue. In each block, the letters randomly appeared 150 times, with each letter for 30 times. Between each block, the subject had a short break and then chose when to continue the next study block. It took approximately 60 min to finish three blocks. Between each block, the recording was paused, and the electrode conductance was examined. The mean of the successful trials used for analysis is  $351 \pm 55$  (mean and SD) over all subjects.

## Data Preprocessing Analysis

Data preprocessing analysis was performed using EEGLAB (Delorme and Makeig, 2004) and included bandpass filtering (0.5–220 Hz), epoch extraction locked to the onset of the letters (–500 to 1,000 ms) and baseline correction (–500 to 0 ms). To avoid confusion, we called these data “wide-band data” to differentiate them from the later narrow-band filtered data such as the alpha band EEG data. Signal artifacts were removed in two steps. First, the data were visually inspected, and epochs containing artifacts such as extremely high-amplitude electrode cable movement-induced fluctuations were rejected. Second, epochs containing typical eye movements and eye-blink artifacts that occurred during the first 800 ms after the onset of the letters were rejected. An independent component analysis (ICA) was applied to decompose the EEG data. After decomposition, 63 time-sequence data of component activations were obtained that corresponded to 63 recording channels for each subject. These component activations were recognized as EEG activity or non-brain artifacts by visual inspection of their scalp topographies, time courses, and frequency spectra. The artifact components related to heart beats, temporal muscle movement, eye movements and eye blinks were removed. The criteria for categorizing component activations as EEG activity included the following: (1) spectral peak(s) at typical EEG frequencies and (2) similar responses across each trials; i.e., an EEG response should not occur in a small number of trials only (Delorme and Makeig, 2004). Based on these criteria, the component activations representing non-brain artifacts were removed (the removed ICAs are  $11.07 \pm 8.62$ , mean and SD, for 14 subjects), and the EEG data were reconstructed from the remaining component activations.

We then employed the Hilbert transform to convert the real-time artifact-cleaned EEG sequence into a complex time sequence. Each complex number has amplitude and angle

information. We derived the amplitude sequence  $A(t)$  and phase sequence  $P(t)$  separately. Then, we applied machine-learning analysis based on the amplitude or phase sequence data. The formula for the Hilbert transform is presented here:

$$Y(t) = H(x(t)) = \int_{-\infty}^{+\infty} x(\tau) * \frac{1}{t - \tau} d\tau$$

Hilbert transformation converts the raw real signal into an imaginary counterpart, and these two parts make a complex signal. The power sequence is defined as the magnitude of this complex signal, and the phase sequence is its phase angle.

Moreover, delta (1–4 Hz), theta (4–8 Hz), alpha (8–14 Hz), beta (14–30 Hz), and gamma (30 Hz above) band oscillations are five typical rhythms observed in the cortex and are thought to be closely related to cognition processes (Kahana et al., 2001; Colgin et al., 2009; Fries, 2015). Additionally, the gamma oscillation can be further divided into low-gamma (30–50 Hz) and high-gamma (50–150 Hz) oscillations. To investigate the functional role of these oscillations in letter classification performance, the original epoched EEG response was filtered into these six bands using a Kaiser window linear phase FIR filter in the MATLAB FDA toolbox. The stop bands were set to attenuate the signal magnitude at –30 dB with a 1 Hz edge band. A Hilbert transformation was then applied to the filtered data.

## Multi-class Classification Analysis and Gradient Ascent Approach

Five-class classification was employed to discriminate the five letters and to investigate the possibility that the EEG phase pattern or power pattern could be used as a feature in EEG-based BCI. A supervised machine-learning algorithm, LIBSVM, a library for support vector machine (SVM) classifiers (Chang and Lin, 2011), was used and implemented in the MATLAB toolbox. The classifications were quinary with a chance level of 20 percent, and the results of these quinary predictions were evaluated electrode by electrode. The Gaussian function was used as the nonlinear transform function in the SVM classifier, and its critical parameter sigma was determined using a gradient ascent approach, which is similar to the steepest descent algorithm, in which the parameter is adaptively adjusted according to the changes in classification accuracy to ensure that it can be



maximized. According to previous research (Schyns et al., 2011), visual stimuli-evoked EEG responses were most informational in the occipital and occipital-temporal cortices. Therefore, the focus was on these 17 electrode sites: P7, P5, P3, P1, Pz, P2, P4, P6, P8, PO7, PO3, POz, PO4, PO8, O1, Oz, and O2. The additional methodological steps encompassing the computational strategy for validating the classification results (cross-validation and shuttered label training sets) are described below.

## Cross-Validation Approaches and Shuffled-Label Training Sets

Cross-validation of the multiclass classification analysis was conducted to obtain robust estimates of the discrimination accuracies and to test the generalization ability of our classifier. In this study, a 30-fold cross-validation approach was adopted. The EEG signal sets were randomly divided into 30 parts, and 29 parts were chosen to train the SVM, which was subsequently used to test the remaining set to obtain the discrimination accuracy (Please note that there are total 450 trials corresponding to five letters for one subject. The 450 trials were divided into 30 parts, with each part contains 15 trials for five letters). This procedure was repeated 30 times, averaging each repetition's accuracy to obtain the final accuracy. To exclude the artificial classification effect caused by the adoption of the SVM classifier and to estimate the validity of the classification result, the labels that indicated the letter for each trial were randomly shuffled 100 times to form 100 random label-training sets. A multiclass classification analysis with a 30-fold cross-validation approach was used on these random label-training sets, and a random label training result ensemble was obtained. In each turn, a subject was randomly selected and the labels of the letters was randomly shuffled. After that, we chose the highest classification accuracy across the electrodes. And then we did this process one-hundred times. Which means we had 100 random-labeled accuracies. We called this a random-label classification accuracies ensemble. A Kolmogorov-Smirnov test (K-S test) was conducted on this ensemble to determine whether the ensemble satisfies a supposed distribution, such as a norm distribution, and if so, to determine its mean value and variance. Finally, the statistical significance was calculated ( $p < 0.0013$ , three sigma standard) based on the mean and variance of this permuted accuracy.

For comparisons of classification accuracy difference between phase and power groups data of 17 electrodes with 12 subjects, we have performed two-way anova analysis and then performed all the pairwise comparisons using Tukey-Kramer's multiple compare method (Specifically, we first applied `[p,~,stats]=anova2(data,12)` in Matlab. Data is a 24\*17 matrix, with the first 12 lines are power accuracy values from 12 subjects, while lines from 13 to 24 are phase accuracy values from 12 subjects; and 17 corresponds to 17 electrodes. Then we have performed multiple comparison with: `C = multcompare(stats)` in Matlab, default is Tukey-Kramer method). Tukey-Kramer Multiple comparison method is one of the best methods for all-possible pairwise comparisons of group means, to determine which are significantly different from which others. Multiple

comparison procedure was performed for significant analysis of pairwise comparison results.

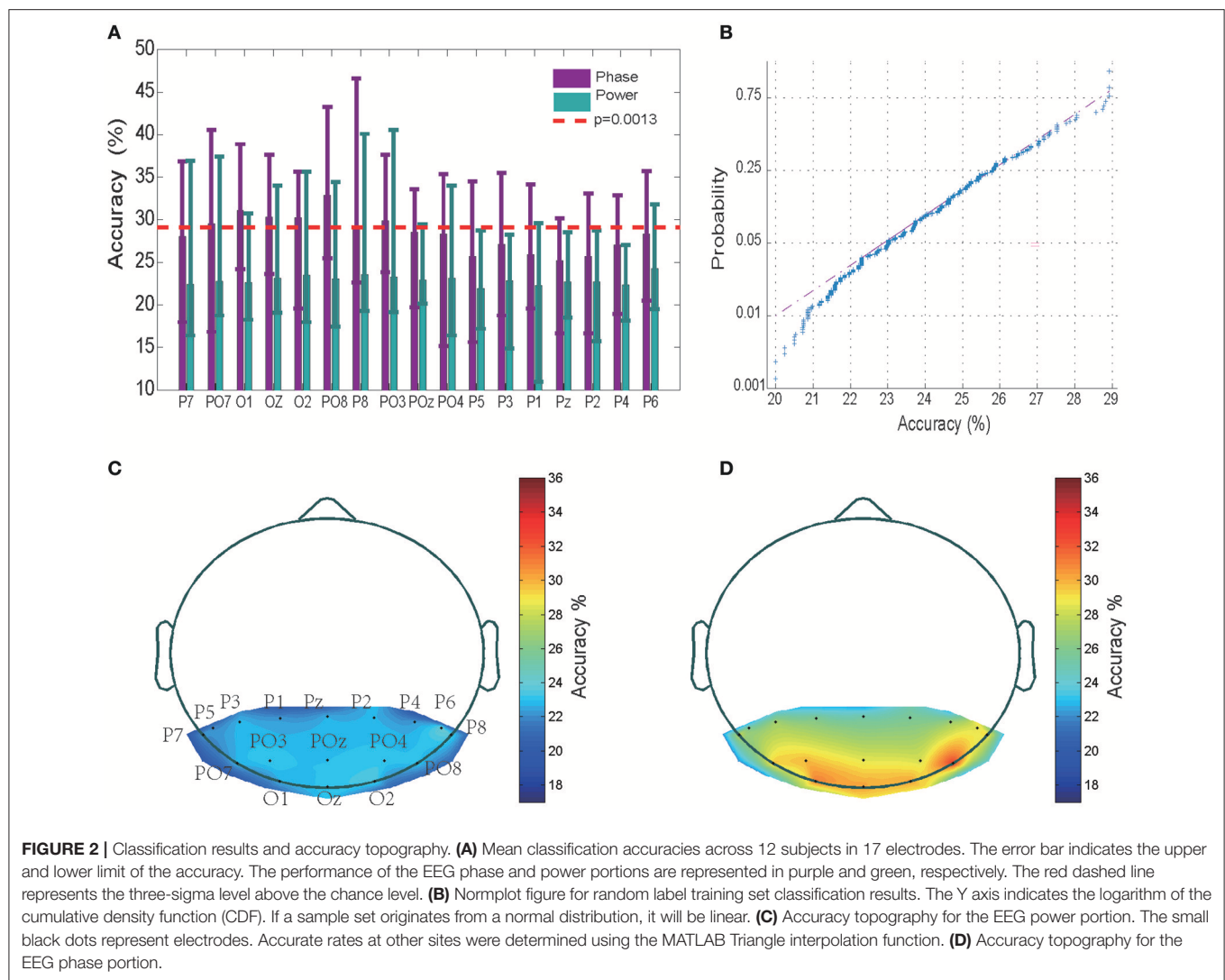
To understand the analysis procedure in a clear way, please see the flowchart Figure S2.

## RESULTS

### Classification Accuracy for Wide-Band EEG Phase and Power Sequences

The power and phase sequences were both 1,500 ms (starting at  $-500$  ms before the appearance of "letter" and stopping at the end of "letter"), and a short 200 ms portion (starting at the 100th ms after the appearance of "letter") was selected for classification accuracy analysis. The reason starting at the timing of 100th ms is based on the following analysis result.

The timing of the appearance of a "letter" is set as 0th ms. Using this 0th ms timing as the starting point, we chose the sequence of different sizes of time window to examine where the valuable information is started to be encoded. The tested time period is from 0 to 600 ms with time step equal to 2 ms. We observed that the classification accuracy is around chance level for the time period  $< 100$  ms, while the accuracy increased rapidly to a 31% high value as the time period was increased to 200 ms, and then fluctuated to reach a saturation level when the time period was further increased to 600 ms (see Figure S1). This analysis suggests that the EEG sequence  $< 100$ th ms may not contain valuable information. Therefore, in the following, the classification accuracy values were obtained by training a SVM classifier using 200 ms EEG power/phase sequences that started at the 100th millisecond after presentation of a letter. The mean and variance of the classification accuracy of each of the 17 electrodes for all 12 subjects are shown in **Figure 2A** (data for the remaining 2 subjects without significant classification power are shown separately in Figure S3). The highest accuracy was 46.61% (chance level of 20%) for a wide-band (0.5–220 Hz) EEG phase sequence (**Figure 2A**). The EEG phase sequence in 17 electrodes of 12 subjects ( $28.42 \pm 3.21$ , mean  $\pm$  SD) showed significantly higher correct rates than the EEG power sequence ( $22.89 \pm 3.02$ , mean  $\pm$  SD) at a  $p < 10^{-9}$  confidence level (two way ANOVA analysis with Tukey-Kramer multiple comparison correction conducted in MATLAB). This implies that the EEG phase portion contains more information than EEG power portion. Multiple comparison procedure was performed for significant analysis of pairwise comparison results, and PO8 was observed to have significantly higher accuracies than P1, P2, P5, Pz ( $0.01 < P < 0.05$ ) while no significant difference was observed between any pair of accuracy values of other electrodes for phase sequences. The confidence interval was determined using the variance of a fully random shuffled label training set classification accuracy. **Figure 2B** shows the normplot figure for random label training set classification results. The Y axis indicates the logarithm of the cumulative density function (CDF). The regression linear fitting analysis suggests that the classification accuracy values  $< 29\%$  are mainly from a normal distribution (K-S test  $p = 0.038$ ). The mean was 23.81%, and the variance was 1.76%; thus, the three-sigma level was 29.09%.



This value was set as the confidence interval with a one-tail confidence level  $P = 0.0013$  (see red dashed line in **Figure 2A**). We observed that 12 of the 14 subjects with 450-trial tests had significant classification power above the three-sigma level, with 29.09% accuracy in at least one electrode; further, 8 subjects had three electrodes, and seven subjects had five powerful electrodes that showed significant classification power  $>29.09\%$ . We also conducted phase and power decoding analyses of the data from the 2 subjects who did not have electrode data with significant classification effects (see **Figure S3**). The highest accuracy for these subjects was only 29% for the phase classification (**Figures S3A,B**) and 27% for the power classification (**Figure S3C**). The mean accuracy value of the phase decoding for all 17 electrodes for the 12 subjects was  $28.42 \pm 3.21$  (mean  $\pm$  SD) and  $27.71 \pm 3.45$  for all 14 subjects. Hence, the following results analysis were mainly based on the 12 subjects. The analysis of the 2 subjects with no significant effects are shown separately in **Figures S3, S4**.

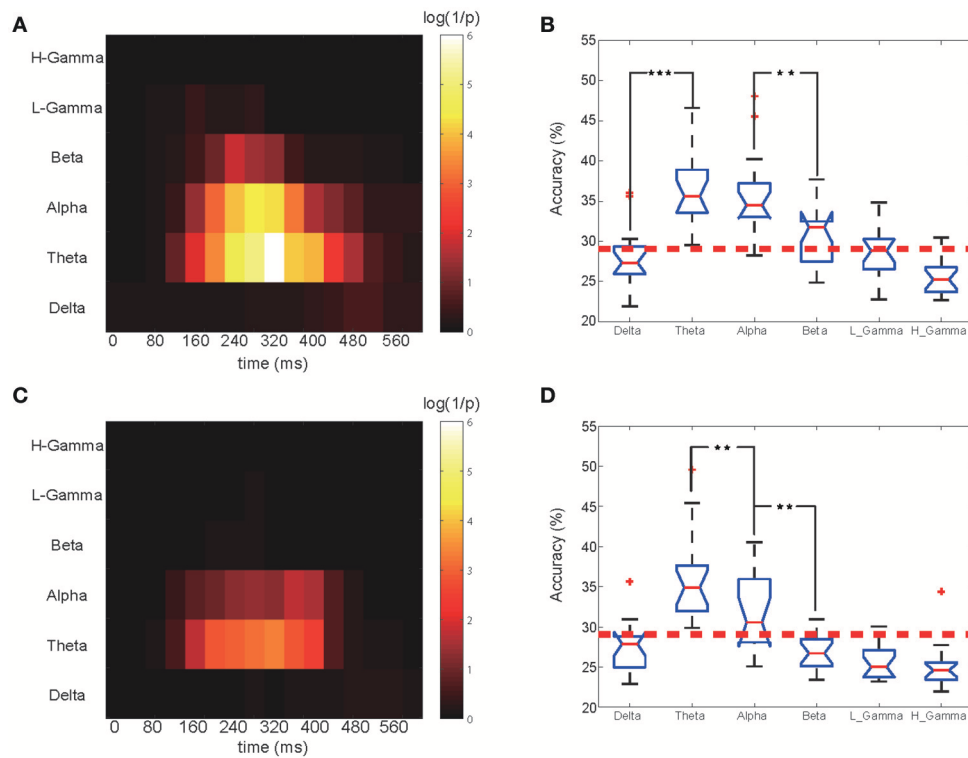
As is shown in **Figures 2C,D** for the averaged spectrum of 12 subjects with at least 1 electrode with significant classification power, the relatively high classification accuracy

appeared in electrodes placed in the left and right posterior regions.

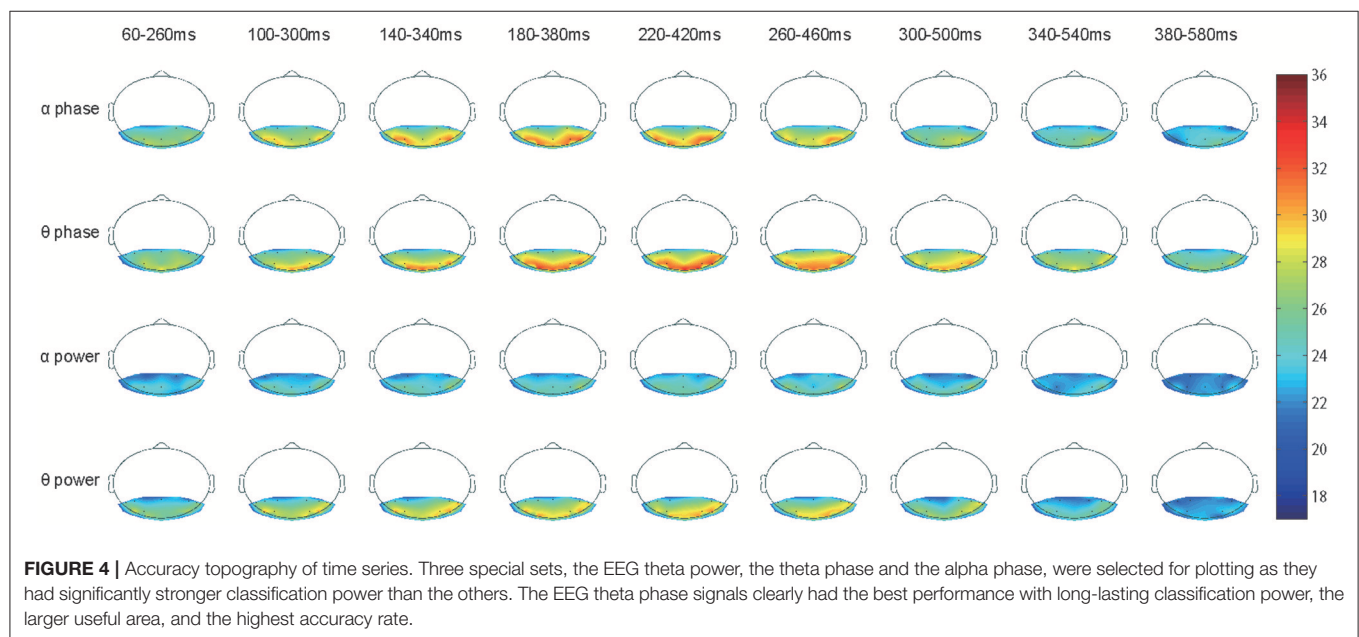
## Different EEG Frequency Bands and Period-Specific Classification Results

To examine the critical period for classification, a shifting 200 ms-long window (from  $-100$  to  $500$  ms, 40 ms per step) was applied to the frequency-filtered power and phase time-courses to extract the training and test sets. We observed that the discrimination accuracy within the first 100 ms period after the presentation of a letter is always approximately equal to chance, while most of the valuable decoded information is in the first half-second period (100–600 ms) after the stimuli's presentation (see **Figures 3, 4**). Hence, our analysis suggested that starting at the 100th millisecond mark after the presentation of a letter may result in a higher classification power than analysis starting from 0 ms after the presentation of a letter (van Gerven et al., 2013; Watrous et al., 2015).

The training and classification processes were employed on these frequency- and time-specific phase signal and power sets



**FIGURE 3 |** Time-frequency classification significance diagram and bands comparison. **(A)** Classification significance for the EEG phase portion in different bands and time periods. for the 12 subjects shown in **Figure 2**. Each small block represents a 200 ms training set. For a particular band and time period, the highest accuracy among all 17 electrodes was chosen and its corresponding  $P$ -value was calculated. The X ticks indicate each periods midpoint, from 0 to 600 ms. **(B)** Comparison of the EEG phase portion classification performance for all bands for the selected optimal time period. One star corresponds to a  $P < 0.05$  significance level for the related two bands, and three stars corresponds to  $P < 0.001$ . **(C)** Classification significance for the EEG power portion in different bands and time periods. **(D)** Comparison of the classification performance of the EEG power portion of all bands.



**FIGURE 4 |** Accuracy topography of time series. Three special sets, the EEG theta power, the theta phase and the alpha phase, were selected for plotting as they had significantly stronger classification power than the others. The EEG theta phase signals clearly had the best performance with long-lasting classification power, the larger useful area, and the highest accuracy rate.

to calculate the mean accuracy over 12 subjects in which we obtained significant results in the previous analysis step. A 2-dimensional accuracy matrix was obtained with the X ticks representing the medial time point of each shifting window and the Y ticks representing all six bands. The classification accuracies were transformed into their *P*-value representations. The *P* value was calculated as the probability that the frequency-filtered power and phase time-courses' accuracy rate can occur from a norm distribution that we obtained from the shuffled label training sets. For a higher accuracy rate, a smaller *P*-value would be obtained. A denary logarithm of  $1/P$  was calculated and chosen as the presentation of the classification performance for illustration purposes. We next compared each frequency band's performance. We selected the best performing time block for each band. **Figures 3A,C** shows the calculated classification significance as a function of time for six bands. Using the calculation, the best performing time block was chosen based on the highest classification significance level for each frequency band, and the corresponding accuracy value was obtained for the same time block. Then, we applied the MATLAB ANOVA toolbox to examine whether these six bands' signals had significantly different classification performance. The EEG phase signal and power signal portions were treated separately.

The phase and power information in different EEG oscillatory band frequencies that contribute to the classification were also studied. **Figure 3A** shows the results of the calculation of the classification significance based on the EEG phase signal, and **Figure 3B** shows a quantification of its classification performance for the 12 subjects who had significant classification power (data for the remaining 2 subjects without significant classification power are shown in Figure S4). The X ticks represent the mid-time point of each shifting 200-ms-long window, which started at 0 ms and ended at 600 ms. As shown in **Figure 3A**, the higher the logarithm value, the higher the accuracy rate it represents. We also calculated the classification significance and performance values based on the EEG power information (**Figure 3C**). For both the EEG power and the phase coding performance, the theta frequency band showed higher classification performance than did the remaining five bands, and the crucial time period began at 60 ms to 580 ms (with a middle time point of 160–480 ms). We found that for theta band, phase part and power part had no significant difference (MATLAB *ttest2*,  $P = 0.89$ ). While in alpha band, phase sequence had a significantly higher accuracy than its power counterpart (*ttest2*,  $P = 0.0341$ ). Also the beta band performed differently ( $P < 0.001$ ).

For both the theta and the alpha frequency bands, the significance and performance levels are generally relatively lower in the power coding than the phase coding (**Figure 3C**). The highest accuracy appeared in the period from 220 to 420 ms for phase coding at the theta band and at 180 to 380 ms for the alpha band.

**Figures 3B,D** shows the calculated classification accuracy for different frequency bands based on EEG oscillatory phase and power components. The EEG rhythmic frequencies significantly influenced the classification accuracy [ $F_{(5, 96)} = 22.64$ ,  $P < 10^{-6}$  MATLAB ANOVA1]. **Figure 3B** shows that, for EEG phase

coding, there was no significant difference in classification between the theta ( $36.70 \pm 4.43$ , mean  $\pm$  SD) and alpha bands ( $35.4 \pm 4.21$ ), but there was a significant difference between the alpha ( $35.40 \pm 4.21$ ) and beta bands ( $30.74 \pm 4.32$ ) ( $p = 0.0037$ , ANOVA1) for the 12 subjects. **Figure 3D** shows that, for power coding, the EEG theta band ( $35.08 \pm 5.32$ ) accuracy was significantly higher than the alpha band ( $31.67 \pm 4.29$ ) accuracy and that the alpha band accuracy was significantly higher than that of the other four frequency bands. The remaining four bands did not show a significant classification effect. In addition, if the data analysis includes the two non-significant subjects, the phase decoding accuracy value for the theta band for all 14 subjects was  $35.50 \pm 5.08$ , which was slightly lower than the  $36.70 \pm 4.43$  result for the 12 subjects.

## Accuracy Topology Map for Shifting Time Window Data

Based on our current decoding methods, we would like to examine the spatial-temporal distribution of classification accuracy values. Here, we focus on the alpha and theta bands because they showed significantly high classification accuracy (**Figures 3B,D**). The accuracy values from the 12 subjects were averaged and represented in color (see **Figure 4**). **Figure 4** shows the classification accuracy map derived from both phase and power information in the alpha and theta bands for the 17 electrodes as a function of time.

Unlike the results shown in **Figure 2D**, there was no strong accuracy lateralization for right hemisphere electrodes, only slightly longer lasting classification power (e.g., the alpha band phase signal from 260 to 460 ms and the theta band phase signal from 300 to 500 ms). The classification power of electrode PO7 had faded but was still in electrode PO8). Interestingly, electrodes O1, O2, and O3 also achieved very high accuracy rates, as PO7 and PO8 did in the theta band phase signal, but presented low values in the alpha band. This difference implies that the theta and the alpha signals may play distinct roles in recognition and have different origins (Fries, 2015).

The classification power in all 17 electrodes clearly faded after 380 ms, and the accuracy decreased to a chance level. Therefore, the remaining topographic maps are not shown.

## DISCUSSION

### Comparison with Existing BCI Methods and Other Phase Coding Research

This study revealed that the phase patterns and power in the theta and alpha bands may contain valuable information about the input stimulus features. This valuable temporal phase coding approach was confirmed with a conclusion consistent with the most recent investigations into decoding other visual and auditory signals in multiple behavior and cognition tasks (Luo and Poeppel, 2007; Schyns et al., 2011; Vanrullen et al., 2011; Wang et al., 2012; ten Oever and Sack, 2015). In addition, decoding of phase and power sequences in different frequency bands suggests different classification powers. Decoding the phase patterns in theta and alpha oscillations provided relatively



higher discrimination accuracy than did the delta, beta and gamma band oscillations. Previous studies suggested that the ventral occipital-temporal (vOT) cortex is involved in the perception of visually presented objects and written words (Dahaene, 1995; Price and Devlin, 2011; Matsuo et al., 2015). Our decoding analysis showed a higher classification power for electrodes placed in occipital-temporal regions compared to other regions, although we should keep in mind that EEG electrodes do not necessarily pick up activity directly under the electrodes. These results provide more evidence to support EEG phase coding in visual perception. Spatially distributed electrodes may encode different preferred stimulus features in this process.

The method used here is not as general as the classic existing BCI methods such as SSVEP and P300 (Zhang et al., 2013; Nezamfar et al., 2016). It also relies on the training of an SVM classifier. The traditional BCI approach often conducts the decoding process in real time. In our approach, we first collected a sufficient amount of EEG response data to input stimuli and then performed the training and decoding processes. In future research, we would expect the faster computer speeds and improved algorithms to allow this decoding approach to occur in real time. In addition, compared with existing BCI approaches, our approach is more reliant on subjects. The performance varied greatly between subjects, similar to the ERD/ERS approach. This implies that we may train the subject in future research to improve the classification performance as in some ERD/ERS research.

Although few studies focus on an EEG phase decoding approach and its performance is not sufficient to evoke more attention, the phase decoding method showed a promising prospect for decoding human brain activities using the mass electromagnetic field. As suggested recently (Panzeri et al., 2015, 2016), this new method and other related methods can be used extensively to improve BMIs, and its performance may be further improved by more sophisticated designs.

Our experimental results are consistent with a previous phase decoding investigation related to an emotional face discrimination EEG experiment (Schyns et al., 2011). Almost similar spatially located electrodes in the theta frequency band and a similar critical time window were obtained. This may suggest a similar cortical pathway involved in the visualization process of alphabet letters and human faces. This similarity also appeared in human fMRI recording (Dehaene and Cohen, 2011). However, in contrast to the face recognition process, our experimental results might include an auditory coding effect in addition to the visualization process. Participants were asked to sit quietly without vocalizing the letters, however, they might read the visualized letters with imaginary pronunciation during the alphabet letter visualization task. The imaginary pronunciation sound duration and intensity might be involved in evoking EEG theta oscillations in the temporal cortex (Luo and Poeppel, 2007; Howard and Poeppel, 2010; Wang et al., 2012; Ng et al., 2013; ten Oever and Sack, 2015) and enhancing psychoacoustic sensitivity (Goswami et al., 2011). Additional experiments must be conducted to identify how much decoded information is purely derived from the visualization process and how much is from an imaginary spoken process. Different from the method

of Schyns et al. (2011), we trained an SVM to perform the classification. The merit of this approach is that it may have a potential BCI application, although the present method cannot distinguish how and to what extent the characteristics of the stimuli are encoded into the EEG oscillation phase patterns that might be limited by the spatial and temporal resolution of the EEG signals. Because SVM and other machine-learning methods are a type of black box, more detailed analytical methods and experimental designs must be used in future research to examine the potential value and limitations of this approach.

How low frequency oscillatory phases represent information in visual perception remains an open issue. In audio perception, the evidence indicates that theta oscillation is a mimic to the input speech envelope (Giraud and Poeppel, 2012; Gross et al., 2014). In this case, the peak (phase zero) of the oscillation may represent a high amplitude of speech envelope, and the trough (phase  $\pi$ ) is related to the quietness.

In addition, recent studies observed that different neuronal oscillations are not independent and isolated (Canolty et al., 2006). They can interact with each other to modulate oscillation amplitude and phase patterns, resulting in a cross-frequency coupling effect. The cross-frequency coupling may include several interactions, such as phase synchronization, amplitude co-modulation and phase-amplitude coupling (PAC). PAC is believed to reflect neural coding of signals within the local microscale and macroscale networks of the brain (Canolty and Knight, 2010). There is increasing experimental evidence suggesting that PAC may provide more useful information for decoding of object categories (Watrous et al., 2015; Jafakesh et al., 2016), which need to be deeply studied in future once high quality data of EEG or ECoG recording is available.

## CONCLUSION

Our experimental results provide strong evidences to confirm that the frequency, phase patterns and power information of cortical oscillation parameters contain important information about stimulus features. First, we found that decoding EEG phase patterns brings higher discrimination accuracy values than decoding EEG power portion. Second, frequency range and cortical spatial location are critical in decoding. We observed that phase patterns of the theta and alpha rhythms recorded in the occipital scalp visual and temporal regions contain more rich information that is valuable for decoding different input visual stimuli compared to other regions. EEG power sequences in the theta oscillation showed a significantly higher discrimination rate than did the chance level, although its classification performance was slightly lower than EEG phase pattern. Decoding the EEG phase and power sequence in the much lower frequency delta band or much higher beta and gamma frequency bands does not result in significant discrimination rates. Third, timing is important. Most of the valuable decoded information is within the first half-second period (100–600 ms) after the stimuli's presentation, and this information is hardly captured by the functional

magnetic resonance imaging technique (with a time resolution of approximately 1 s).

In sum, our experimental results support that low-frequency cortical oscillations are actively involved in coding sensory information. Directly decoding the phase and power sequences of EEG signals in the theta band may have great potential in brain-computer interface applications for English alphabet letter discrimination. Although the present EEG study showed that electrodes sited in the occipital scalp visual and temporal regions had higher accuracy rates and always reached the peak first, future research with combined EEG and functional MRI experiments may provide better spatial resolution in distinguishing the precise cortical locations in visual stimulus-encoding sites.

## AUTHOR CONTRIBUTIONS

YY, PW, and YW designed the research, YY and YW performed the research, and YW and YY wrote the paper. All authors reviewed the manuscript.

## REFERENCES

- Busch, N. A., Dubois, J., and Vanrullen, R. (2009). The phase of ongoing EEG oscillations predicts visual perception. *J. Neurosci.* 29, 7869–7876. doi: 10.1523/JNEUROSCI.0113-09.2009
- Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., et al. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313, 1626–1628. doi: 10.1126/science.1128115
- Canolty, R. T., and Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends Cogn. Sci.* 14, 506–515. doi: 10.1016/j.tics.2010.09.001
- Chang, C.-C., and Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–27. doi: 10.1145/1961189.1961199
- Colgin, L. L., Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., et al. (2009). Frequency of gamma oscillations routes flow of information in the *Hippocampus*. *Nature* 462, 353–357. doi: 10.1038/nature08573
- Dahaene, S. (1995). Electrophysiological evidence for category-specific word processing in the normal human brain. *Neuronreport* 6, 2153–2157. doi: 10.1097/00001756-199511000-00014
- Dehaene, S., and Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends Cogn. Sci. (Regul. Ed.)* 15, 254–262. doi: 10.1016/j.tics.2011.04.003
- 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
- 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
- Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034
- Fries, P., Nikolić, D., and Singer, W. (2007). The gamma cycle. *Trends Neurosci.* 30, 309–316. doi: 10.1016/j.tics.2007.05.005
- Giraud, A.-L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat. Neurosci.* 15, 511–517. doi: 10.1038/nn.3063
- Goswami, U., Wang, H.-L. S., Cruz, A., Fosker, T., Mead, N., and Huss, M. (2011). Language-universal sensory deficits in developmental dyslexia: English, Spanish, and Chinese. *J. Cogn. Neurosci.* 23, 325–337. doi: 10.1162/jocn.2010.21453
- Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., et al. (2014). Speech rhythms and multiplexed oscillatory sensory coding

## ACKNOWLEDGMENTS

We are grateful of Dr. Ying Mao and Liang Chen for their great helps in discussions and experimental protocol. YY thanks for the support from the National Natural Science Foundation of China (31571070, 81761128011), Shanghai Science and Technology Committee support (16410722600), the program for the Professor of Special Appointment (Eastern Scholar SHH1140004) at Shanghai Institutions of Higher Learning, the Research Fund for the Doctoral Program of Higher Education of China (1322051) and Omics-based precision medicine of epilepsy entrusted by the Key Research Project of the Ministry of Science and Technology of China (Grant No. 2016YFC0904400) for their support.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2018.00062/full#supplementary-material>

- in the human brain. *PLoS Biol.* 11:e1001752. doi: 10.1371/journal.pbio.1001752
- Heusser, A. C., Poeppel, D., Ezzyat, Y., and Davachi, L. (2016). Episodic sequence memory is supported by a theta-gamma phase code. *Nat. Neurosci.* 19, 1374–1380. doi: 10.1038/nn.4374
- Howard, M. F., and Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *J. Neurophysiol.* 104, 2500–2511. doi: 10.1152/jn.00251.2010
- Iemi, L., Chaumon, M., Crouzet, S. M., and Busch, N. A. (2017). Spontaneous neural oscillations bias perception by modulating baseline excitability. *J. Neurosci.* 37, 807–819. doi: 10.1523/JNEUROSCI.1432-16.2016
- Jafakesh, S., Jahromy, F. Z., and Daliri, M. R. (2016). Decoding of object categories from brain signals using cross frequency coupling methods. *Biomed. Signal Process. Control* 27, 60–67. doi: 10.1016/j.bspc.2016.01.013
- Kahana, M. J., Seelig, D., and Madsen, J. R. (2001). Theta returns. *Curr. Opin. Neurobiol.* 11, 739–744. doi: 10.1016/S0959-4388(01)00278-1
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Brain Res. Rev.* 29, 169–195. doi: 10.1016/S0165-0173(98)00056-3
- Krüger, N., Janssen, P., Kalkan, S., Lappe, M., Leonardis, A., Piater, J., et al. (2013). Deep hierarchies in the primate visual cortex: what can we learn for computer vision? *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 1847–1871. doi: 10.1109/TPAMI.2012.272
- Lange, J., Oostenveld, R., and Fries, P. (2013). Reduced occipital alpha power indexes enhanced excitability rather than improved visual perception. *J. Neurosci.* 33, 3212–3220. doi: 10.1523/JNEUROSCI.3755-12.2013
- Lisman, J. E., and Idiart, M. A. (1995). Storage of 7 +/- 2 short-term memories in oscillatory subcycles. *Science* 267, 1512–1515. doi: 10.1126/science.7878473
- Luo, H., and Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54, 1001–1010. doi: 10.1016/j.neuron.2007.06.004
- Matsuo, T., Kawasaki, K., Kawai, K., Majima, K., Masuda, H., Murakami, H., et al. (2015). Alternating zones selective to faces and written words in the human ventral occipitotemporal cortex. *Cerebral Cortex* 25, 1265–1277. doi: 10.1093/cercor/bht319
- Nezamfar, H., Salehi, S. S. M., Moghadamfalahi, M., and Erdogmus, D. (2016). FlashType<sup>TM</sup>: a context-aware c-VEP-based BCI typing interface using EEG signals. *IEEE J. Sel. Top. Signal Process.* 10, 932–941. doi: 10.1109/JSTSP.2016.2552140

- Ng, B. S., Logothetis, N. K., and Kayser, C. (2013). EEG Phase Patterns Reflect the Selectivity of Neural Firing. *Cerebral Cortex* 23, 389–398. doi: 10.1093/cercor/bhs031
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Panzeri, S., Macke, J. H., Gross, J., and Kayser, C. (2015). Neural population coding: combining insights from microscopic and mass signals. *Trends Cogn. Sci.* 19, 162–172. doi: 10.1016/j.tics.2015.01.002
- Panzeri, S., Safaai, H., De Feo, V., and Vato, A. (2016). Implications of the dependence of neuronal activity on neural network states for the design of brain-machine interfaces. *Front. Neurosci.* 10:165. doi: 10.3389/fnins.2016.00165
- Price, C. J., and Devlin, J. T. (2011). The Interactive Account of ventral occipitotemporal contributions to reading. *Trends Cogn. Sci.* 15, 246–253. doi: 10.1016/j.tics.2011.04.001
- Schyns, P. G., Thut, G., and Gross, J. (2011). Cracking the code of oscillatory activity. *PLoS Biol.* 9:e1001064. doi: 10.1371/journal.pbio.1001064
- Siegel, M., Warden, M. R., and Miller, E. K. (2009). Phase-dependent neuronal coding of objects in short-term memory. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21341–21346. doi: 10.1073/pnas.0908193106
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.* 19, 109–139. doi: 10.1146/annurev.ne.19.030196.000545
- ten Oever, S., and Sack, A. T. (2015). Oscillatory phase shapes syllable perception. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15833–15837. doi: 10.1073/pnas.1517519112
- Tomassini, A., Ambrogioni, L., Medendorp, W. P., and Maris, E. (2017). Theta oscillations locked to intended actions rhythmically modulate perception. *eLife* 6:e25618. doi: 10.7554/eLife.25618
- van Dijk, H., Schoffelen, J. M., Oostenveld, R., and Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J. Neurosci.* 28, 1816–1823. doi: 10.1523/JNEUROSCI.1853-07.2008
- van Gerven, M. A., Maris, E., Sperling, M., Sharan, A., Litt, B., Anderson, C., et al. (2013). Decoding the memorization of individual stimuli with direct human brain recordings. *Neuroimage* 70, 223–232. doi: 10.1016/j.neuroimage.2012.12.059
- Vanrullen, R., Busch, N. A., Drewes, J., and Dubois, J. (2011). Ongoing EEG phase as a Trial-by-Trial predictor of perceptual and attentional variability. *Front. Psychol.* 2:60. doi: 10.3389/fpsyg.2011.00060
- VanRullen, R., and Macdonald, J. S. P. (2012). Perceptual echoes at 10 hz in the human brain. *Curr. Biol.* 22, 995–999. doi: 10.1016/j.cub.2012.03.050
- Wang, R., Perreau-Guimaraes, M., Carvalhaes, C., and Suppes, P. (2012). Using phase to recognize English phonemes and their distinctive features in the brain. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20685–20690. doi: 10.1073/pnas.1217500109
- Watrous, A. J., Deuker, L., Fell, J., and Axmacher, N. (2015). Phase-amplitude coupling supports phase coding in human ECoG. *eLife* 4:e07886. doi: 10.7554/eLife.07886
- Won, D.-O., Zhang, H. H., Guan, C., and Lee, S.-W. (2014). “A BCI speller based on SSVEP using high frequency stimuli design,” in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (San Diego, CA: IEEE), 1068–1071. doi: 10.1109/SMC.2014.6974055
- Worden, M. S., Foxe, J. J., Wang, N., and Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific  $\alpha$ -band electroencephalography increases over occipital cortex. *J. Neurosci.* 20:RC63.
- Zhang, D., Song, H., Xu, R., Zhou, W., Ling, Z., and Hong, B. (2013). Toward a minimally invasive brain-computer interface using a single subdural channel: a visual speller study. *Neuroimage* 71, 30–41. doi: 10.1016/j.neuroimage.2012.12.069

**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 © 2018 Wang, Wang and Yu. 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) and the copyright owner 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.



# Temporal Dissociation of Neocortical and Hippocampal Contributions to Mental Time Travel Using Intracranial Recordings in Humans

Roey Schurr<sup>1,2</sup>, Mor Nitzan<sup>2,3†</sup>, Ruth Eliahou<sup>4</sup>, Laurent Spinelli<sup>5</sup>, Margitta Seeck<sup>5</sup>, Olaf Blanke<sup>5,6</sup> and Shahar Arzy<sup>1,2\*</sup>

<sup>1</sup> Neuropsychiatry Lab, Department of Neurology, Hadassah Hebrew University Medical Center, Jerusalem, Israel, <sup>2</sup> Faculty of Medicine, Hadassah Hebrew University Medical School, Jerusalem, Israel, <sup>3</sup> Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem, Israel, <sup>4</sup> Department of Radiology, Hadassah Hebrew University Medical Center, Jerusalem, Israel, <sup>5</sup> Department of Neurology, University Hospital, Geneva, Switzerland, <sup>6</sup> Laboratory of Cognitive Neuroscience, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

James M. Broadway,  
University of California, Santa Barbara,  
United States  
Isabel Maria Martin Monzon,  
Universidad de Sevilla, Spain

### \*Correspondence:

Shahar Arzy  
shahar.arzy@ekmd.huji.ac.il

### †Present Address:

Mor Nitzan,  
The Rachel and Selim Benin School of  
Computer Science and Engineering,  
The Hebrew University of Jerusalem,  
Jerusalem, Israel

**Received:** 09 January 2018

**Accepted:** 12 February 2018

**Published:** 28 February 2018

### Citation:

Schurr R, Nitzan M, Eliahou R, Spinelli L, Seeck M, Blanke O and Arzy S (2018) Temporal Dissociation of Neocortical and Hippocampal Contributions to Mental Time Travel Using Intracranial Recordings in Humans. *Front. Comput. Neurosci.* 12:11. doi: 10.3389/fncom.2018.00011

In mental time travel (MTT) one is “traveling” back-and-forth in time, remembering, and imagining events. Despite intensive research regarding memory processes in the hippocampus, it was only recently shown that the hippocampus plays an essential role in encoding the temporal order of events remembered, and therefore plays an important role in MTT. Does it also encode the temporal relations of these events to the remembering self? We asked patients undergoing pre-surgical evaluation with depth electrodes penetrating the temporal lobes bilaterally toward the hippocampus to project themselves in time to a past, future, or present time-point, and then make judgments regarding various events. Classification analysis of intracranial evoked potentials revealed clear temporal dissociation in the left hemisphere between lateral-temporal electrodes, activated at ~100–300 ms, and hippocampal electrodes, activated at ~400–600 ms. This dissociation may suggest a division of labor in the temporal lobe during self-projection in time, hinting toward the different roles of the lateral-temporal cortex and the hippocampus in MTT and the temporal organization of the related events with respect to the experiencing self.

**Keywords:** episodic memory, mental time travel, self-projection, self-reference, hippocampus, lateral temporal, sEEG

## INTRODUCTION

A fundamental trait of human cognition is the capacity to engage in “mental time travel” (MTT), to remember past events or imagine possible future ones (Tulving, 1985). When Tulving first presented the concept of MTT, it was proposed as a means of extending and binding together the two more basic functions of episodic memory and episodic future thinking, also known as “prospection” (Schacter and Addis, 2007; Suddendorf and Corballis, 2007; Bar, 2009; Spreng et al., 2009; Schacter et al., 2012). Over the years, the concept of MTT was developed beyond the common neurocognitive basis of past and future thinking to include several different functions (Spreng et al., 2009; Schacter et al., 2012). The process of “scene construction” has been suggested as a key component of MTT, allowing the retrieval of relevant



elements from memory and their subsequent binding into a coherent spatial scene (Hassabis et al., 2007a; Maguire and Mullally, 2013). Another process suggested as a fundamental aspect of MTT is self-projection in time, namely the ability to disengage from the immediate environment and mentally “project” oneself to a new “self-location” in time, either in the past or in the future (Buckner and Carroll, 2007; Arzy et al., 2008; Nyberg et al., 2010; Markowitsch and Staniloiu, 2011; Klein, 2013; Kurczek et al., 2015). It is from this “self-location” in time that the individual re-orient herself with respect to different events, in past or future (Arzy et al., 2009a; Peer et al., 2015). To reiterate, MTT comprises of several distinct processes, among them: self-projection to a specific self-location in time, imagination of the relevant event (that is, the act of remembering a past event or of prospecting a future one), and self-orientation with respect to other events (Peer et al., 2014, 2015).

Similarly to the way in which the field of memory research has progressed from focusing on autobiographical memory to the broader notion of MTT and related concepts, the study of their neuroanatomical substrate has also advanced. Whereas, early studies of memory functions focused on the hippocampus, various studies have since established the existence of a large-scale brain network supporting MTT-related processes (Buckner and Carroll, 2007; Hassabis et al., 2007a; Arzy et al., 2009a; Schacter and Addis, 2009; Spreng et al., 2009; Nyberg et al., 2010; Benoit and Schacter, 2015). The key regions of this network include the medial prefrontal, posterior parietal, and lateral temporal cortices, and the medial temporal lobe, including the hippocampus (Addis et al., 2007; Arzy et al., 2009a; Spreng et al., 2009; Rugg and Vilberg, 2013). Notably, although the hippocampus is considered a key region in this “core” network (McNaughton and Morris, 1987; Squire, 1992, 2004; Carpenter and Grossberg, 1993; Moll and Miikkulainen, 1997; Scoville and Milner, 2000; Yonelinas, 2002; Burgess et al., 2007; Bird and Burgess, 2008), its specific involvement in MTT is still debated. For example, while some reported hippocampal involvement in future thinking (Okuda et al., 2003; Hassabis et al., 2007b; Schacter and Addis, 2009), others reported evidence suggesting that future thinking could be independent of the hippocampus (Squire et al., 2010; Hurley et al., 2011).

Moreover, elucidating the differential contributions of the hippocampus and neocortical regions to MTT may have profound implications for the ongoing debate regarding the role of the hippocampus in both memory functions and spatial cognition, including representation of the immediate space, navigation and spatial orientation (O’Keefe and Dostrovsky, 1971; Doeller et al., 2008; Dombeck et al., 2010; Buzsáki and Moser, 2013; Eichenbaum and Cohen, 2014; Hartley et al., 2014). Several attempts have been made to reconcile the role of the hippocampus in memory functions and spatial cognition. The “relational memory theory” suggests that the hippocampus offers a general relational processing mechanism, providing similar computations for the encoding of episodes as sequences of events, and the encoding of routes as sequences of places traversed (Konkel and Cohen, 2009; Eichenbaum and Cohen, 2014). Alternatively, the abovementioned “scene construction theory” asserts that the hippocampus supports episodic memories and

imagined future events by facilitating the generation of atemporal scenes, binding together the event’s disparate elements into a coherent whole (Maguire and Mullally, 2013). Under this view, the hippocampus is thought to support spatial navigation by virtue of ongoing anticipatory scene construction, giving rise to a continuous representation of the upcoming spatial environment. While different empirical results support both theories, decisive experimental evidence for the role of the hippocampus in MTT is still required.

To investigate the role of the hippocampus in MTT we recorded intracranial evoked potentials (iEPs) in response to an established task of self-projection in time (Arzy et al., 2008, 2009a; **Figure 1**) in three patients with epilepsy undergoing pre-surgical evaluation. Patients were requested to imagine themselves either in the present self-location in time (“now”) or in another self-location, either 10 years toward the past or toward the future (“then”). It is from this self-location in time that they had to make judgments with respect to different events. For control purposes, iEPs were recorded also when patients performed a spatial task requiring self-projection in space (Arzy et al., 2006). Patients were implanted with bitemporal depth electrodes, penetrating both the hippocampus and the lateral temporal cortex (LTC), a major region in the cortical network involved in MTT (Svoboda et al., 2006; Arzy et al., 2008; Spreng et al., 2009; Benoit and Schacter, 2015; Peer et al., 2015). Such stereo-electroencephalography (sEEG) depth electrodes enable the separation of neocortical and hippocampal activities in both the time and space domains, unlike other neuroimaging methods, with lower spatial or temporal resolution (such as EEG and functional MRI, respectively). This setting enabled us to classify the temporal dynamics of brain activity in the hippocampus and LTC, to better understand the role of these regions in MTT.

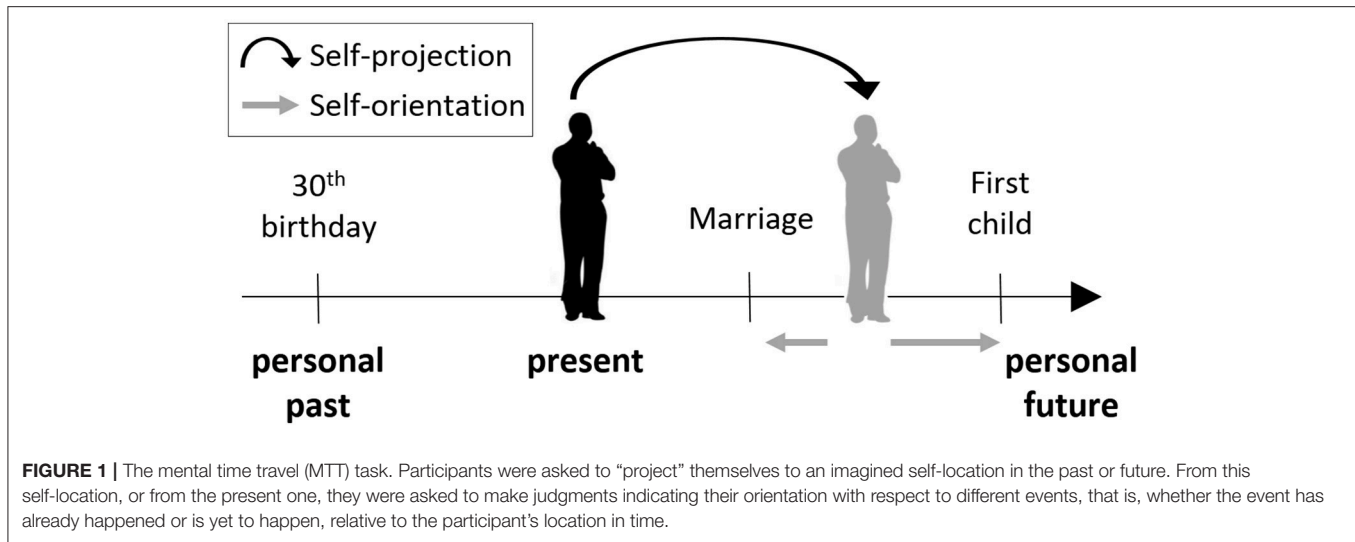
## MATERIALS AND METHODS

### Participants

Participants were three right-handed epileptic patients (17, 18, and 40 years old) who suffered from complex partial seizures resistant to pharmacological treatment, with no history of psychiatric or other neurological disorders. In order to localize the seizure onset zone and to dissociate it from essential cortex, intracranial electrodes were implanted. One patient was diagnosed with an epileptic focus in the right temporal pole, one with a left frontal focus, and in one the epileptic focus was found in the left amygdala. Written informed consent was obtained, and the procedures were approved by the Ethical Committee of the University Hospital of Geneva.

### Stimuli and Procedures

In the MTT task (Arzy et al., 2008) participants are first asked to imagine themselves either at the present time (“now”), or in another time point (“then”), 10 years in the past or in the future. Participants are then presented with events from personal life (e.g., car license; first child) or non-personal world events (e.g., Challenger explosion; Obama’s election), and are asked to indicate whether this event takes place before or after the currently imagined time-point (**Figure 1**). Thus, participants are



requested to mentally “project” themselves in time in order to accomplish the task. Stimuli were designed to be in the range of  $\pm 10$  years of the imagined time-point, and included events that were chosen from a validated list of common personal life events for the personal items, and from major headline news events for the non-personal items (Arzy et al., 2008, 2009a). Stimuli appeared for 700 ms in the center of a computer screen with an inter-stimulus interval of 2,000 ms as used previously (Arzy et al., 2008). Judgments were given using index and middle fingers of the left and right hand in alternating blocks as a button press on a serial response box. Participants were instructed to respond as quickly and precisely as possible while maintaining a mental image of themselves in the appropriate time-point (“now,” “past,” or “future”). These conditions were performed in different blocks and counterbalanced across participants. Each block included 120 stimuli, equally distributed among four groups appearing in random order: personal-events/world-events  $\times$  before/after.

As a control task, participants also performed a spatial task involving own-body transformation (Blanke et al., 2005). This task presents participants with a schematic human figure, either facing toward them or away from them, with the figure’s right or left hand marked by a ribbon. Participants either responded from their present location (“here”), or were asked to mentally “project” themselves to the location represented by the schematic figure (“there”). It is from this perspective that they made judgments regarding the presented figure (Figure S1; Blanke et al., 2005; Arzy et al., 2006). In the “there” condition, participants were instructed to indicate whether the figure’s marked hand is the right or left hand. They were instructed to respond as fast and precise as possible, yet always perform the mental projection of their body before responding. In the “here” condition the same visual stimuli were used, and participants were asked to decide from their habitual location whether the indicated hand was on the right or the left side of the computer screen (Blanke et al., 2005). Stimuli appeared for 300 ms in the center of the computer screen. The interstimulus interval was 2,000 ms. Each block included 120 stimuli, equally distributed

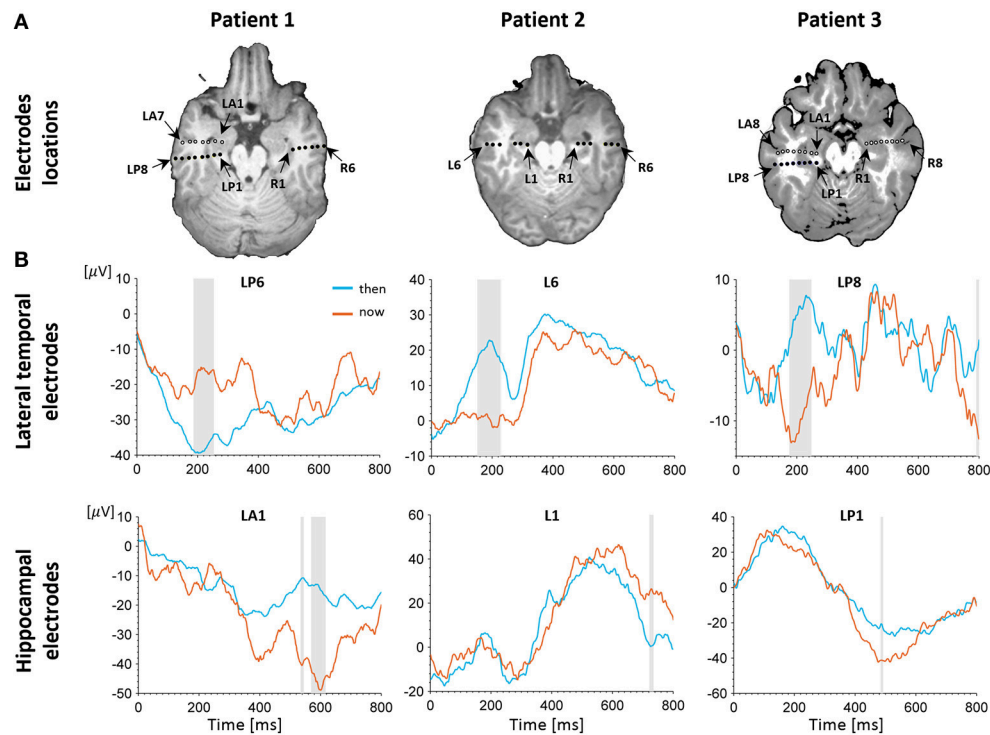
among the four conditions, counterbalanced across subjects. Since the analysis is done within-task, an optimal duration for stimulus presentation was chosen separately for each task, based on previous studies.

## Overview of Implanted Electrodes

Patients were implanted with depth electrodes penetrating the temporal lobe from the neocortex to the MTL bilaterally according to strict clinical criteria. In total, we have analyzed 57 electrodes implanted in all three patients (Figure 2A).

## EEG Acquisition and Analysis

Continuous intracranial EEG was acquired with a Deltamed<sup>®</sup> system [1,024 Hz (patients 1,2) or 512 Hz (patient 3) digitization]. Depth electrodes had a center-to-center distance of 1 cm (Ad-Tech, Racine, WI). Electrode location was determined by three dimensional MRI of the brain as well as CT scan with the implanted electrodes (Blanke et al., 1999, 2005). Preprocessing and analyses were conducted using Cartool software (Brunet et al., 2011; <https://sites.google.com/site/cartoolcommunity/>), Brainstorm toolbox (Tadel et al., 2011; <http://neuroimage.usc.edu/brainstorm>), FieldTrip toolbox (Oostenveld et al., 2011; <http://www.ru.nl/neuroimaging/fieldtrip>), and Matlab<sup>®</sup> (Mathworks, inc.). Epochs of EEG from 100 ms before to 800 ms after stimulus onset were bandpass filtered (1–120 Hz), and averaged for each of the stimulus conditions to calculate the intracranial evoked potential (iEPs). In the MTT task, the past and future conditions were collapsed into one condition (“then”), allowing a simpler  $2 \times 2$  design (now/then  $\times$  before/after) in accordance with previous studies showing similar response to past and future events (e.g., Arzy et al., 2008, 2009a; Anelli et al., 2016; Gauthier and van Wassenhove, 2016; for review see Schacter et al., 2012). Data were inspected visually to reject epochs with epileptic discharges as well as epochs with other types of transient noise.



**FIGURE 2 |** Electrophysiological results. **(A)** Depth electrodes locations in the hippocampus and lateral temporal cortex (LTC), shown for each patient on a co-registration of post-operative CT scan and pre-operative MRI (white circles depict electrodes projected on this slice for visualization purposes; for more precise localization of these electrodes see **Figure S2**). Exact neuroanatomical position of each electrode as verified by two certified neuro-radiologists is available in **Table S1**. **(B)** Intracranial evoked potentials (IEPs) recorded at representative electrodes in the left LTC (top) and left hippocampus (bottom) during MTT. LTC electrodes show high early task modulation, whereas electrodes in the hippocampus show high late task modulation. Shaded areas show time points of significant differences between conditions in a two-tailed independent samples  $t$ -test ( $p < 0.05$ , uncorrected).

## Electrode Selection

We aimed to differentiate between lateral cortical and hippocampal activations in response to the MTT and the spatial tasks. To this end, we identified hippocampal and LTC electrodes according to their apparent location on a post-implantation CT, co-registered with the pre-implantation MRI images. Exact neuroanatomical position of each electrode was verified by two certified neuro-radiologists using a neuroanatomical atlas (Harnsberger et al., 2006). Electrodes that showed clearly defective iEPs were excluded from the analyses.

## Electrodes Classification

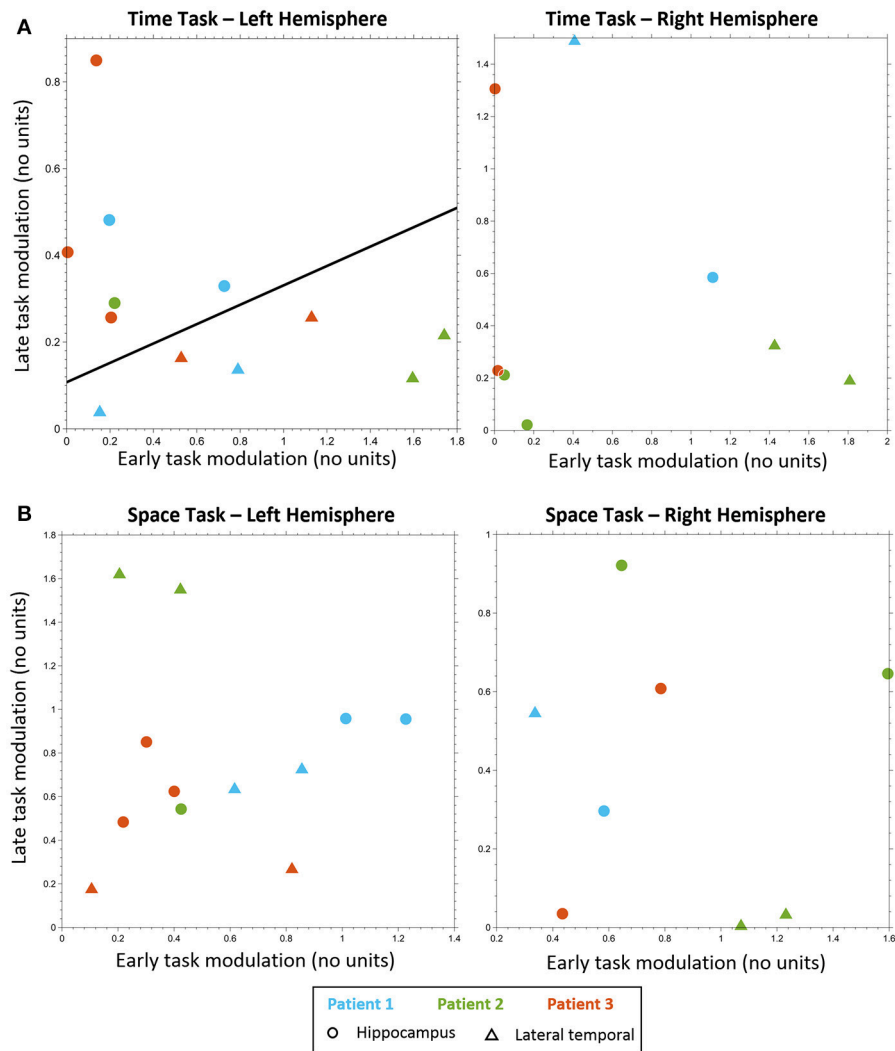
Following our previous findings using EEG (Arzy et al., 2008), we defined two time periods of interest: an early period ranging from 100 to 400 ms post stimulus onset that encompassed the initial peak responses at the LTC, and a late period ranging from 400 to 800 ms post stimulus onset that captured a second peak response in the hippocampus (**Figure 2B**; Staresina et al., 2012). To differentiate between LTC and hippocampal electrodes we defined early and late modulation features for each electrode and task, as follows (**Figure 4**): For each condition and period, the raw modulation was defined as the absolute value of the sum of differences between iEPs deflections in the two conditions (the signed area between the two iEPs deflections). Subsequently, the

modulation was normalized by the area under the curve of the “now” (or “here”) condition in the same period. Accordingly, the early modulation of electrode  $i$  in the time-task is given by:

$$\text{Early modulation} = \frac{\left| \int_{100 \text{ ms}}^{400 \text{ ms}} S_{\text{then}}^i(t) - S_{\text{now}}^i(t) dt \right|}{\left| \int_{100 \text{ ms}}^{400 \text{ ms}} S_{\text{now}}^i(t) dt \right|} \quad (1)$$

Where  $S_{\text{now}}^i(t)$  and  $S_{\text{then}}^i(t)$  are the mean iEPs recorded in electrode  $i$  in the “now” and “then” conditions, respectively. Likewise, the late modulation is defined with integration limits of 400–800 ms.

Each electrode’s position in the two dimensional feature space was thus determined by its early and late task modulations (**Figure 4D**). When lateral temporal and hippocampal electrodes seemed separable in this representation, we tested for significance of this separation using Support Vector Machine with a linear kernel (SVM; Cortes and Vapnik, 1995). Linear SVM is a supervised learning algorithm that performs linear classification of the data by constructing the optimal hyperplane with largest margin for separating data into two groups. To avoid domination of small numeric results by greater ones we scaled the data by



**FIGURE 3 |** Electrodes classification. Electrodes classification using linear SVM, based on early task modulation value (X-axis) and late task modulation value (Y-axis). **(A)** MTT task: left hippocampal electrodes (circles) are clearly separable from left lateral temporal cortex (LTC) electrodes (triangles) on the plane of early and late task modulations (see **Figure S3** and **Table S2**). A separating line is shown, as obtained from SVM classification of all electrodes (left,  $p < 0.005$ ). No such separation was found for electrodes in the right hemisphere (right, see **Figure S4**). **(B)** Spatial task: no separation between lateral temporal and hippocampal electrodes was found neither in the left hemisphere (left, see **Figure S6**) nor in the right (right, see **Figure S7**).

Z-score procedure for each of the two features (Chang and Lin, 2011).

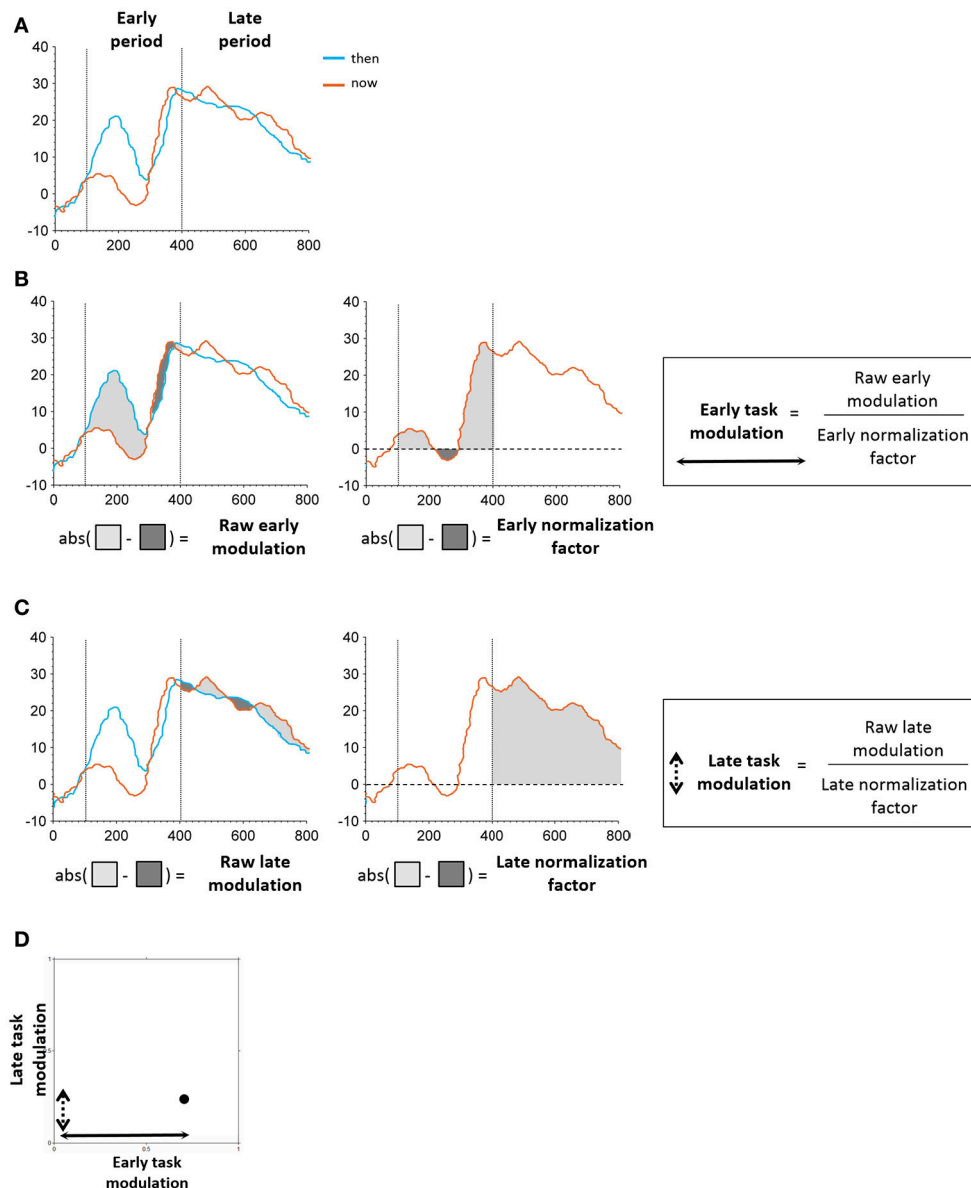
SVM uses a penalty parameter  $C > 0$  that determines the tradeoff between margin maximization and training error minimization. An optimal value for this parameter had to be determined. Ten different  $C$ -values equally spaced on a log-scale in the range of  $[10^{-3}, 10^3]$  were tested, each yielding a cross-validation classification accuracy using the N-fold cross-validation procedure (Chang and Lin, 2011). The  $C$ -value yielding the highest cross-validation accuracy was subsequently used for training the classifier and for statistical tests.

To statistically validate our classification results, we used a non-parametric permutation test (Ojala and Garriga, 2010). The null hypothesis of this test is that the dataset labels (LTC or hippocampal) are independent of the features (early and

late modulations). We re-trained the classifier on all possible permutations of the dataset labels, and calculated the N-fold cross-validation accuracy for each permutation. This allowed the derived classification accuracy to be assigned a  $p$ -value. In case the dataset labels and features are independent in the original data, one can expect to obtain high  $p$ -values (Ojala and Garriga, 2010).

### iEP-Amplitude Analysis

We examined whether iEPs significantly differed between conditions (“now”/“then” and “here”/“there”). To this aim, statistical analysis ( $t$ -tests, two tailed,  $p < 0.05$ , uncorrected) was used on the amplitude of the single unaveraged epochs over trials, comparing the different experimental conditions in each time-frame, and searching for significant differences. Since



**FIGURE 4 |** Schematic illustration of task modulation extraction. **(A)** Early and late periods identified in the time windows of 100–400 and 400–800 ms post stimulus onset, respectively. **(B)** Extraction of early modulation value. The raw modulation was defined as the absolute value of the sum of differences between iEPs in the two conditions (“then”–“now,” “there”–“here”; left). The normalization factor was defined as the area under the curve of the “now”/“here” condition in the respective period (middle). The raw modulation was subsequently normalized by the normalization factor of the respective period, resulting in the final task modulation value (right). **(C)** Extraction of early modulation value. Same procedure as applied for the early task modulation was used here. **(D)** Each electrode’s position in the two-dimensional feature space was determined by its early and late task modulation values.

iEP values at adjacent time-frames are highly dependent, one cannot use conventional methods of correction for the multiple comparisons. We therefore used a cluster-based nonparametric randomization test (Maris and Oostenveld, 2007). In short, clusters were defined as continuous time-frames in which the t-statistic exceeded a given threshold (corresponding to  $p < 0.05$ ). A cluster-level test statistic was defined as the sum of all t-statistics in the cluster, and the type-I error rate was controlled by evaluating the cluster-level test statistic under the

randomization null distribution of the maximum cluster-level test statistic, using 1,000 random permutations between the two conditions and  $p < 0.05$ .

## RESULTS

A behavioral self-projection effect was found in two out of the three patients, with longer reaction times for the “past” and



“future” conditions compared with the “now” condition ( $p < 0.05$  for all tests), comparable to previous studies using the same paradigm in larger number of subjects (e.g., Arzy et al., 2008, 2009a). To distinguish between LTC and hippocampal involvement we used data from all patients and analyzed 12 electrodes in the left hemisphere (six in the LTC and six in the hippocampus) and eight electrodes in the right hemisphere (three in the LTC and five in the hippocampus; **Figure 2A**, **Figure S2**). Analysis of iEPs in the left hemisphere in the MTT task showed a significant early task modulation in the time window of  $\sim 100$ – $300$  ms ( $p < 0.05$  uncorrected) in five out of six LTC electrodes (**Figure 2B**, upper row; **Figure S3**). A late task modulation was found in the time window of  $\sim 400$ – $600$  ms in all hippocampal electrodes (**Figure 2B**, lower row; **Figure S3**). Such consistent effects were not found in the right hemisphere (**Figure S4**), nor in the spatial task in either hemisphere (**Figures S6, S7**).

Classification analysis based on the early and late task modulations (**Figure 3**) yielded a significant separation between LTC and hippocampal electrodes in the MTT task in the left hemisphere (cross-validation accuracy 100%,  $p = 0.004$ ; **Figure 3A**). Five out of six electrodes which showed late hippocampal modulation were located in the hippocampal formation (HF) and one in the parahippocampal gyrus. No significant separation was found in the right hemisphere (cross-validation accuracy 75%,  $p = 0.304$ ; **Figure 3B**), nor in the spatial task either for the left or right hemispheres (cross-validation accuracy 33.33, 62.5%;  $p = 0.847, 0.982$ , respectively; **Figures 3C,D**). No significant difference between conditions was found in the MTT task nor in the spatial task using the cluster-based nonparametric randomization test.

## DISCUSSION

The present study used the high temporal and spatial resolution of intracranial recordings and employed a classification analysis in order to distinguish between LTC and hippocampal involvement in self-projection in time, a key component in MTT. Our iEP data revealed that LTC and hippocampal contributions to self-projection in time display distinct temporal dynamics. Classification analysis of electrodes in the left hemisphere showed a clear temporal dissociation between LTC electrodes that exhibited an early self-projection component ( $\sim 100$ – $300$  ms), and hippocampal electrodes that exhibited a late component ( $\sim 400$ – $600$  ms). No such effect was found either in the right hemisphere or in a control task of self-projection in space.

Our results suggest the involvement of both LTC and the hippocampus in MTT. Several neuroimaging studies involving MTT-related tasks revealed increased activation in both the medial temporal lobe and the LTC (Addis et al., 2007, 2009a, 2011; Buckner and Carroll, 2007; Schacter and Addis, 2007; Botzung et al., 2008; Arzy et al., 2009a; Spreng et al., 2009; Spreng and Grady, 2010; Schacter et al., 2012; Benoit and Schacter, 2015). The high spatial and temporal resolution of iEPs enabled us to temporally dissociate the contributions of these two regions during MTT. We believe these results could not be explained by mere temporal delay in the processing of the same

information at the circuit level, since other sEEG studies have identified hippocampal responses within the first few 100 ms of stimulus/task onset (Axmacher et al., 2007, 2010; Olsen et al., 2012), while here hippocampal activity was found significantly later ( $\sim 400$ – $600$  ms). Therefore, these results suggest a division of labor in the temporal lobe: Early processing of self-projection takes place in the LTC, to establish one's self-location on the mental time line (the first step in the MTT task). Subsequently, hippocampal activity possibly reflects the required computations for orienting oneself with respect to the presented events (the second step in the MTT task). These results are in line with patient data revealing preservation of self-projection effects despite hippocampal lesions (Arzy et al., 2009b). This latter implication of the hippocampus in MTT may be related to its role in determining the temporal order of events, in accordance with the “relational memory theory” (Eichenbaum and Cohen, 2014). According to this theory, the hippocampus serves as a general relational processing mechanism, involving, among other representational schemes, the representation of episodes as the flow of events across time. The hippocampus may be similarly involved in the task used here, in determining the temporal relations of the events to one's imagined self-location in time. This is also in line with previous clinical and neuroimaging studies that found hippocampal activity in tasks involving general relational processing (Giovanello et al., 2004; Preston et al., 2004; Prince et al., 2005; Konishi et al., 2006), and specifically in the context of the temporal order of events (Reber and Squire, 1998; Hopkins et al., 2004; Lehn et al., 2009; Paz et al., 2010; Davachi and DuBrow, 2015; Rubin et al., 2015; Jenkins and Ranganath, 2016). Impaired ability to explicitly remember the sequential order of events was also found in studies in amnesic patients with hippocampal damage (Reber and Squire, 1998; Hopkins et al., 2004) as well as lesion studies in nonhuman animals (DeCoteau and Kesner, 2000; Fortin et al., 2002; Kesner et al., 2002).

Most hippocampal electrodes that showed late hippocampal modulation were located in the hippocampal formation (HF). The HF has been shown to be involved in MTT and autonoetic consciousness in a unique model of patient population with a specific lesion in the CA1 part of the HF (Bartsch et al., 2011). In a more precise manner, the HF also contains the recently discovered time-cells. Accumulating experimental evidence, mostly in rodents but also in humans, suggest that the hippocampus plays a central role in the temporal organization of memories (Devito and Eichenbaum, 2011; for review see Eichenbaum, 2013). Notably, these cells share similar properties with place-cells, which encode one's location in the environment (Kraus et al., 2015). Likewise, a time-space similarity was recently found in the distributed manner in which episodic or atemporal spatial memories are represented along the hippocampal axis, based on their temporal or spatial scale (Collin et al., 2015). However, such a similarity between the hippocampal responses to the MTT and spatial tasks was not evident in our results. A potential reason for that is that the spatial task here is not equivalent to the MTT task. Future studies may better address this point by designing more comparable temporal and spatial tasks (e.g., Gauthier and van Wassenhove, 2016). Another possibility is that higher-order functions as examined here are

**BOX 1 |** The effect of reducing the number of electrodes used in the classification analysis.

In our study we found significant separation of the LTC and the hippocampus based on their temporal pattern of activity only in the left hemisphere during the time task. Although these results seem to support left lateralization, the lack of clear separation in the right hemisphere should be interpreted with caution. Due to the small number of electrodes that met inclusion criteria in the right hemisphere (8 overall, where no LTC electrodes were included for subject 3, compared with 12 overall in the left hemisphere), classification in this hemisphere is of limited value. In other words, it is possible that the power of the statistical method used in this study is too low to reveal an effect in the right hemisphere, even if it exists. In principle, one could estimate the number of electrodes required to obtain a certain power level of the test, yet general procedures for planning sample size are yet to be developed in the case of classification based tests (Maxwell et al., 2008).

To assess the effect of the small number of electrodes in the right hemisphere, we conducted an additional analysis in which the number of electrodes in the left hemisphere was reduced to match that of the right hemisphere. The same classification analysis was done for all 120 possible subsets of electrodes in the left hemisphere which include exactly 5 hippocampal electrodes and 3 lateral temporal electrodes, as in the right hemisphere. For each subset we calculated the cross-validation accuracy and its *p*-value (see Materials and Methods). **Figure S8** shows the distribution of resulting accuracy values and their corresponding *p*-values. Although high accuracy values (>75%) were found in a large number of electrodes subsets (84/120), these findings were significant ( $p < 0.05$ ) for only a small fraction of the subsets (33/120). These results suggest that the lack of significant temporal separation in the right hemisphere could be the result of reduced power of the statistical analysis due to the small number of electrodes in this hemisphere.

not directly related to time- and place-cells, which could be responsible for encoding much shorter distances and time-scales.

Previous studies established the LTC as part of the MTT network, supporting both episodic memory and episodic future thinking (Svoboda et al., 2006; Hassabis et al., 2007a; Addis et al., 2009a; Spreng et al., 2009; Markowitsch and Staniloiu, 2011; Benoit and Schacter, 2015). Nevertheless, its exact role in the different processes comprising MTT is not completely clear. Much evidence has accumulated relating LTC activity to retrieval of semantic memory, by means of neuroimaging studies of various memory tasks in healthy subjects (Martin and Chao, 2001; McClelland and Rogers, 2003; Konishi et al., 2006), as well as studies in patients who suffered damage to the LTC (Hodges et al., 1992; Gilboa et al., 2005; Addis et al., 2009b). Retrieval of semantic knowledge has been suggested to subserve both recollection and future thinking, and thus support MTT (Tulving, 2002; Levine, 2004; Schacter et al., 2012). Recruitment of LTC was found in tasks involving decision making with respect to personal events (Andrews-Hanna et al., 2010), self-projection in time (St Jacques et al., 2011), construction and elaboration of past and future events (Addis et al., 2007), and orientation with respect to different events in time (Peer et al., 2015). The early iEP modulation we found in LTC further established the notion that the LTC supports MTT not only via retrieval of semantic information, but also through direct involvement in the act of self-projection in time.

Significant separation of the LTC and the hippocampus based on their temporal pattern of activity was found in our study only in the left hemisphere. Lateralization in the hippocampi has been known for a long time, but less so is the lateralization in the LTC. Our results are concordant with previous studies that found predominant left lateralization in various tasks involving autobiographic memory and orientation in time (Maguire, 2001; Levine, 2004; Svoboda et al., 2006; Arzy et al., 2008; Spreng et al., 2009; Peer et al., 2015), though some other studies have suggested right predominance (Fink et al., 1996; Gilboa et al., 2005; Arzy et al., 2009a). It should be noted that while our results suggest left lateralization, the lack of effect in the right hemisphere should be interpreted with caution. Due to the small number of electrodes that met inclusion criteria in the right hemisphere (8 overall, where no LTC electrodes were

included for subject 3), classification in this hemisphere is of limited value. In an additional analysis in which the number of electrodes in the left hemisphere was reduced to match that of the right hemisphere, the power of the test was indeed reduced, as expected (see **Box 1** and **Figure S8**). This is indeed a main limitation of this study, which includes a relatively small number of patients. However, this sample size is comparable to several other studies that include intracranial recording in human hippocampus (Vanni-Mercier et al., 2009; Staresina et al., 2012; Kurczek et al., 2015). Such small samples are customary due to the rare opportunity to record intracranial artifact-free high-quality electrophysiological data in response to high-cognitive tasks such as MTT and self-projection, which is not applicable even in primates. Notably, most patients with temporal electrodes suffer from hippocampal sclerosis and frequent electrical discharges, which contaminate the data. Such patients were not included in our study, making the study sample of high quality, though small. Moreover, our results were consistent across all subjects. Subjects were nevertheless epileptic patients in whom interictal epileptic activity may influence results. To avoid such a disturbance we applied several methods: First, in two of our patients epileptic foci were identified elsewhere and in one aberrant epileptic activity was absent during recording as well as 2 days later. The data was also inspected visually to exclude any epileptic artifacts. Stimulus-locked iEPs were clear and similar among patients. Most late modulations were found in the HF. However, more electrodes in other hippocampal locations may show responses as well. This was nevertheless impossible to test in our study, due to strict clinical considerations regarding electrodes implantation. It should thus be noted that the HF effect found here does not exclude a parallel parahippocampal effect.

As noted earlier, the spatial task is not equivalent to the time task. However, in both tasks patients had to imagine themselves in a different self-location—in time or in space. The absence of a significant early component for space in the LTC is also supported by fMRI and EEG studies using the same space task, which did not show such an activation (Arzy et al., 2006; Ionta et al., 2013). The late hippocampal modulation which relates stimuli to the projected self may be absent due to the nature of the spatial task used. Further study of a comparable spatial task involving relational organization of self and landmarks in space

**BOX 2 | Statistical learning and classification in the analysis of intracranial data.**

Intracranial electrophysiological recording in awake human patients is the most accurate existing method in the cognitive neurosciences. Unlike non-invasive methods—such as functional MRI, MEG or EEG—it enables direct recording of neural activity in exceptionally high spatial and temporal resolutions, as well as a high signal to noise ratio (SNR; Lachaux et al., 2003; Ball et al., 2009). It is therefore the only manner by which electrophysiological correlates of high cognitive functions may be recorded invasively, since such functions cannot be controlled in non-human animals, including primates. However, statistical group analysis—a common approach in the abovementioned modalities—is difficult to employ in iEPs. This is due to the strict clinical considerations regarding location of electrodes implantation and experimental settings, which ultimately lead to significant variability among individual patients. Therefore, whereas other neuroimaging methods are used to identify group effects across many subjects, in iEPs experiments, where only a handful of patients are usually recruited, analysis is effectuated in the individual subject level (Kramer et al., 2011; Peer et al., 2015). While the high quality of the data could enable the detection of significant effects on the level of individual subjects, it is not free of limitations. Statistics is done over trials, which do not necessarily reflect the cognitive effect; the number of repetitions affects both subjects' performance and statistical power; correction for multiple comparisons is dependent on the number of electrodes, which, in turn, are inserted according to clinical considerations and differ between patients. Needless to mention, even classical group effects are prone to invalid statistical inferences due to low statistical power, improper circular analysis, or other biases that tend to increase false-positive rates (Kriegeskorte et al., 2009; Simmons et al., 2011; Button et al., 2013).

A statistical method that may overcome these caveats, and therefore is appropriate for the analysis of iEP data, is *statistical learning*, and specifically *classification* (Arzy et al., 2014; Shalev-Shwartz and Ben-David, 2014). Here we use a distribution-free framework, aiming to identify a classification rule by which a new observation can be classified as belonging to one class or another. The classification process and resulting predictions are based on a set of features inherent to the data (e.g., in iEPs features may be comprised of amplitude, latency or power spectra, or as in our case: late and early task modulations). Each observation, or *instance*, is represented as a “vector of features” in the features space. Instances are further *labeled* as belonging to one of two or more predefined *classes* (e.g., in iEPs classes may consist of anatomical electrode location such as hippocampal vs. LTC, different frequency bands, or experimental conditions). In the framework of *supervised learning*, a finite set of labeled instances is defined as the *training data*. Subsequently, the procedure produces a *predictor*, or *classifier*, which can be used to predict the label of new instances, by separating the instances to different classes according to a certain *classification rule* (e.g., distance to its nearest neighbors or linear separation). The *accuracy* of a classifier is the probability that it will predict the correct label on a randomly generated set of instances and can be estimated on a given instance set using the N-fold cross-validation procedure (also termed “leave-one-out cross-validation”; Chang and Lin, 2011). In this procedure, classification is learned using N-1 instances, and then used to predict the label of the remaining instance. The process is repeated N times, and the fraction of instances classified correctly is used as the estimated classifications accuracy. In addition, one may estimate the statistical significance of classification accuracy by using methods such as non-parametric permutation tests on the dataset labels. Overall, such a statistical learning approach may therefore fit well iEPs analysis, as long as the research question may be reformulated as a classification problem into two (or several) predefined *classes*.

could shed light on the role of the hippocampus in non-temporal relational organization (Gauthier and van Wassenhove, 2016). We therefore refer in this study mostly to results found in the MTT task and mention spatial task results with caution.

Our small number of patients did not allow for reliable statistical testing using conventional approaches. Specifically in intracranial studies, it is difficult to delineate consistent iEPs across individuals, in part due to varying relative positions of the electrodes across different subjects. For example, such variability leads to “polarity reversal” (Halgren et al., 1982): When recording iEPs from local generators, the polarity of the resulting iEP reverses as one records from two opposite sides of this generator (Figure S5). We therefore suggest that classification, done at a low dimensional feature space that summarizes the iEPs recorded at each electrode, is a more suitable statistical method in such cases, and may serve as a useful tool in analyses of other neuroscientific data as well (Box 2; see also Arzy et al., 2014). While classification reliably distinguishes between predefined classes, the applied predefinition inevitably influences the results. Classification here was nevertheless based on previous results using fMRI and EEG, enabling a precise predefinition of classes with respect to neuroanatomical localization and appropriate time windows, respectively.

To conclude, in the present study we found that both the LTC and the hippocampus are involved in MTT; however, while the first is involved early in the process, as subjects “project” themselves in time, the latter is only involved later, when subjects relate the different events to the “projected” self. This division of labor may contribute to the reconciliation of the major debate regarding the role of the hippocampus in MTT.

## AUTHOR CONTRIBUTIONS

SA and OB: Designed the study; SA, LS, and MS: Performed the study; RS, MN, and SA: Analyzed the data; RE: Analyzed the neuroanatomical structures; RS and SA: Wrote the manuscript.

## FUNDING

The study was supported by the Israel Science Foundation and the Agnes Ginges Center for Neurogenetics.

## ACKNOWLEDGMENTS

We would like to thank our patients for their kind agreement to participate in the study. MN is grateful to the Azrieli Foundation for the award of an Azrieli Fellowship.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2018.00011/full#supplementary-material>

**Figure S1 |** Own-body transformation task: Participants viewed a schematic human figure with one hand marked, facing either toward them or away from them. In the ‘here’ condition participants were asked to judge from their own self-location whether the marked hand was on the right or the left side of the computer screen. In the ‘there’ condition, participants were asked to “project” themselves to the position represented by the schematic human figure, and from this self-location to indicate whether the marked hand would be their right or left hand. Correct responses for each case are indicated below each figure.



**Figure S2** | Depth electrodes locations in the hippocampus and lateral temporal cortex (LTC), shown on individual patients' MRI scans.

**Figure S3** | Electrophysiological results for the time-task in the left hemisphere. Intracranial evoked potentials (iEPs) from all electrodes used in the classification analysis are presented. LTC electrodes (up) show high early task modulation, whereas electrodes in the hippocampus (bottom) show high late task modulation. Shaded areas show time points of significant differences between conditions in two-tailed independent samples *t*-test ( $p < 0.05$ , uncorrected).

**Figure S4** | Electrophysiological results for the time-task in the right hemisphere. iEPs recorded at electrodes in the right LTC and right hippocampus. No clear distinction in task modulation is apparent between LTC electrodes and electrodes in the hippocampus. Shaded areas show time points of significant differences between conditions in two-tailed independent samples *t*-test ( $p < 0.05$ , uncorrected).

**Figure S5** | Demonstration of iEPs polarity-reversal in the electrodes shown in **Figure 1**. Some iEPs in **Figure 1** are of seemingly opposite polarity between Patients. This is the result of "polarity reversal" (Halgren et al., 1982). When recording iEPs from local generators, the polarity of the resulting iEP reverses as one records from two opposite sides of this generator. Observing such reversal in our data is expected since the exact relative position of electrodes differed between subjects. Note the iEPs similarity when plotting the reverse iEP (marked with an asterisk) in some of the electrodes.

**Figure S6** | Electrophysiological results for the space-task in the left hemisphere. iEPs from all electrodes used in the classification analysis are presented. No clear distinction in task modulation is apparent between LTC electrodes and electrodes in the hippocampus. Shaded areas show time points of significant differences between conditions in two-tailed independent samples *t*-test ( $p < 0.05$ , uncorrected).

**Figure S7** | Electrophysiological results for the space-task in the right hemisphere. iEPs recorded at electrodes in the right LTC and right hippocampus. No clear distinction in task modulation is apparent between LTC electrodes and electrodes in the hippocampus. Shaded areas show time points of significant differences between conditions in two-tailed independent samples *t*-test ( $p < 0.05$ , uncorrected).

**Figure S8** | The effect of reducing the number of electrodes used in the classification analysis. The distribution of cross-validation accuracy and corresponding *p*-values in the classification analysis of the MTT task, for subsets of 8 electrodes in the left hemispheres. Each subsets includes exactly 5 hippocampal electrodes and 3 lateral temporal electrodes, as in the right hemisphere. Although high accuracy values ( $>75\%$ ) were found in a large number of electrodes subsets (84/120), these findings were significant ( $p < 0.05$ ) for only a small fraction of the subsets (33/120).

**Table S1** | Electrodes locations.

**Table S2** | Early and late modulation in time task, left hemisphere.

## REFERENCES

- Addis, D. R., Pan, L., Vu, M. A., Laiser, N., and Schacter, D. L. (2009a). Constructive episodic simulation of the future and the past: distinct subsystems of a core brain network mediate imagining and remembering. *Neuropsychologia* 47, 2222–2238. doi: 10.1016/j.neuropsychologia.2008.10.026
- Addis, D. R., Roberts, R. P., and Schacter, D. L. (2011). Age-related neural changes in autobiographical remembering and imagining. *Neuropsychologia* 49, 3656–3669. doi: 10.1016/j.neuropsychologia.2011.09.021
- Addis, D. R., Sacchetti, D. C., Ally, B. A., Budson, A. E., and Schacter, D. L. (2009b). Episodic simulation of future events is impaired in mild Alzheimer's disease. *Neuropsychologia* 47, 2660–2671. doi: 10.1016/j.neuropsychologia.2009.05.018
- Addis, D. R., Wong, A. T., and Schacter, D. L. (2007). Remembering the past and imagining the future: common and distinct neural substrates during event construction and elaboration. *Neuropsychologia* 45, 1363–1377. doi: 10.1016/j.neuropsychologia.2006.10.016
- Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., and Buckner, R. L. (2010). Functional-anatomic fractionation of the brain's default network. *Neuron* 65, 550–562. doi: 10.1016/j.neuron.2010.02.005
- Anelli, F., Ciaramelli, E., Arzy, S., and Frassinetti, F. (2016). Prisms to travel in time: investigation of time-space association through prismatic adaptation effect on mental time travel. *Cognition* 156, 1–5. doi: 10.1016/j.cognition.2016.07.009
- Arzy, S., Bick, A., and Blanke, O. (2009b). Mental time in amnesia: evidence from bilateral medial temporal damage before and after recovery. *Cogn. Neuropsychol.* 26, 503–510. doi: 10.1080/02643290903439178
- Arzy, S., Collette, S., Ionta, S., Fornari, E., and Blanke, O. (2009a). Subjective mental time: the functional architecture of projecting the self to past and future. *Eur. J. Neurosci.* 30, 2009–2017. doi: 10.1111/j.1460-9568.2009.06974.x
- Arzy, S., Halje, P., Schechter, D. S., Spinelli, L., Seeck, M., and Blanke, O. (2014). Neural generators of psychogenic seizures: evidence from intracranial and extracranial brain recordings. *Epilepsy Behav.* 31, 381–385. doi: 10.1016/j.yebeh.2013.10.017
- Arzy, S., Molnar-Szakacs, I., and Blanke, O. (2008). Self in time: imagined self-location influences neural activity related to mental time travel. *J. Neurosci.* 28, 6502–6507. doi: 10.1523/JNEUROSCI.5712-07.2008
- Arzy, S., Thut, G., Mohr, C., Michel, C. M., and Blanke, O. (2006). Neural basis of embodiment: distinct contributions of temporoparietal junction and extrastriate body area. *J. Neurosci.* 26, 8074–8081. doi: 10.1523/JNEUROSCI.0745-06.2006
- Axmacher, N., Cohen, M. X., Fell, J., Haupt, S., Dümpelmann, M., Elger, C. E., et al. (2010). Intracranial EEG correlates of expectancy and memory formation in the human hippocampus and nucleus accumbens. *Neuron* 65, 541–549. doi: 10.1016/j.neuron.2010.02.006
- Axmacher, N., Mormann, F., Fernández, G., Cohen, M. X., Elger, C. E., and Fell, J. (2007). Sustained neural activity patterns during working memory in the human medial temporal lobe. *J. Neurosci.* 27, 7808–7816. doi: 10.1523/JNEUROSCI.0962-07.2007
- Ball, T., Kern, M., Mutschler, I., Aertsen, A., and Schulze-Bonhage, A. (2009). Signal quality of simultaneously recorded invasive and non-invasive EEG. *Neuroimage* 46, 708–716. doi: 10.1016/j.neuroimage.2009.02.028
- Bar, M. (2009). The proactive brain: memory for predictions. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 364, 1235–1243. doi: 10.1098/rstb.2008.0310
- Bartsch, T., Döhring, J., Rohr, A., Jansen, O., and Deuschl, G. (2011). CA1 neurons in the human hippocampus are critical for autobiographical memory, mental time travel, and autonoetic consciousness. *Proc. Natl. Acad. Sci. U.S.A.* 108, 17562–17567. doi: 10.1073/pnas.1110266108
- Benoit, R. G., and Schacter, D. L. (2015). Specifying the core network supporting episodic simulation and episodic memory by activation likelihood estimation. *Neuropsychologia* 75, 450–457. doi: 10.1016/j.neuropsychologia.2015.06.034
- Bird, C. M., and Burgess, N. (2008). The hippocampus and memory: insights from spatial processing. *Nat. Rev. Neurosci.* 9, 182–194. doi: 10.1038/nrn2335
- Blanke, O., Mohr, C., Michel, C. M., Pascual-Leone, A., Brugger, P., Seeck, M., et al. (2005). Linking out-of-body experience and self processing to mental own-body imagery at the temporoparietal junction. *J. Neurosci.* 25, 550–557. doi: 10.1523/JNEUROSCI.2612-04.2005
- Blanke, O., Morand, S., Thut, G., Michel, C. M., Spinelli, L., Landis, T., et al. (1999). Visual activity in the human frontal eye field. *Neuroreport* 10, 925–930. doi: 10.1097/00001756-199904060-00006
- Botzung, A., Denkova, E., and Manning, L. (2008). Experiencing past and future personal events: functional neuroimaging evidence on the neural bases of mental time travel. *Brain Cogn.* 66, 202–212. doi: 10.1016/j.bandc.2007.07.011
- Brunet, D., Murray, M. M., and Michel, C. M. (2011). Spatiotemporal analysis of multichannel EEG: CARTOOL. *Comput. Intell. Neurosci.* 2011:813870. doi: 10.1155/2011/813870
- Buckner, R. L., and Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57. doi: 10.1016/j.tics.2006.11.004
- Burgess, N., Barry, C., and O'Keefe, J. (2007). An oscillatory interference model of grid cell firing. *Hippocampus* 17, 801–812. doi: 10.1002/hipo.20327
- Button, K. S., Ioannidis, J. P., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14, 365–376. doi: 10.1038/nrn3475

- Buzsáki, G., and Moser, E. I. (2013). Memory, navigation and theta rhythm in the hippocampal-entorhinal system. *Nat. Neurosci.* 16, 130–138. doi: 10.1038/nn.3304
- Carpenter, G. A., and Grossberg, S. (1993). Normal and amnesic learning, recognition and memory by a neural model of cortico-hippocampal interactions. *Trends Neurosci.* 16, 131–137. doi: 10.1016/0166-2236(93)90118-6
- Chang, C. C., and Lin, C. J. (2011). LIBSVM: a library for support vector machines. *ACM Trans. Intell. Syst. Technol.* 2, 1–27. doi: 10.1145/1961189.1961199
- Collin, S. H., Milivojevic, B., and Doeller, C. F. (2015). Memory hierarchies map onto the hippocampal long axis in humans. *Nat. Neurosci.* 18, 1562–1564. doi: 10.1038/nn.4138
- Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018
- Davachi, L., and DuBrow, S. (2015). How the hippocampus preserves order. *Trends Cogn. Sci.* 19, 92–99. doi: 10.1016/j.tics.2014.12.004
- DeCoteau, W. E., and Kesner, R. P. (2000). A double dissociation between the rat hippocampus and medial caudoputamen in processing two forms of knowledge. *Behav. Neurosci.* 114, 1096–108. doi: 10.1037/0735-7044.114.6.1096
- Devito, L. M., and Eichenbaum, H. (2011). Memory for the order of events in specific sequences: contributions of the hippocampus and medial prefrontal cortex. *J. Neurosci.* 31, 3169–3175. doi: 10.1523/JNEUROSCI.4202-10.2011
- Doeller, C. F., King, J. A., and Burgess, N. (2008). Parallel striatal and hippocampal systems for landmarks and boundaries in spatial memory. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5915–5920. doi: 10.1073/pnas.0801489105
- Dombeck, D. A., Harvey, C. D., Tian, L., Looger, L. L., and Tank, D. W. (2010). Functional imaging of hippocampal place cells at cellular resolution during virtual navigation. *Nat. Neurosci.* 13, 1433–1440. doi: 10.1038/nn.2648
- Eichenbaum, H. (2013). Memory on time. *Trends Cogn. Sci.* 17, 81–88. doi: 10.1016/j.tics.2012.12.007
- Eichenbaum, H., and Cohen, N. J. (2014). Can we reconcile the declarative memory and spatial navigation views on hippocampal function? *Neuron* 83, 764–770. doi: 10.1016/j.neuron.2014.07.032
- Fink, G. R., Markowitsch, H. J., Reinkemeier, M., Bruckbauer, T., Kessler, J., and Heiss, W. D. (1996). Cerebral representation of one's own past: neural networks involved in autobiographical memory. *J. Neurosci.* 16, 4275–4282.
- Fortin, N. J., Agster, K. L., and Eichenbaum, H. B. (2002). Critical role of the hippocampus in memory for sequences of events. *Nat. Neurosci.* 5, 458–462. doi: 10.1038/nn834
- Gauthier, B., and van Wassenhove, V. (2016). Cognitive mapping in mental time travel and mental space navigation. *Cognition* 154, 55–68. doi: 10.1016/j.COgnITION.2016.05.015
- Gilboa, A., Ramirez, J., Köhler, S., Westmacott, R., Black, S. E., and Moscovitch, M. (2005). Retrieval of autobiographical memory in Alzheimer's disease: relation to volumes of medial temporal lobe and other structures. *Hippocampus* 15, 535–550. doi: 10.1002/hipo.20090
- Giovanello, K. S., Schnyer, D. M., and Verfaellie, M. (2004). A critical role of the anterior hippocampus in relational memory: evidence from an fMRI study comparing associative and item recognition. *Hippocampus* 14, 5–8. doi: 10.1002/hipo.10182
- Halgren, E., Squires, N. K., Wilson, C. L., and Crandall, P. H. (1982). Brain generators of evoked potentials: the late (endogenous) components. *Bull. Los Angeles Neurol. Soc.* 47, 108–23.
- Harnsberger, H. R., Osborn, G. A., Ross, J., and Andre, M. (2006). *Diagnostic and Surgical Imaging Anatomy: Brain, Head and Neck, Spine*. Lippincott Williams & Wilkins, Phi. Available online at: from www.amazon.com/Diagnostic-Surgical-Imaging-Anatomy-Published/dp/1931884293
- Hartley, T., Lever, C., Burgess, N., and O'Keefe, J. (2014). Space in the brain: how the hippocampal formation supports spatial cognition. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 369:20120510. doi: 10.1098/rstb.2012.0510
- Hassabis, D., Kumaran, D., and Maguire, E. A. (2007a). Using imagination to understand the neural basis of episodic memory. *J. Neurosci.* 27, 14365–14374. doi: 10.1523/JNEUROSCI.4549-07.2007
- Hassabis, D., Kumaran, D., Vann, S. D., and Maguire, E. A. (2007b). Patients with hippocampal amnesia cannot imagine new experiences. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1726–1731. doi: 10.1073/pnas.0610561104
- Hodges, J. R., Patterson, K., Oxbury, S., and Funnell, E. (1992). Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy. *Brain* 115(Pt 6), 1783–1806. doi: 10.1093/brain/115.6.1783
- Hopkins, R. O., Waldram, K., and Kesner, R. P. (2004). Sequences assessed by declarative and procedural tests of memory in amnesic patients with hippocampal damage. *Neuropsychologia* 42, 1877–1886. doi: 10.1016/j.neuropsychologia.2004.05.008
- Hurley, N. C., Maguire, E. A., and Vargha-Khadem, F. (2011). Patient HC with developmental amnesia can construct future scenarios. *Neuropsychologia* 49, 3620–3628. doi: 10.1016/j.neuropsychologia.2011.09.015
- Ionta, S., Sforza, A., Funato, M., and Blanke, O. (2013). Anatomically plausible illusory posture affects mental rotation of body parts. *Cogn. Affect. Behav. Neurosci.* 13, 197–209. doi: 10.3758/s13415-012-0120-z
- Jenkins, L. J., and Ranganath, C. (2016). Distinct neural mechanisms for remembering when an event occurred. *Hippocampus* 26, 554–559. doi: 10.1002/hipo.22571
- Kesner, R. P., Gilbert, P. E., and Barua, L. A. (2002). The role of the hippocampus in memory for the temporal order of a sequence of odors. *Behav. Neurosci.* 116, 286–290. doi: 10.1037/0735-7044.116.2.286
- Klein, S. B. (2013). The complex act of projecting oneself into the future. *Wiley Interdisc. Rev.* 4, 63–79. doi: 10.1002/wcs.1210
- Konishi, S., Asari, T., Jimura, K., Chikazoe, J., and Miyashita, Y. (2006). Activation shift from medial to lateral temporal cortex associated with recency judgements following impoverished encoding. *Cereb. Cortex* 16, 469–474. doi: 10.1093/cercor/bhi126
- Konkel, A., and Cohen, N. J. (2009). Relational memory and the hippocampus: representations and methods. *Front. Neurosci.* 3, 166–174. doi: 10.3389/neuro.01.023.2009
- Kramer, M. A., Eden, U. T., Lepage, K. Q., Kolaczky, E. D., Bianchi, M. T., and Cash, S. S. (2011). Emergence of persistent networks in long-term intracranial EEG recordings. *J. Neurosci.* 31, 15757–15767. doi: 10.1523/JNEUROSCI.2287-11.2011
- Kraus, B. J., Brandon, M. P., Robinson, R. J., Connerney, M. A., Hasselmo, M. E., and Eichenbaum, H. (2015). During running in place, grid cells integrate elapsed time and distance run. *Neuron* 88, 578–589. doi: 10.1016/j.neuron.2015.09.031
- Kriegeskorte, N., Simmons, W. K., Bellgowan, P. S., and Baker, C. I. (2009). Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* 12, 535–540. doi: 10.1038/nn.2303
- Kurczek, J., Wechsler, E., Ahuja, S., Jensen, U., Cohen, N. J., Tranel, D., et al. (2015). Differential contributions of hippocampus and medial prefrontal cortex to self-projection and self-referential processing. *Neuropsychologia* 73, 116–126. doi: 10.1016/j.neuropsychologia.2015.05.002
- Lachaux, J. P., Rudrauf, D., and Kahane, P. (2003). Intracranial EEG and human brain mapping. *J. Physiol. Paris* 97, 613–628. doi: 10.1016/j.jphysparis.2004.01.018
- Lehn, H., Steffenach, H. A., van Strien, N. M., Veltman, D. J., Witter, M. P., and Häberg, A. K. (2009). A specific role of the human hippocampus in recall of temporal sequences. *J. Neurosci.* 29, 3475–3484. doi: 10.1523/JNEUROSCI.5370-08.2009
- Levine, B. (2004). Autobiographical memory and the self in time: brain lesion effects, functional neuroanatomy, and lifespan development. *Brain Cogn.* 55, 54–68. doi: 10.1016/S0278-2626(03)00280-X
- Maguire, E. A. (2001). Neuroimaging studies of autobiographical event memory. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 356, 1441–1451. doi: 10.1098/rstb.2001.0944
- Maguire, E. A., and Mullally, S. L. (2013). The hippocampus: a manifesto for change. *J. Exp. Psychol.* 142, 1180–1189. doi: 10.1037/a0033650
- 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
- Markowitsch, H. J., and Staniloiu, A. (2011). Memory, autoegetic consciousness, and the self. *Conscious. Cogn.* 20, 16–39. doi: 10.1016/j.concog.2010.09.005
- Martin, A., and Chao, L. L. (2001). Semantic memory and the brain: structure and processes. *Curr. Opin. Neurobiol.* 11, 194–201. doi: 10.1016/S0959-4388(00)00196-3
- Maxwell, S. E., Kelley, K., and Rausch, J. R. (2008). Sample size planning for statistical power and accuracy in parameter estimation. *Annu. Rev. Psychol.* 59, 537–563. doi: 10.1146/annurev.psych.59.103006.093735

- McClelland, J. L., and Rogers, T. T. (2003). The parallel distributed processing approach to semantic cognition. *Nat. Rev. Neurosci.* 4, 310–322. doi: 10.1038/nrn1076
- McNaughton, B. L., and Morris, R. G. M. (1987). Hippocampal synaptic enhancement and information storage within a distributed memory system. *Trends Neurosci.* 10, 408–415. doi: 10.1016/0166-2236(87)90011-7
- Moll, M., and Miikkulainen, R. (1997). Convergence-zone episodic memory: analysis and simulations. *Neural Netw.* 10, 1017–1036. doi: 10.1016/S0893-6080(97)00016-6
- Nyberg, L., Kim, A. S., Habib, R., Levine, B., and Tulving, E. (2010). Consciousness of subjective time in the brain. *Proc. Natl. Acad. Sci. U.S.A.* 107, 22356–22359. doi: 10.1073/pnas.1016823108
- Ojala, M., and Garriga, G. C. (2010). Permutation tests for studying classifier performance. *J. Mach. Learn. Res.* 11, 1833–1863. doi: 10.1109/ICDM.2009.108
- O'Keefe, J., and Dostrovsky, J. (1971). The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* 34, 171–175. doi: 10.1016/0006-8993(71)90358-1
- Okuda, J., Fujii, T., Ohtake, H., Tsukiura, T., Tanji, K., Suzuki, K., et al. (2003). Thinking of the future and past: the roles of the frontal pole and the medial temporal lobes. *Neuroimage* 19, 1369–1380. doi: 10.1016/S1053-8119(03)00179-4
- Olsen, R. K., Moses, S. N., Riggs, L., and Ryan, J. D. (2012). The hippocampus supports multiple cognitive processes through relational binding and comparison. *Front. Hum. Neurosci.* 6:146. doi: 10.3389/fnhum.2012.00146
- 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
- Paz, R., Gelbard-Sagiv, H., Mukamel, R., Harel, M., Malach, R., and Fried, I. (2010). A neural substrate in the human hippocampus for linking successive events. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6046–6051. doi: 10.1073/pnas.0910834107
- Peer, M., Lyon, R., and Arzy, S. (2014). Orientation and disorientation: lessons from patients with epilepsy. *Epilepsy Behav.* 41, 149–157. doi: 10.1016/j.yebeh.2014.09.055
- Peer, M., Salomon, R., Goldberg, I., Blanke, O., and Arzy, S. (2015). Brain system for mental orientation in space, time, and person. *Proc. Natl. Acad. Sci. U.S.A.* 112, 11072–11077. doi: 10.1073/pnas.1504242112
- Preston, A. R., Shrager, Y., Dudukovic, N. M., and Gabrieli, J. D. (2004). Hippocampal contribution to the novel use of relational information in declarative memory. *Hippocampus* 14, 148–152. doi: 10.1002/hipo.20009
- Prince, S. E., Daselaar, S. M., and Cabeza, R. (2005). Neural correlates of relational memory: successful encoding and retrieval of semantic and perceptual associations. *J. Neurosci.* 25, 1203–1210. doi: 10.1523/JNEUROSCI.2540-04.2005
- Reber, P. J., and Squire, L. R. (1998). Encapsulation of implicit and explicit memory in sequence learning. *J. Cogn. Neurosci.* 10, 248–263. doi: 10.1162/089892998562681
- Rubin, A., Geva, N., Sheintuch, L., and Ziv, Y. (2015). Hippocampal ensemble dynamics timestamp events in long-term memory. *Elife* 4:e12247. doi: 10.7554/eLife.12247
- Rugg, M. D., and Vilberg, K. L. (2013). Brain networks underlying episodic memory retrieval. *Curr. Opin. Neurobiol.* 23, 255–260. doi: 10.1016/j.conb.2012.11.005
- Schacter, D. L., and Addis, D. R. (2007). The cognitive neuroscience of constructive memory: remembering the past and imagining the future. *Philos. Trans. R. Soc. B* 362, 773–786. doi: 10.1098/rstb.2007.2087
- Schacter, D. L., and Addis, D. R. (2009). On the nature of medial temporal lobe contributions to the constructive simulation of future events. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 364, 1245–1253. doi: 10.1098/rstb.2008.0308
- Schacter, D. L., Addis, D. R., Hassabis, D., Martin, V. C., Spreng, R. N., and Szpunar, K. K. (2012). The future of memory: remembering, imagining, and the brain. *Neuron* 76, 677–694. doi: 10.1016/j.neuron.2012.11.001
- Scoville, W. B., and Milner, B. (2000). Loss of recent memory after bilateral hippocampal lesions. 1957. *J. Neuropsychiatry Clin. Neurosci.* 12, 103–113. doi: 10.1176/jnp.12.1.103-a
- Shalev-Shwartz, S., and Ben-David, S. (2014). *Understanding Machine Learning: from Theory to Algorithms*. New York, NY: Cambridge University Press.
- Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011). False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol. Sci.* 22, 1359–1366. doi: 10.1177/0956797611417632
- Spreng, R. N., and Grady, C. L. (2010). Patterns of brain activity supporting autobiographical memory, prospection, and theory of mind, and their relationship to the default mode network. *J. Cogn. Neurosci.* 22, 1112–1123. doi: 10.1162/jocn.2009.21282
- Spreng, R. N., Mar, R. A., and Kim, A. S. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J. Cogn. Neurosci.* 21, 489–510. doi: 10.1162/jocn.2008.21029
- Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychol. Rev.* 99, 195–231. doi: 10.1037/0033-295X.99.2.195
- Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiol. Learn. Mem.* 82, 171–177. doi: 10.1016/j.nlm.2004.06.005
- Squire, L. R., Van Der Horst, A. S., McDuff, S. G., Frascino, J. C., Hopkins, R. O., and Mauldin, K. N. (2010). Role of the hippocampus in remembering the past and imagining the future. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19044–19048. doi: 10.1073/pnas.1014391107
- Staresina, B. P., Fell, J., Do Lam, A. T., Axmacher, N., and Henson, R. N. (2012). Memory signals are temporally dissociated in and across human hippocampus and perirhinal cortex. *Nat. Neurosci.* 15, 1167–1173. doi: 10.1038/nn.3154
- St Jacques, P. L., Conway, M. A., Lowder, M. W., and Cabeza, R. (2011). Watching my mind unfold versus yours: an fMRI study using a novel camera technology to examine neural differences in self-projection of self versus other perspectives. *J. Cogn. Neurosci.* 23, 1275–1284. doi: 10.1162/jocn.2010.21518
- Suddendorf, T., and Corballis, M. C. (2007). The evolution of foresight: what is mental time travel, and is it unique to humans? *Behav. Brain Sci.* 30, 299–313. doi: 10.1017/S0140525X07001975
- Svoboda, E., McKinnon, M. C., and Levine, B. (2006). The functional neuroanatomy of autobiographical memory: a meta-analysis. *Neuropsychologia* 44, 2189–2208. doi: 10.1016/j.neuropsychologia.2006.05.023
- Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., and Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011:879716. doi: 10.1155/2011/879716
- Tulving, E. (1985). Memory and consciousness. *Can. Psychol.* 26, 1–12. doi: 10.1037/h0080017
- Tulving, E. (2002). Episodic memory: from mind to brain. *Annu. Rev. Psychol.* 53, 1–25. doi: 10.1146/annurev.psych.53.100901.135114
- Vanni-Mercier, G., Mauguière, F., Isnard, J., and Dreher, J. C. (2009). The hippocampus codes the uncertainty of cue-outcome associations: an intracranial electrophysiological study in humans. *J. Neurosci.* 29, 5287–5294. doi: 10.1523/JNEUROSCI.5298-08.2009
- Yonelinas, A. (2002). The nature of recollection and familiarity: a review of 30 years of research. *J. Mem. Lang.* 46, 441–517. doi: 10.1006/jmla.2002.2864

**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 © 2018 Schurr, Nitzan, Eliahou, Spinelli, Seeck, Blanke and Arzy. 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) and the copyright owner 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.



# The Amount of Time Dilation for Visual Flickers Corresponds to the Amount of Neural Entrainments Measured by EEG

Yuki Hashimoto and Yoko Yotsumoto\*

Department of Life Sciences, University of Tokyo, Tokyo, Japan

## OPEN ACCESS

### Edited by:

Arpan Banerjee,  
National Brain Research Centre  
(NBRC), India

### Reviewed by:

Cota Navin Gupta,  
Indian Institute of Technology  
Guwahati, India  
Jeffrey Valla,  
Cornell University, United States

### \*Correspondence:

Yuko Yotsumoto  
cyuko@mail.ecc.u-tokyo.ac.jp

**Received:** 09 August 2017

**Accepted:** 19 April 2018

**Published:** 07 May 2018

### Citation:

Hashimoto Y and Yotsumoto Y (2018)  
The Amount of Time Dilation for Visual  
Flickers Corresponds to the Amount  
of Neural Entrainments Measured by  
EEG. *Front. Comput. Neurosci.* 12:30.  
doi: 10.3389/fncom.2018.00030

The neural basis of time perception has long attracted the interests of researchers. Recently, a conceptual model consisting of neural oscillators was proposed and validated by behavioral experiments that measured the dilated duration in perception of a flickering stimulus (Hashimoto and Yotsumoto, 2015). The model proposed that flickering stimuli cause neural entrainment of oscillators, resulting in dilated time perception. In this study, we examined the oscillator-based model of time perception, by collecting electroencephalography (EEG) data during an interval-timing task. Initially, subjects observed a stimulus, either flickering at 10-Hz or constantly illuminated. The subjects then reproduced the duration of the stimulus by pressing a button. As reported in previous studies, the subjects reproduced 1.22 times longer durations for flickering stimuli than for continuously illuminated stimuli. The event-related potential (ERP) during the observation of a flicker oscillated at 10-Hz, reflecting the 10-Hz neural activity phase-locked to the flicker. Importantly, the longer reproduced duration was associated with a larger amplitude of the 10-Hz ERP component during the inter-stimulus interval, as well as during the presentation of the flicker. The correlation between the reproduced duration and the 10-Hz oscillation during the inter-stimulus interval suggested that the flicker-induced neural entrainment affected time dilation. While the 10-Hz flickering stimuli induced phase-locked entrainments at 10-Hz, we also observed event-related desynchronizations of spontaneous neural oscillations in the alpha-frequency range. These could be attributed to the activation of excitatory neurons while observing the flicker stimuli. In addition, neural activity at approximately the alpha frequency increased during the reproduction phase, indicating that flicker-induced neural entrainment persisted even after the offset of the flicker. In summary, our results suggest that the duration perception is mediated by neural oscillations, and that time dilation induced by flickering visual stimuli can be attributed to neural entrainment.

**Keywords:** time perception, duration perception, neural entrainment, time, EEG



## INTRODUCTION

A major focus of interval-timing studies has been the mechanism of how physical time-flow is converted into a mental representation of duration. Many studies have proposed that neural oscillators with periodic activations are utilized for the physical-mental conversion of time, but physiologic aspects of the oscillators are still controversial (Gibbon et al., 1984; Treisman et al., 1994; Matell and Meck, 2004).

In psychophysical studies, flickering visual stimuli have been widely used to investigate the function of the oscillators. It has been reported that a flickering stimulus causes observers to overestimate the duration of the stimulus, and such an overestimation is called “time dilation” (Treisman and Brogan, 1992; Kanai et al., 2006; Ortega and López, 2008). Hashimoto and Yotsumoto (2015) examined time dilation using various flickering frequencies, and conducted simulations by a model that integrated flicker-induced neural entrainments with a previously proposed oscillator-based model (Matell and Meck, 2004). The behavioral results were consistent with the simulations, indicating that neural entrainment can account for flicker-induced time dilation.

The neurophysiological aspect of the flicker-induced time dilation has also been investigated. Herbst et al. (2013) reported that a set of stimuli flickering above the flicker fusion frequency (Landis, 1954) evoked steady-state visually evoked potentials (SSVEPs; Regan, 1977), as the stimuli were not perceived as a flicker and did not cause time dilation. Therefore, they concluded that conscious perception of a flicker, instead of neural activity triggered by a flickering stimulus, played a crucial role in time dilation.

However, EEGs were not recorded in Herbst et al. (2013) during their interval-timing task. Morillon et al. (2009) showed that neural activity differs when attending to the temporal aspect of an event, and when attending to other aspect such as color. They reported larger BOLD activities in dorsolateral prefrontal cortex and temporal-parietal junction when the subjects attended to the temporal aspect, suggesting the temporal processing network is controlled by attention. Therefore, neural activities while attending to the duration of the flicker might be different from neural activities during passive observation of the flicker. To investigate the effect of a flicker on the interval-timing network, it is essential to examine neural activity while subjects attend to the temporal aspect of the flicker stimuli.

In this study, we investigated the physiological relations between neural oscillation and time perception. Recently, we proposed a model which assumed that multiple oscillators with various intrinsic frequencies process interval-timing (Hashimoto and Yotsumoto, 2015). The model extended the striatal beat-frequency model (Buhusi and Meck, 2005) which hypothesized that a duration is encoded as the timing on which a specific subset of oscillatory neurons simultaneously activates. We further simulated the activity of the oscillators when a flicker entrained the oscillators. We demonstrated that when the frequencies of oscillators were drawn to the flickering frequency, the simultaneous activation of the oscillatory neurons occurred earlier than the encoded duration, which in turn caused time

dilation. In previous studies of time distortion, entrainment of oscillators was considered a factor capable of inducing time distortion (Treisman and Brogan, 1992; Treisman et al., 1994), while the presentation of flickering stimuli mainly caused time dilation (Treisman and Brogan, 1992; Kanai et al., 2006; Ortega and López, 2008; Kaneko and Murakami, 2009). Hence, neural entrainment was not considered the dominant source of flicker-induced time dilation; instead, changes in arousal level (Treisman and Brogan, 1992; Ortega and López, 2008) and temporal cueing (Kanai et al., 2006; Kaneko and Murakami, 2009; Herbst et al., 2013) became the focus of increased research. Hashimoto and Yotsumoto’s model successfully demonstrated time dilation and lack of time contraction by combining neural entrainment with an existing neural model of time perception. In addition, their model can be physiologically verified because it is directly linked to a neural model of time perception and flicker-induced time dilation.

In the present study, we took our model as a working hypothesis, and recorded EEG data while subjects performed a duration reproduction task. First, we measured the EEG power spectrum while subjects observed a flicker. The model predicts that presentation of a flicker would entrain the time-encoding neural network and affect neural oscillations in the brain. Consequently, the neural activities would be phase-locked to the flicker, and the neural activities may last even after the disappearance of the flicker. Second, we examined whether the reproduced duration of a flickering stimulus and the amplitude of the SSVEP would correlate trial by trial. The model predicts that an increase in time dilation would be observed with an increase in neural entrainment, which would be observed as a greater SSVEP.

Additionally, we analyzed the EEG recordings for duration reproduction when no flicker was presented. Some previous studies have reported that the effect of flicker-induced time dilation lasted after the offset of the flicker (Johnston et al., 2006; Burr et al., 2007). In addition, neural entrainment was reported to last  $\sim 0.5$  s after the offset of the flicker (Spaak et al., 2014). Therefore, we analyzed EEG recordings during the reproduction phase as well as the flicker observation phase.

## METHODS

### Subjects

Thirteen volunteers (4 men; age range, 18–23 years) with normal or corrected-to-normal vision participated in the experiment. One subject was excluded from the analyses because Fp1 and Fp2 channel malfunctioned resulting in failure in detecting eye movements and blinks. The data collected from the other 12 subjects (4 men; age range, 18–23 years) were used in the following analyses. All participants were blind to the purpose of the study. This study was carried out in accordance with the recommendations of the ethics boards of the University of Tokyo with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the institutional review boards of the University of Tokyo.



## Apparatus

The experiment was conducted in a dark soundproof room. The stimuli were presented on a 23.6-inch LCD monitor with a 120-Hz refresh rate, 1,920-pixel width and 1,090-pixel height (VIEWPixx 3D; VPixx Technologies Inc., Saint-Bruno, QC Canada). The viewing distance was set to 57.3 cm with a chin rest. The experiment was conducted with MATLAB 2014 (Mathworks, Natick, MA USA) and the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007).

EEG recordings were obtained at a sampling rate of 512 Hz using a 32-channel EEG system with a signal amplifier, active electrodes, a battery box (g.USBamp, g.LADYbird, and g.GAMMABox, respectively; g.tec medical engineering, Schiedlberg, Austria) and Simulink with MATLAB 2012. The electrodes were mounted using an AsiaCap (BrainProducts, Gilching, Germany) on the following positions: Fp1, Fp2, AFz, Fz, F3, F4, F7, F8, FC1, FC2, FC5, FC6, T7, T8, Cz, C3, C4, CP1, CP2, CP5, CP6, Pz, P3, P4, P7, P8, POz O1, O2. Additionally, three electrodes were mounted on the left-side of the left-eye, the right-side of the right-eye, and the bottom of the left-eye to monitor for eye movements and blinking. The ground electrode was mounted on Fpz and the reference electrode was mounted on the left earlobe. The ERP and time-frequency representation analyses were conducted with Fieldtrip software (Oostenveld et al., 2011) and custom MATLAB scripts.

## Stimuli

A circular disc with a  $4^\circ$  radius was presented against a black background at the center of the display, and a fixation cross was overlaid on the circular disc. For each subject, the luminance of the circular disc was set to be subjectively equiluminant to 25 cd/m<sup>2</sup> white by the heterochromatic flicker photometry with a 20-Hz square-wave (Bone and Landrum, 2004) to reduce eyestrain and the effect of luminance adaptation. The luminance of the fixation cross was also set to be subjectively equiluminant to 12.5 cd/m<sup>2</sup> white.

## Procedure

EEG data were recorded while the subjects performed a duration reproduction task. The time course of the reproduction task is illustrated in **Figure 1**.

Before each trial, a green circular disc with a red edge annulus was presented on the display for 1.2 s; during this time the subjects were allowed to blink, but not otherwise. A trial started with a green circular disc overlaid by a gray fixation cross that was presented for 0.6–1.4 s. It was followed by the presentation of a white circular disc, which defined the standard duration that the subjects were asked to remember. The white circular disc was either continuously illuminated or flickering at 10 Hz. After a reproduction cue was delivered by changing the color of the fixation cross from gray to red, the subjects reproduced the remembered standard duration by pressing a space key with a right index finger. The inter-stimulus interval (ISI) between the first presentation of a white circular disc and the reproduction cue was randomly chosen from a range of 0.4–1.0 s. During the reproduction phase, a white circular disc was again presented on the display as visual feedback. As soon as the key was released, the

white circular disc disappeared, and a green circular disc replaced it for 0.6 s.

It should be noted that the luminance of the white circular disc was set at the same value for all conditions. Therefore, the maximum luminance was similar for all stimuli, while the temporal average of the luminance was lower for the flickering than for the constantly illuminated stimuli. This setting was a conservative choice in order to avoid false positives of time dilation and SSVEP. If the averaged luminance had been controlled using a brighter flicker, it might have led to an overestimation of time dilation (Xuan et al., 2007) and visually evoked potentials (Norcia et al., 2015), due to brightness. To avoid these critical false positives, we chose to control for maximum luminance and not averaged luminance.

Each subject performed 300 experimental trials and 60 catch trials, resulting in a total of 360 trials. In the experimental trials, the standard duration was always 1.0 s, while in the catch trials the duration of the standard duration was jittered between 0.5 and 1.5 s. There were 150 experimental trials and 30 catch trials in which the first white circular disc was continuously illuminated (“static” condition). In the other 150 experimental trials and 30 catch trials, the first white circular disc was flickering (“flickering” condition). The 360 trials were divided into 30 blocks. The subjects were allowed to take a break anytime between these 12 blocks.

## Behavioral Data Analysis

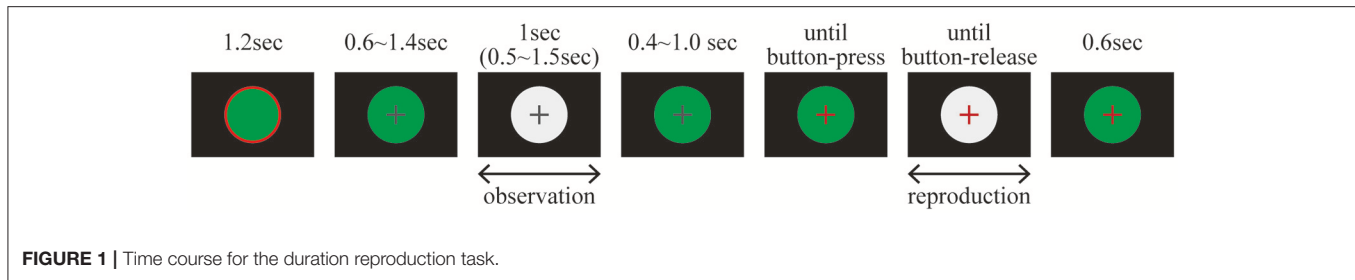
To remove trials with extraordinary short or long reproduction, which was caused by instantaneous button press and overlook of the offset signal, the extraordinary reproductions were detected by applying distanced-based outlier detection with  $\epsilon = 0.1$  s and  $\pi = 10^{-5}$  (Knox and Ng, 1998) for each subject and condition, resulting excluding 0.6% of trials from the following analyses.

For each subject, a *t*-test was applied to evaluate the difference in the reproduced durations in the experimental trials between the “static” and “flickering” conditions to determine whether the subjective duration of a flickering stimulus dilated. In addition, for the catch trials in which the standard duration differed trial-by-trial, the correlation between the standard durations and the reproduced durations were tested for each subject and for each condition to confirm that the subjects attended to the duration of the stimuli.

For subsequent EEG analysis, the 150 experimental trials for each stimulus type were sorted in accordance with the subjects’ reproduced durations. The top 50 trials were classified as the “long” reproduction trials; the bottom 50, “short”; the intermediate 50, “middle.”

## Electroencephalogram Data Analysis Preprocessing

An online high-pass filter at 0.1 Hz and an online notch filter at 50 Hz were applied to the EEG data, while recording. The recordings were divided into epochs from the beginning to the end of the trial, followed by an application of an offline band-pass filter from 0.2 to 128 Hz and an offline notch filter. Eye movements and blinks were detected as transient fluctuations in the electrooculogram (EOG), and



the trials containing the artifacts were excluded. After the trial removal, independent component analysis (ICA) was applied to the 29 EEG channels without EOGs to correct for muscle artifacts (Makeig et al., 1996). By visual inspection of the independent components for each subject, low-frequency (<1 Hz) components bisymmetrically distributed around frontal electrodes and high-frequency (>20 Hz) components evident only at a few parietotemporal electrodes were excluded as the muscle artifacts originated from the forehead and temple. Subsequently, the EEG was reconstructed by the remaining independent components and the following analyses were applied to the reconstructed EEG (Jung et al., 2000).

### Event-Related Potentials

Each trial was divided into two phases. The observation phase was defined as a period of  $-0.6$  to  $1.6$  s, time-locked to the onset of the standard duration; and the reproduction phase was defined as a period of  $-0.84$  to  $1.14$  s, time-locked to the onset of reproduction. For each subject and stimulus type, the ERP was calculated by averaging the preprocessed signals. For each ERP, baseline correction was applied by subtracting the mean signal within an interval of  $-0.1$  to  $0$  s. Then, the ERPs acquired from Pz, P3, P4, POz, O1, and O2 were averaged.

For each of the 1-s standard duration intervals and the 0.4-s interval occurring just after the offset of the standard duration, the SSVEP amplitude was calculated by applying the discrete Fourier transform (Cooley et al., 1969) to the ERP for the intervals measured for each subject. The detailed formulation is provided as Supplementary Formula 1. For each of the intervals, a within-subject  $t$ -test was applied to evaluate the differences in the SSVEP amplitude between “static” and “flickering” conditions. The amplitudes of the second (20 Hz), third (30 Hz), and fourth (40 Hz) harmonics were also compared for the “static” and “flickering” conditions using within-subject  $t$ -tests.

To further investigate the relationship between behavior and SSVEP, the SSVEP amplitudes of the “long,” “middle,” and “short” reproduction trials were modeled and evaluated using analysis of variance (ANOVA) followed by a *post-hoc* Tukey’s honest significant difference (HSD) test. The type of reproduced duration (“long,” “middle,” or “short”) was set as a fixed effect and the subject was set as a random effect in the model. Cohen’s  $d$ s (Cohen, 1988) was calculated for each difference (Equation 1).

$$\text{Cohen's } d = \frac{\text{sample mean}}{\text{sample standard deviation}} \times \sqrt{2} \quad (1)$$

In addition, for each of the second, third, and fourth SSVEP harmonic values, the differences in the respective amplitudes of the “long,” “middle,” and “short” reproduction trials were tested by ANOVA and the *post-hoc* Tukey’s HSD test. The fixed and random effects were set as those for the ANOVA for the base frequency. Cohen’s  $d$ s values were also calculated in the same manner.

### Time-Frequency Representations

The time-frequency representation was calculated for each trial by projecting the preprocessed signal onto the time-frequency representation using a short-term Fourier transform and applying an adaptive Hanning window length. The window length of the Hanning taper was set at 7 cycles per window. For each time-frequency representation, a baseline-correction was applied by subtracting the mean amplitude within an interval of  $-0.1$  to  $0$  s for each frequency. Following the baseline-correction, the time-frequency representations were averaged for each subject and stimulus type and subsequently, the time-frequency representations acquired from Pz, P3, P4, POz, O1, and O2 were averaged.

The difference in time-frequency representation between the “static” and “flickering” conditions was tested by a cluster-based permutation (Maris and Oostenveld, 2007) with *ft\_timelockstatistics* function in the Fieldtrip software. In the cluster-based permutation test, two conditions were compared by calculating  $t$ -values for every time-frequency data point. A continuum in which the  $t$ -value exceeded a certain criterion was clustered and  $t$ -values in the cluster were summed up, resulting in a  $T$ -value of the cluster. If there were multiple  $T$ -values originating from multiple clusters,  $T$ -values, except the largest one, were rejected. To compute the distribution of  $T$ -values based on the null hypothesis, the label of condition was randomly assigned to the data sets, and  $T$ -values were resampled repeatedly. The position of the original  $T$ -value and the resampled  $T$ -values were sorted and the percentile of the original  $T$ -value was calculated. If the percentile of original  $T$ -value was smaller than 2.5% or larger than 97.5%, the two conditions were concluded to be significantly different.

## RESULTS

### Behavior

The mean reproduced duration for the “static” condition was 0.95 s (SD: 0.08 s) and 1.16 s (SD: 0.19 s) for the “flickering” condition. The difference was significant for all subjects ( $p <$

$1.0 \times 10^{-5}$  for each subject,  $p < 5.0 \times 10^{-4}$  with Bonferroni correction), which indicated that the flicker was perceived to be longer than the constantly illuminated stimulus. The correlation between the standard duration and the reproduced duration in catch trials was 0.78 (SD: 0.11) for the “static” condition, and 0.76 (SD: 0.09) for the “flickering” condition. The correlation was significant in all subjects and conditions ( $p < 1.0 \times 10^{-3}$  for all subjects and conditions,  $p < 0.05$  with Bonferroni correction), which indicated that the subjects attended to each standard duration accurately. Additional information regarding the distribution of the reproduced duration for each subject and condition is provided in Supplementary Figure 1, and the detailed results of the correlation analyses are reported in Supplementary Table 1. The reproduced durations in the “long,” “middle,” and “short” reproduction trials in the “static” and “flickering” conditions were shown in **Table 1**. The mean differences of reproduced duration between the “long” and “short” reproduction trials were 0.24 s (SD: 0.06 s) and 0.30 s (SD: 0.10 s) in the “static” and “flickering” conditions, respectively.

## Electroencephalogram

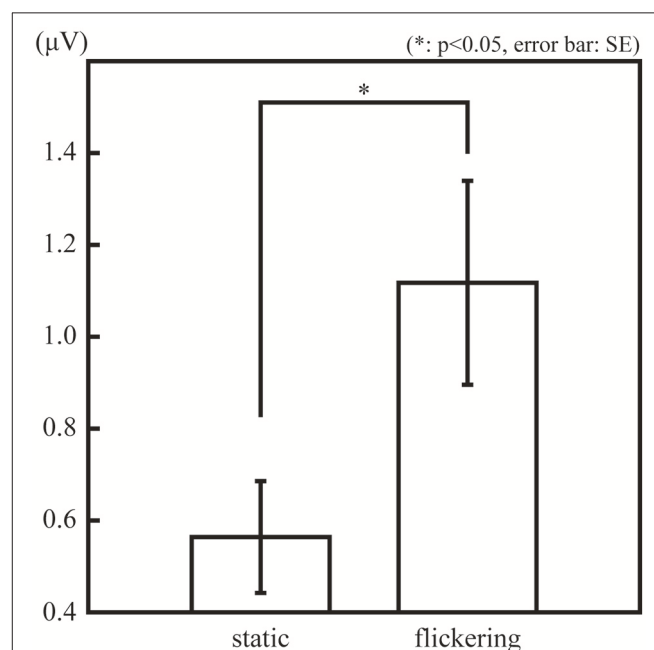
### Steady State Visually Evoked Potential

**Figure 2** illustrates the amplitude of the SSVEP (10-Hz ERP component) in the observation phase. The amplitude of the SSVEP was significantly larger in the “flickering” condition than in the “static” condition [ $t_{(11)} = 3.02$ ,  $p = 0.01$ ], suggesting the flickering stimulus evoked a 10-Hz neural activity phase-locked to the change in luminance of the flicker. The amplitudes of the second, third, and fourth SSVEP harmonics were also significantly larger in the “flickering” condition than in the “static” condition [ $t_{(11)} = 6.36, 2.41, 3.85$ ;  $p = 0.0001, 0.03, 0.002$ , respectively]. The topographic representation and the frequency spectrum of the SSVEP are illustrated in Supplementary Figures 2,3, respectively.

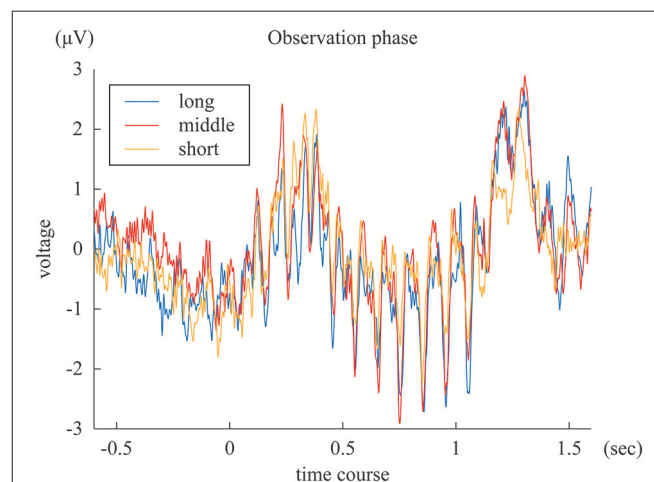
The SSVEPs during the 1-s standard duration in the “long,” “middle,” and “short” reproduction trials were illustrated in **Figure 3**. The amplitudes of the SSVEP in the “long,” “middle,” and “short” reproduction trials were 1.35  $\mu$ V (SD: 0.89  $\mu$ V), 1.26  $\mu$ V (SD: 0.81  $\mu$ V), and 0.98  $\mu$ V (SD: 0.66  $\mu$ V), which are illustrated in **Figure 4A**. ANOVA revealed that the amplitudes of the SSVEP were different across the types of reproduced duration [ $F_{(2, 22)} = 8.20$ ,  $p = 0.004$  with and without Greenhouse-Geisser’s sphericity correction,  $\epsilon = 0.98$ ]. The *post-hoc* Tukey’s HSD test showed significant differences in the amplitude of the SSVEP between the “long” and “short” reproduction trials, and between

the “middle” and “short” reproduction trials ( $p < 0.005$  and  $p < 0.05$ , Cohen’s  $d = 1.26$  and  $d = 1.53$ , respectively). There was no significant difference in the SSVEP between the “long” and “middle” reproduction trials.

In addition, the SSVEPs during the 0.4 s immediately after the offset of the standard duration were calculated for each subject and stimulus type. Because this 0.4-s interval was an inter-stimulus interval and the same green circular disc was



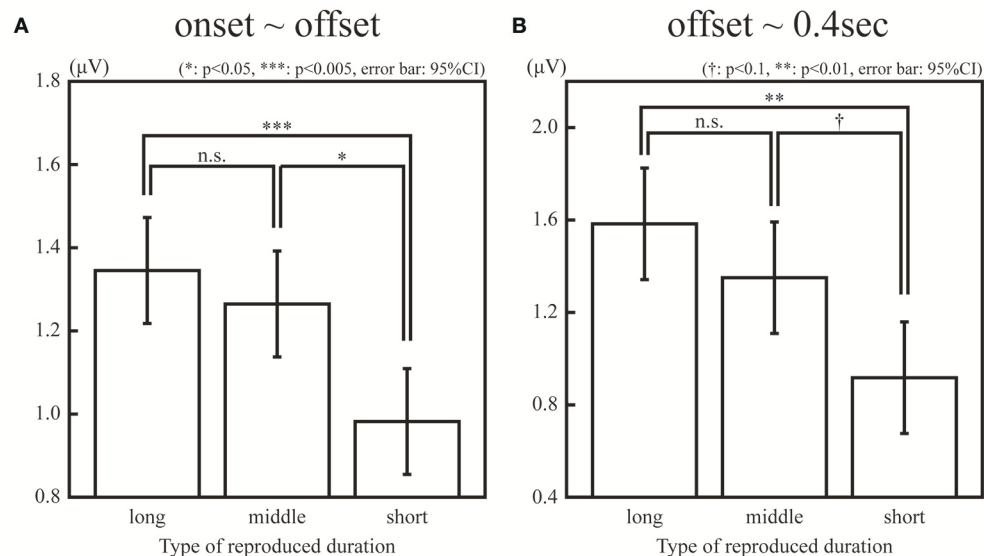
**FIGURE 2** | The amplitude of the 10-Hz ERP components in the observation phase in the “static” and “flickering” conditions. The asterisk indicates the significant difference with  $\alpha = 0.05$ . The error bars represent the standard errors of the mean.



**FIGURE 3** | The blue, red, and yellow lines represent ERP of “long,” “middle,” and “short” reproduction trials, averaged across subjects.

**TABLE 1** | Mean and SD of reproduced duration of the trials categorized as “long,” “middle,” and “short.”

		“Long”	“Middle”	“Short”
Constantly illuminated	Mean	1.07	0.95	0.83
	(S.D.)	(0.09)	(0.08)	(0.09)
Flickering	Mean	1.31	1.15	1.00
	(S.D.)	(0.22)	(0.19)	(0.18)



**FIGURE 4 |** Amplitudes of the 10-Hz ERP components in the “long,” “middle,” “short” reproduction trials during observation of a flicker (**A**) and after the offset of the flicker (**B**). The single, double, and triple asterisk(s) indicate the significant difference with  $\alpha = 0.05$ , 0.01, and 0.005, respectively. The dagger (†) indicates the marginally significant difference with  $\alpha = 0.1$ .

continuously presented both in the “static” and “flickering” conditions, the difference of the SSVEPs in this interval was not due to the difference of the presented stimuli, but it reflected the difference of the neural entrainment lasting even after the offset of the stimuli. The SSVEP after the offset of the standard duration was significantly larger in the “flickering” condition than in the “static” condition [ $t_{(11)} = 2.46$ ,  $p < 0.05$ ]. The amplitudes in the “long,” “middle,” and “short” reproduction trials in the “flickering” condition were  $1.58 \mu\text{V}$  (SD:  $1.44 \mu\text{V}$ ),  $1.35 \mu\text{V}$  (SD:  $0.98 \mu\text{V}$ ), and  $0.92 \mu\text{V}$  (SD:  $0.72 \mu\text{V}$ ) respectively, which are illustrated in **Figure 4B**. ANOVA revealed that amplitudes were different across the types of reproduced durations [ $F_{(2, 22)} = 6.18$ ,  $p = 0.007$  and  $0.02$  with and without Greenhouse-Geisser’s sphericity correction,  $\epsilon = 0.67$ ]. The Tukey’s HSD test showed a significant difference in the amplitude between the “long” and “short” reproduction trials ( $p < 0.01$ , Cohen’s  $d = 1.09$ ). There was a marginally significant difference between the “middle” and the “short” reproduction trials ( $p < 0.1$ , Cohen’s  $d = 1.34$ ). There was no significant difference between the “long” and “middle” reproduction trials.

The second, third, and forth harmonics of the SSVEP during the standard duration were not significantly different among the “long,” “middle,” and “short” reproduction trials [ $F_{(2, 22)} = 1.82$ ,  $0.14$ ,  $0.91$ , respectively;  $p > 0.1$  for all harmonics]. The amplitude of the SSVEP harmonics during  $0.4 \text{ s}$  just after the offset were not significantly different either [ $F_{(2, 22)} = 0.35$ ,  $0.70$ ,  $0.51$ , respectively;  $p > 0.1$  for all harmonics].

### Event-Related Potential and Time-Frequency Representation

The ERPs in the observation phase were averaged across subjects, and shown in **Figure 5A**. In the observation phase,

the 10-Hz flicker caused oscillatory EEG fluctuations at 10 Hz (SSVEP). **Figure 5B** illustrates the difference in time-frequency representation in the observation phase between the “static” and “flickering” condition, represented by the  $t$ -values calculated across subjects. The presentation of a flicker decreased the EEG amplitude in the wide range of frequency around 10 Hz ( $p = 0.010$ ), suggesting large event-related desynchronization (Klimesch et al., 2007) at approximately the alpha band in the “flickering” condition.

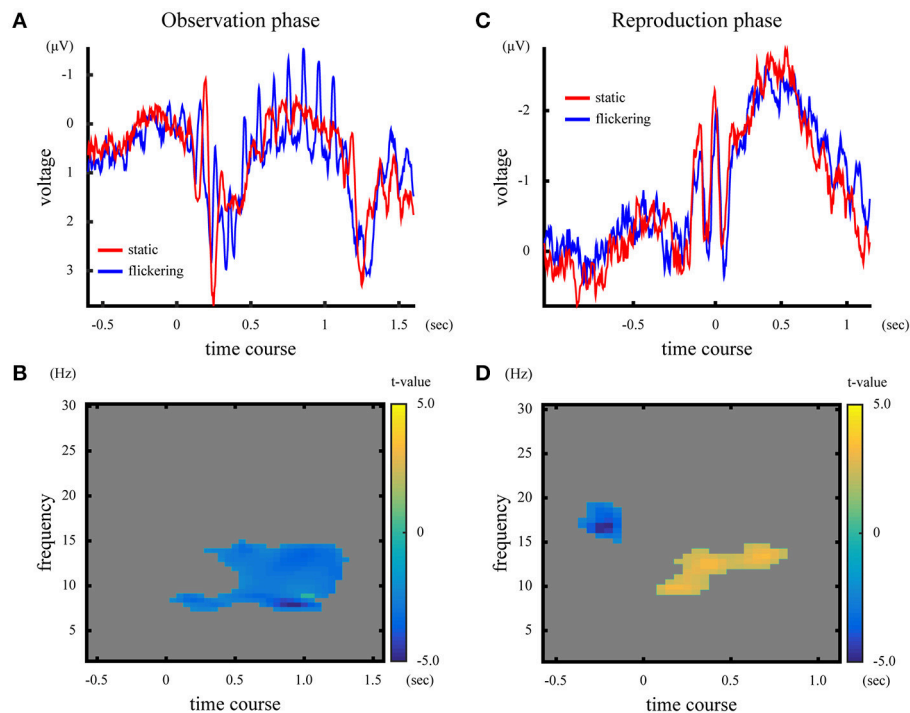
The averaged ERP and the  $t$ -values in time-frequency representation in the reproduction phase are illustrated in **Figures 5C,D**. In the “flickering” condition, the EEG amplitude at approximately the alpha band during the reproduction increased, while before the onset of reproduction the amplitude at approximately the beta band decreased ( $p = 0.018$  for the alpha band activity,  $p = 0.048$  for the beta band activity). Note that the visual stimuli presented during the reproduction phase did not flicker even in the flickering condition. Additionally, the time-frequency representations of EEG averaged across subjects for each condition and phase are shown in Supplementary Figure 4.

## DISCUSSION

In the present study, we measured the neural correlations between EEG and time dilation in order to evaluate the effects of neural entrainment in time dilation. We found that 10-Hz flickers induced time dilation replicating the results of previous studies, and that the reproduced duration correlated with the amplitude of the 10-Hz ERP component both before and after the flicker was offset.

Flicker-induced oscillations that lasted even after the disappearance of the flicker were also reported by the previous





**FIGURE 5 |** [Upper panels] ERP and time-frequency representation in the observation phase (A) and the reproduction phase (C). The blue and red lines represent ERPs in the “static” and “flickering” condition, respectively. [Lower panels] Difference of time-frequency representation between the “static” and “flickering” condition in the observation phase (B), and in the reproduction phase (D). The yellow area indicates that the time-frequency amplitude was larger in the “flickering” condition, while the blue area shows that the time-frequency amplitude was larger in the “static” condition. The gray area indicates no significant differences between the two conditions. The color bar indicates the *t*-value at the time-frequency data point.

study (Spaak et al., 2014), and the EEG oscillations were considered to reflect neural activity being entrained to the flicker. Therefore, the correlation between the reproduced duration and the amplitude of the flicker-induced 10-Hz oscillation supported our hypothesis that the presentation of a flicker can induce neural entrainment, and that this neural entrainment would cause flicker-induced time dilation. It should be mentioned that a significant difference in the 10-Hz ERP amplitude was not observed comparing long and middle reproduction trials. However, the lack of difference observed in the 10-Hz ERP amplitude does not necessarily suggest a lack of difference in the magnitude of neural entrainment. Rather, it might be more plausible to attribute this lack of difference to the relationship between neural entrainment and the consequent ERP. Because the EEG is the summation of neural activities, it is natural to assume that the oscillatory 10-Hz ERP observed in our results represented the collective activity of the entrained neural oscillators. Mathematically, the phase coherence among oscillators and the amplitude of their mean activity are associated by an S-shaped function (Supplementary Figure 5), thus when comparing two conditions having a stronger neural entrainment, the difference in the ERP amplitude gradually decreases and is less detectable. This property may explain why there was no difference observed in the 10-Hz ERP amplitude comparing the middle and long reproduction trials. Conversely, it is less likely that there was no change in ERP because there was no

difference in the magnitude of neural entrainment. Had the magnitude of the neural entrainment been small in the short reproduction trials, and large but identical in the middle and long reproduction trials, the distribution of the reproduced duration would have been skewed. However, such skewness was not observed in our results (Supplementary Figure 1). Therefore, the lack of difference in the ERP amplitude should be attributed to the relationship between the neural entrainment and the resulting ERP, rather than on an similar neural entrainment.

The ERP amplitudes of the harmonics were also larger for the flickering stimuli than for the continuously illuminated stimuli. However, the amplitudes of the harmonics did not correlate with the reproduced duration. The difference might be attributed to the distribution of oscillators in the time-encoding network. In the model proposed in Hashimoto and Yotsumoto (2015), the oscillating frequencies of the time-encoding network distribute more densely at around alpha frequencies compared to other frequencies. Therefore, the entrainment of the oscillators at around 10 Hz would have larger impact on time perception than the entrainment of the other oscillators. The difference of impact might have led to the result that the amplitude of the 10-Hz ERP component correlated with the amount of time dilation while the amplitude of the harmonics did not.

Although the ERP analysis revealed the flicker induced phase-locked 10-Hz neural activity during the observation of a flicker (Figure 4A), decrease of EEG amplitude at ~10 Hz was



observed through time-frequency representation (**Figure 4B**). The decrement of neural activity at  $\sim 10$  Hz could be attributed to the large amplitude of event-related desynchronization. Klimesch et al. (2007) reviewed that EEG-components at approximately the alpha band often decrease while performing cognitive tasks, and such a decrement is considered to reflect a decrease in spontaneous alpha activity caused by excitatory brain processing. In our temporal reproduction task, the subjects' visual system had to process a greater amount of change in luminance when the stimulus flickered, which could result in higher event-related desynchronization. If the event-related desynchronization exceeded the amplitude of 10-Hz neural activity evoked by the presentation of a flicker, there would have been no cluster exhibiting increased amplitude at 10 Hz with the flickering stimulus. In fact, the 10-Hz neural activation was not evident during the visual inspections of the trial-by-trial EEG analysis. Conversely, in the ERP analysis, averaging canceled out spontaneous neural activities that had randomly distributed phases. Therefore, the 10-Hz SSVEP phase locked to the flicker was clearly observable.

Interestingly, the presentation of a flicker also evoked neural activity at approximately the alpha band, even in the reproduction phase when no flicker was presented. There are some possible explanations for this phenomenon: First, this oscillation might be attributed to the "replay of the flicker in the mind" phenomenon, as in this phase, the subjects remembered the standard duration defined by the presentation of a flicker. However, this explanation is unlikely because the frequency of increased activity was slightly higher than 10 Hz. If the EEG oscillation was because of the "replay of the flicker in the mind" phenomenon, the reproduced duration would be  $< 1.0$  s because the higher EEG frequency would indicate subjectively faster elapse of time in the reproduction phase than that in the observation phase. This was not the case with the findings of our behavioral tests. The second possibility is the aftereffect of the neural activation induced by observation of a flicker. Previously, studies have reported that presentation of a flicker altered perception of the subsequent stimulus. Johnston et al. (2006) reported that presentation of a flicker compressed the perceived duration of the subsequent stimulus presented 0.5 s later, suggesting an aftereffect of a flicker on interval-timing. Droit-Volet and Wearden (2002) conducted an experiment with children, and similarly reported the effect of a flicker on the duration perception of the subsequent stimulus. In line with these studies, the increased activity during the reproduction phase could be interpreted as an aftereffect of the neural activation induced by the preceding flicker. This explanation is congruent with the model of neural entrainment. Alpha-band neural entrainment induced by the presentation of a flicker have been reported to last around 0.5 s (Spaak et al., 2014), and in our results, the sustained neural entrainment also sustained for 0.4 s after the offset of the flicker. Therefore, the alpha oscillatory EEG data observed in our study might reflect the aftereffect.

In our experiments, we evaluated the correlation between time dilation and neural entrainment by conducting both inter-stimulus comparisons ("static" and "flickering" conditions) and intra-stimulus comparisons ("long" and "short" reproduction trials). The results supported the hypothesis that neural entrainment induces subjective time dilation. However, in our studies, we only measured EEGs having continuously illuminated stimuli and 10-Hz flickers. It would be of value to conduct additional experiments with stimuli flickering at different frequencies in order to better examine whether the correlation between neural entrainment and time dilation is a general phenomenon. In addition, recording EEGs with arrhythmic flickers will also be helpful in distinguishing the neural entrainment due to oscillators, the neural activities reflecting each flash in the flicker and, in particular, it will contribute to identify the ultimate source of event-related desynchronization reflecting excitatory brain processing induced by the flicker. Despite these reservations, our results clearly showed that the observation of a flicker during an interval-timing task evoked a periodic neural activity, which persisted even after the offset of the flicker, and the prolonged perception was associated with a larger periodic neural activity.

In summary, (1) presentation of a flicker induced subjective time dilation, and the amount of time dilation correlated with the amplitude of neural entrainment induced by the flicker. (2) The observation of a flicker evoked large event-related desynchronization at approximately the alpha band, suggesting excitatory brain processing. (3) An aftereffect of the flicker was observed during the reproduction phase because of the increase in EEG amplitude at approximately the alpha band. These results indicate that neural entrainment can be triggered by the presentation of a flicker, and support the working hypothesis that neural entrainment results in the distortion of interval-timing perception.

## AUTHOR CONTRIBUTIONS

YH and YY conceived and designed the experiments; YH performed the experiments and analyzed the data; YH and YY wrote the manuscript.

## FUNDING

This work was supported by CisHub of U-Tokyo, and Grants-in-Aid for Scientific Research for YY (KAKENHI-2511903, 16H03749, 17K18693).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2018.00030/full#supplementary-material>

## REFERENCES

- Bone, R. A., and Landrum, J. T. (2004). Heterochromatic flicker photometry. *Arch. Biochem. Biophys.* 430, 137–142. doi: 10.1016/j.abb.2004.04.003
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- 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
- Burr, D., Tozzi, A., and Morrone, M. C. (2007). Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nat. Neurosci.* 10:423. doi: 10.1038/nn1874
- Cohen, J. (1988). “Case 3: One sample of n observations,” in *Statistical Power Analysis for the Behavioral Science, 2nd Edn.* (Hillsdale, NJ: Lawrence Erlbaum Associates), 45–48.
- Cooley, J., Lewis, P., and Welch, P. (1969). The finite Fourier transform. *IEEE Trans. Audio Electroacoust.* 17, 77–85. doi: 10.1109/TAU.1969.1162036
- Droit-Volet, S., and Wearden, J. (2002). Speeding up an internal clock in children? Effects of visual flicker on subjective duration. *Q. J. Exp. Psychol. Sect. B* 55, 193–211. doi: 10.1080/02724990143000252
- 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
- Hashimoto, Y., and Yotsumoto, Y. (2015). Effect of temporal frequency spectra of flicker on time perception: behavioral testing and simulations using a striatal beat frequency model. *Tim. Time Percept.* 3, 201–222. doi: 10.1163/22134468-03002049
- 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
- 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
- Jung, T. P., Makeig, S., Humphries, C., Lee, T. W., Mckeown, M. J., Iragui, V., et al. (2000). Removing electroencephalographic artifacts by blind source separation. *Psychophysiology* 37, 163–178. doi: 10.1111/1469-8986.3720163
- Kanai, R., Paffen, C. L., Hogendoorn, H., and Verstraten, F. A. (2006). Time dilation in dynamic visual display. *J. Vis.* 6, 1421–1430. doi: 10.1167/6.12.8
- Kaneko, S., and Murakami, I. (2009). Perceived duration of visual motion increases with speed. *J. Vis.* 9:14. doi: 10.1167/9.7.14
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., et al. (2007). What's new in Psychtoolbox-3. *Perception* 36, 1–16. doi: 10.1068/v070821
- 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
- Knox, E. M., and Ng, R. T. (1998). “Algorithms for mining distancebased outliers in large datasets,” in *Proceedings of the International Conference on Very Large Data Bases* (New York, NY), 392–403.
- Landis, C. (1954). Determinants of the critical flicker-fusion threshold. *Physiol. Rev.* 34, 259–286. doi: 10.1152/physrev.1954.34.2.259
- Makeig, S., Bell, A. J., Jung, T.-P., and Sejnowski, T. J. (1996). “Independent component analysis of electroencephalographic data,” in *Advances in Neural Information Processing Systems 8*, eds D. S. Touretzk, M. C. Mozer, and M. E. Hasselmo (Cambridge, MA: MIT Press), 145–151.
- Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and ME G-data. *J. Neurosci. Methods* 164, 177–190. doi: 10.1016/j.jneumeth.2007.03.024
- 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
- Morillon, B., Kell, C. A., and Giraud, A. L. (2009). Three stages and four neural systems in time estimation. *J. Neurosci.* 29, 14803–14811. doi: 10.1523/JNEUROSCI.3222-09.2009
- Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottareau, B. R., and Rossion, B. (2015). The steady-state visual evoked potential in vision research: a review. *J. Vis.* 15:4. doi: 10.1167/15.6.4
- 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
- Ortega, L., and López, F. (2008). Effects of visual flicker on subjective time in a temporal bisection task. *Behav. Process.* 78, 380–386. doi: 10.1016/j.beproc.2008.02.004
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366
- Regan, D. (1977). Steady-state evoked potentials. *J. Opt. Soc. Am.* 67, 1475–1489. doi: 10.1364/JOSA.67.001475
- Spaak, E., de Lange, F. P., and Jensen, O. (2014). Local entrainment of alpha oscillations by visual stimuli causes cyclic modulation of perception. *J. Neurosci.* 34, 3536–3544. doi: 10.1523/JNEUROSCI.4385-13.2014
- Treisman, M., and Brogan, D. (1992). Time perception and the internal clock: effects of visual flicker on the temporal oscillator. *Eur. J. Cogn. Psychol.* 4, 41–70. doi: 10.1080/09541449208406242
- 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. Sect. A* 47, 241–289. doi: 10.1080/14640749408401112
- Xuan, B., Zhang, D., He, S., and Chen, X. (2007). Larger stimuli are judged to last longer. *J. Vis.* 7:2. doi: 10.1167/7.10.2

**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 © 2018 Hashimoto and Yotsumoto. 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) and the copyright owner 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.



# Temporal Reference, Attentional Modulation, and Crossmodal Assimilation

Yingqi Wan and Lihan Chen\*

School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, China

Crossmodal assimilation effect refers to the prominent phenomenon by which ensemble mean extracted from a sequence of task-irrelevant distractor events, such as auditory intervals, assimilates/biases the perception (such as visual interval) of the subsequent task-relevant target events in another sensory modality. In current experiments, using visual Ternus display, we examined the roles of temporal reference, materialized as the time information accumulated before the onset of target event, as well as the attentional modulation in crossmodal temporal interaction. Specifically, we examined how the global time interval, the mean auditory inter-intervals and the last interval in the auditory sequence assimilate and bias the subsequent percept of visual Ternus motion (element motion vs. group motion). We demonstrated that both the *ensemble (geometric) mean* and the *last interval* in the auditory sequence contribute to bias the percept of visual motion. Longer mean (or last) interval elicited more reports of group motion, whereas the shorter mean (or last) auditory intervals gave rise to more dominant percept of element motion. Importantly, observers have shown dynamic adaptation to the temporal reference of crossmodal assimilation: when the target visual Ternus stimuli were separated by a long gap interval after the preceding sound sequence, the assimilation effect by ensemble mean was reduced. Our findings suggested that crossmodal assimilation relies on a suitable temporal reference on adaptation level, and revealed a general temporal perceptual grouping principle underlying complex audio-visual interactions in everyday dynamic situations.

**Keywords:** temporal window, temporal ventriloquism effect, central tendency effect, assimilation, attention

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

Keita Mitani,  
Tokyo Institute of Technology, Japan  
Hakan Karşilar,  
Koç University, Turkey

### \*Correspondence:

Lihan Chen  
CLH@pku.edu.cn

**Received:** 20 February 2018

**Accepted:** 16 May 2018

**Published:** 05 June 2018

### Citation:

Wan Y and Chen L (2018) Temporal Reference, Attentional Modulation, and Crossmodal Assimilation. *Front. Comput. Neurosci.* 12:39. doi: 10.3389/fncom.2018.00039

## INTRODUCTION

Multisensory interaction has been traditionally revealed to take place over a narrowed window time—i.e., within a presumed “temporal window” (Meredith et al., 1987; Powers et al., 2009; Vroomen and Keetels, 2010; Wallace and Stevenson, 2014; Gupta and Chen, 2016). For example, paired sound/tactile events presented in temporal proximity to paired visual events can alter the perceived interval between the visual stimuli, and hence bias the perception of visual apparent motion (Keetels and Vroomen, 2008; Chen et al., 2010; Shi et al., 2010). The above illusions have been typically known as temporal ventriloquism (Chen and Vroomen, 2013). Studies on temporal ventriloquism indeed suggested that crossmodal events appearing in temporal proximities have higher probabilities of “correlation” and even “causation” relations (Ernst and Di Luca, 2011; Parise et al., 2012). Based on those relations, sensory events with higher functional priorities (such

as “precision” in timing) would calibrate/attract the counterpart events (with lower functional appropriateness) from the other modalities, give rise to successful *multisensory integration*. During the integration, multisensory events within a presumed short time window will largely obey the “assumption of unity,” in which the coherent representation of multiple events become possible when they have been deemed as coming from a common source (Vatakis and Spence, 2007, 2008; Misceo and Taylor, 2011; Chuen and Schutz, 2016; Chen and Spence, 2017). As a result, the effectiveness of crossmodal interaction is enhanced.

However, the presumed “temporal window” for integration has often been violated in many ecological scenarios. Take an example: upon hearing the whistle of a running car behind us, after a decent long delay, we can know exactly what kind of the “car” is approaching and then make prompt avoidance. This indicates that humans can adaptively use the prior knowledge and employ the temporal/spatial information (including environmental cues associated with the sound) to facilitate the perceptual decision. This daily scenario, however, imposes a great challenge for human perception. How are perceptual grouping and correspondences between events achieved when the crossmodal events are separated both in longer temporal ranges and with larger temporal disparities? Moreover, for the longer temporal range, observers have difficulties in memorizing all the events and the processing of the sensory properties (including time information) would probably exceed their working memory capacities (Cowan, 2001; Klemen et al., 2009; Klemen and Chambers, 2012; Cohen et al., 2016). Therefore, the efficiency of crossmodal interaction will be reduced accordingly. The complex timing scenario as well as the challenge for time cognition also stems from the variance of the multiple time intervals. In short temporal range (such as around 2 s), human observers could discriminate the short temporal intervals when the coefficient of variance (i.e., “CV,” the ratio of the interval deviation to its baseline value) is less than 0.3. The discrimination ability is greatly reduced when the CV is above 0.3 (Allan, 1974; Getty, 1975; Penney et al., 2000).

To cope with the above constraints, human observers adopt one of the efficient perceptual strategies—“ensemble coding” to process the mean properties of multiple events. For example, people can extract the mean rhythm of a given sound sequence and use this information to allocate visual attention and facilitate the detection of target events (Miller et al., 2013). Recent studies have shown that this averaging process is highly dependent on the temporal reference. The temporal reference included the generally global time interval before the onset of target event(s), the variabilities of the multiple intervals and the critical information of the last interval (Jones and McAuley, 2005; Acerbi et al., 2012; Cardinal, 2015; Karaminis et al., 2016). One compelling example is the central tendency effect within the broader framework of Bayesian optimization (Jazayeri and Shadlen, 2010; Shi et al., 2013; Shi and Burr, 2016; Roach et al., 2017), whereby incorporating the mean of the statistical distribution in the estimation would assimilate the estimates toward the mean (Jazayeri and Shadlen, 2010; Burr et al., 2013; Karaminis et al., 2016). For example, the estimation of a target property, such as the duration of an event, is assimilated toward

to the mean duration of previously encountered target events (i.e., event history) (Nakajima et al., 1992; Burr et al., 2013; Shi et al., 2013; Roach et al., 2017). The central tendency effect indicates that human observers exploit predictive coding using the averaged sensory properties (Shi and Burr, 2016). The predictive coding framework states that the brain produces a Bayesian estimate of the environment (Friston, 2010). A strong mismatch between the prediction and the actual sensory input leads to an update of the internal model, and could trigger observable changes in perceptual decision. During this updating, attentional process can be considered as a form of predictive coding to establish an expectation of the moments in time until the task-relevant, to be integrated stimulus inputs arrive (Klemen and Chambers, 2012). On the other hand, the temporal reference (including temporal window) for crossmodal interaction is flexible by perceptual training (Powers et al., 2009, 2012), repeated exposure (adaptation) to the sensory stimuli (Mégevand et al., 2013), or recalibration process through experience (Sugano et al., 2010, 2012, 2016; Bruns and Röder, 2015; Habets et al., 2017). The flexibility of temporal window has also been shown to be shaped by the individual differences (Hillock et al., 2011; Stevenson et al., 2012, 2014; Lewkowicz and Flom, 2014; Chen et al., 2016; Hillock-Dunn et al., 2016).

Time perception is intrinsically related with attention and memory (Block and Gruber, 2014). Attention has been revealed to act as an essential cognitive faculty in integrating information in the multisensory mind (Duncan et al., 1997; Talsma et al., 2007, 2010; Donohue et al., 2011, 2015; Tang et al., 2016). (Selective) attention improves the efficiency of pooling task-relevant information - multiple (complex) properties (Buchan and Munhall, 2011; Li et al., 2016). Withdrawing attention has been shown in other tasks/paradigms to degrade the representation of individual sensory properties (Alsus et al., 2005, 2014). In the central tendency effect, observers processed task-relevant sensory properties to obtain the subsequent perceptual decision. However, whether/how attentional modulation would deplete the limited attentional resources for ensemble coding and hence play a role in the crossmodal assimilation, has not been empirically examined.

Therefore, in the present study, we aimed to examine how the temporal reference and the attentional processing would affect the crossmodal assimilation. We adopted “temporal ventriloquism effect” with visual Ternus display. We investigated how the temporal configurations between an auditory sequence (with multiple inter-intervals) and the visual Ternus display (with one interval) modulate the visual apparent-motion percepts. Ternus display can elicit two distinct percepts of visual apparent motion: “element” motion or “group” motion, determined by the visual inter-stimulus-interval ( $ISI_V$ ) between the two display frames (with other stimulus settings being fixed). Element motion is typically observed with short  $ISI_V$  (e.g., of 50 ms), and group motion with long  $ISI_V$  (e.g., of 230 ms) (Ternus, 1926; Shi et al., 2010) (see Supplement 1 for visual animation of Ternus display). Previously we have shown that when two beeps were presented in temporal proximity to, or synchronously with, the two visual frames respectively, the beeps can systematically bias the transitional threshold of visual apparent motion (Shi et al.,



2010). Here we extended the Ternus temporal ventriloquism paradigm to investigate the temporal crossmodal ensemble coding. We implemented five experiments to address this issue. Experiments 1 and 2 examined the role of temporal window-interval gap between the offset of sound sequence and the onset of target Ternus display, to show the temporal constraints of central tendency effect. Experiment 3 compared the central tendency effect with the recency effect, by manipulating both the mean auditory interval and the last auditory interval. In Experiment 4, we fixed the last interval to be equal to the transitional threshold of perceiving element vs. group motion in the pretest, and manipulated the mean auditory inter-interval to show a genuine central tendency effect during crossmodal assimilation. In Experiment 5, we implemented dual-tasks and asked observers to perform the visual Ternus task while fulfilling a concurrent task of counting oddball sounds. Overall, the current results revealed that crossmodal central tendency effect is subject to the temporal reference (including the length of global time interval, the mean interval and the last interval for a given sound sequence) but less dependent on attentional modulation.

## MATERIALS AND METHODS

### Participants

A total of 60 participants (14, 13, 7, 12, 14 in Experiments 1–5), ages ranging from 18 to 33 years, took part in the main experiments. A *post-hoc* power estimation has shown the statistical powers are generally approaching or above 0.8 for the given sample sizes. All observers had normal or corrected-to-normal vision and reported normal hearing. The experiments were performed in compliance with the institutional guidelines set by the Academic Affairs Committee, School of Psychological and Cognitive Sciences, Peking University. The protocol was approved by the Committee for Protecting Human and Animal Subjects, School of Psychological and Cognitive Sciences, Peking University. All participants gave written informed consent in accordance with the Declaration of Helsinki, and were paid for their time on a basis of 40 CNY/hour, i.e., 6.3 US dollars/hour.

### Apparatus and Stimuli

The experiments were conducted in a dimly lit (luminance: 0.09 cd/m<sup>2</sup>) room. Visual stimuli were presented at the center of a 22-inch CRT monitor (FD 225P) at a screen resolution of 1024 × 768 pixels and a refresh rate of 100 Hz. Viewing distance was 57 cm, maintained by using a chin rest. A Ternus display consisted of two stimulus frames, each containing two black discs (10.30 cd/m<sup>2</sup>; disc diameter and separation between discs: 1.6° and 3° of visual angle, respectively) presented on a gray background (16.3 cd/m<sup>2</sup>). The two frames shared one element location at the center of the monitor, while containing two other elements located at horizontally opposite positions relative to the center (see **Figure 1A**). Each frame was presented for 30 ms; the inter-stimulus interval (ISI<sub>V</sub>) between the two frames was randomly selected from the range of 50–230 ms, with a step size of 30 ms.

Mono sound beeps (1,000 Hz pure tone, 65 dB SPL, 30 ms, except in Experiment 5 where pure tones with pitches of either 1,000 Hz or 500 Hz were given) were generated and delivered via

an M-Audio card (Delta 1010) to a headset (Philips, SHM1900). No ramps were applied to modulate the shape of the tone envelope. To ensure accurate timing of the auditory and visual stimuli, the duration of the visual stimuli and the synchronization of the auditory and visual stimuli were controlled via the monitor's vertical synchronization pulses. The experimental program was written with Matlab (Mathworks Inc.) and the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007).

## Experimental Design

### Practice

Prior to the formal experiment, participants were familiarized with Ternus displays of either typical “element motion” (with an interval of 50 ms) or “group motion” (with an interval of 260 ms) in a practice block. They were asked to discriminate the two types of apparent motion by pressing the left or the right mouse button, respectively. The mapping between response button and type of motion was counterbalanced across participants. During practice, when an incorrect response was made, immediate feedback appeared on the screen showing the correct response (i.e., element or group motion). The practice session continued until the participant reached a mean accuracy of 95%. All participants achieved this within 120 trials.

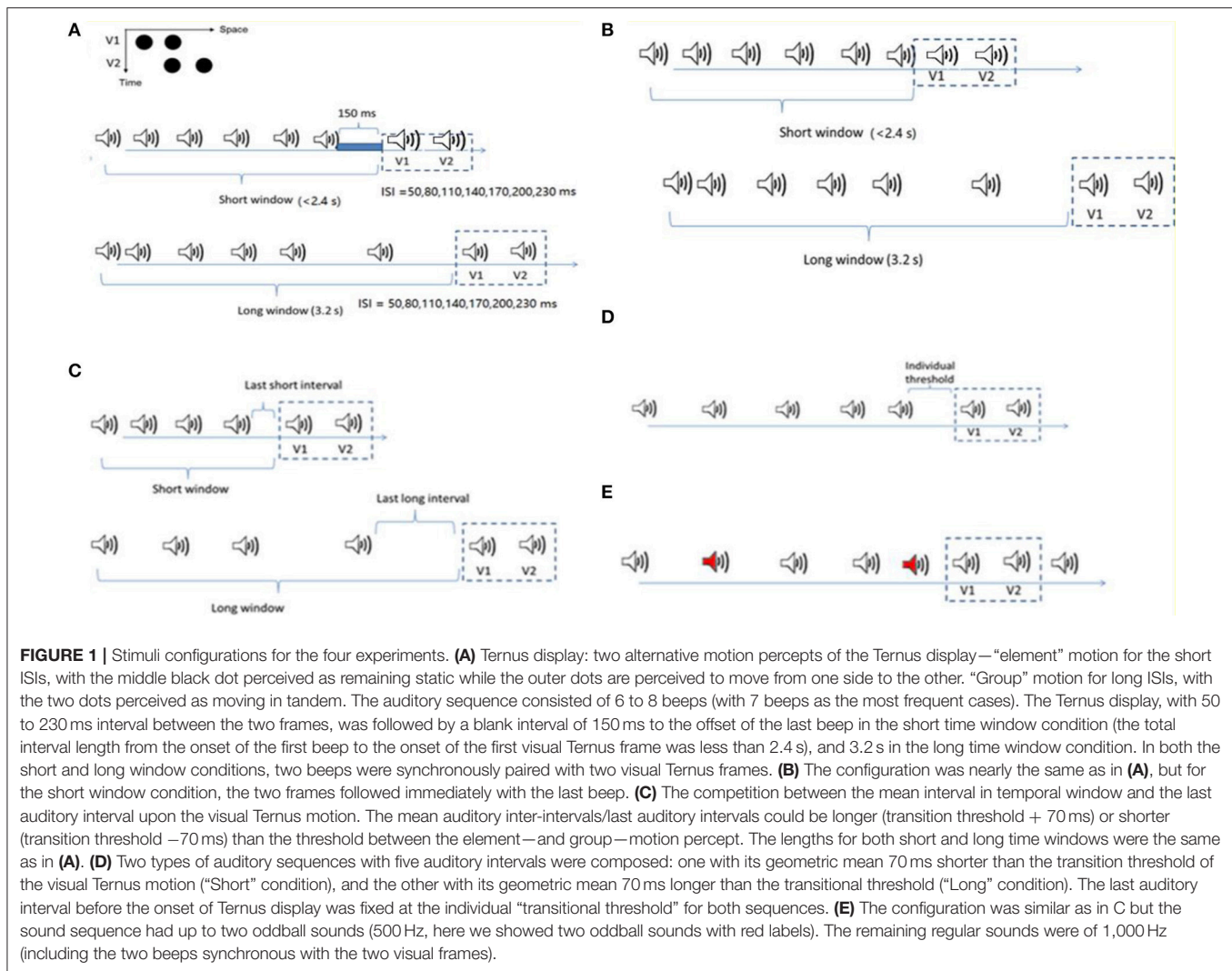
### Pre-test

For each participant, the transition threshold between element and group motion was determined in a pre-test session. A trial began with the presentation of a central fixation cross lasting 300 to 500 ms. After a blank screen of 600 ms, the two Ternus frames were presented, synchronized with two auditory tones [i.e., baseline: ISI<sub>V</sub>(visual) = ISI<sub>A</sub>(auditory)]; this was followed by a blank screen of 300 to 500 ms, prior to a screen with a question mark prompting the participant to make a two-alternative forced-choice response indicating the type of perceived motion (element or group motion). The ISI<sub>V</sub> between the two visual frames was randomly selected from one of the following seven intervals: 50, 80, 110, 140, 170, 200, and 230 ms. There were 40 trials for each level of ISI<sub>V</sub>, counterbalanced with left- and rightward apparent motion. The presentation order of the trials was randomized for each participant. Participants performed a total of 280 trials, divided into 4 blocks of 70 trials each. After completing the pre-test, the proportions of the group motion responses across seven intervals were fitted to the psychometric curve using a logistic function (Treutwein and Strasburger, 1999; Wichmann and Hill, 2001). The transitional threshold, that is, the point of subjective equality (PSE) at which the participant was likely to report the two motion percepts equally, was calculated by estimating 50% of reporting of group motion on the fitted curve. The just noticeable difference (JND), an indicator of the sensitivity of apparent motion discrimination, was calculated as half of the difference between the lower (25%) and upper (75%) bounds of the thresholds from the psychometric curve.

### Main Experiments

In the main experiments, the procedure for presenting visual stimuli was the same as in the pre-test session, except that prior to the occurrence of two Ternus-display frames, an





auditory sequence consisting a variable number of 6–8 beeps was presented (see below for the details of the onset of Ternus-display frames relative to that of the auditory sequence). A trial began with the presentation of a central fixation marker, randomly for 300 to 500 ms. After a 600-ms blank interval, the auditory train and the visual Ternus frames were presented (see **Figure 1A**), followed sequentially by a blank screen of 300 to 500 ms and a screen with a question mark at the screen center prompting participants to indicate the type of motion they had perceived: element vs. group motion (non-speeded response). During the experiment, observers were simply asked to indicate the type of visual motion ("element" or "group" motion) that they perceived, while ignoring the beeps. After the response, the next trial started following a random inter-trial interval of 500 to 700 ms.

In Experiment 1, the visual Ternus frames were preceded by an auditory sequence of 6–8 beeps with the geometric mean of inter-stimulus interval [ISIA(uditory), i.e.,  $ISI_A$ ], manipulated to be 70 ms shorter than, or 70 ms longer than the transition threshold estimated in the pre-test. The [ISIV(isual), i.e.,  $ISI_V$ ] between the two visual Ternus frames was randomly selected

from one of the following seven intervals: 50, 80, 110, 140, 170, 200, and 230 ms. The total auditory sequence consisted of 6–8 beeps. Visual Ternus frames were presented on most of all trials (672 trials in total) following the last beep; the remaining were catch trials (72 trials) in which the frames were inset in the sound sequence to break up anticipatory processes. For the short time window of the auditory sequence, the time interval from the onset of the first beep to the onset of the first visual frame was less than 2.4 s, and the gap interval between the offset of the last beep and the onset of the first Ternus frame was 150 ms. For the long time window, the total interval from the onset of the sound to the first visual frame was 3.2 s. In both the short and long window conditions, two beeps were synchronously paired with two visual Ternus frames. All the trials were randomized and organized in 12 blocks (62 trials for each block).

In Experiment 2, the settings were the same as in Experiment 1, except for the condition: the visual frames were following immediately with the offset of the last beep.

In Experiment 3, we introduced two factors of interval modulations: the mean interval of temporal window and the last

auditory interval. The mean auditory inter-intervals and the last auditory intervals could be larger (transition threshold + 70 ms) or shorter (transition threshold - 70 ms) than the threshold between the element- and group- motion percept. Therefore, there were four combinations of the “interval” conditions: both the mean interval and the last interval were shorter (i.e., “MeanSLastS”); the mean interval was shorter but the last interval was longer (“MeanSLastL”); the mean was longer but the last interval was shorter (“MeanLLastS”); and both the mean interval and the last interval were longer (“MeanLLastL”). The onset of the two visual Ternus frames (30 ms) was accompanied by a (30-ms) auditory beep (i.e.,  $ISI_V = ISI_A$ ).

In Experiment 4 we compared two auditory sequences: one with its geometric mean 70 ms shorter than the transition threshold of the visual Ternus motion (hereafter the “Short” condition), and the other with its geometric mean 70 ms longer than the transitional threshold (hereafter the “Long” condition). Instead of randomization of the five auditory intervals (excluding the final synchronous auditory interval with the visual Ternus interval), the last auditory interval before the onset of Ternus display was fixed at the “transitional threshold” for both sequences. The rest four intervals were chosen randomly such that the coefficient of variance (CV) of the auditory sequence was in the range between 0.1 and 0.2, which is the normal range of CV for human observers (Allan, 1974; Getty, 1975; Penney et al., 2000). By this manipulation, we expected to minimize the influence of the potential recency effect caused by the last auditory interval. The audiovisual Ternus frames were appended at the end of these sequences for 85.7% trials (with 672 trials out of 784 trials), in which the Ternus display appeared at the end of the sound sequence (the “onset” of first visual frame was synchronized with 6th beep). The remaining were 112 catch trials, in which 56 trials had the Ternus displays at the beginning of the sound sequence (i.e., the “onset” of the first visual frame was synchronized with the second beep), and the rest 56 trials at middle temporal locations (i.e., the “onset” of the first visual frame was synchronized with the 4th beep). Those catch trials were used to avoid potential anticipatory attending to the visual events appearing at the end of the sound sequence. The total 784 trials were randomized and organized in 14 blocks, with each of 56 trials.

In Experiment 5, we used three types of auditory sequences, in which the mean auditory interval was either *shorter than*, *equal to* or *longer than* the individual transitional threshold of Ternus motion. The auditory sequence consisted of 8 to 10 beeps, including those accompanying the two visual Ternus frames, with the latter being inserted mainly at the 6th–7th positions (504 trials), and followed by 0–2 beeps (number selected at random), to minimize expectations for the onset of the visual Ternus frames. Two of the beeps (the 6th and the 7th) were synchronously paired with two visual Ternus frames which were separated by a visual ISI ( $ISI_V$ ) that varied from 50 to 230 ms (for the critical beeps,  $ISI_V = ISI_A$ ). There were up to two oddball tones (500 Hz) in the sound sequence, while the remaining regular sounds were of 1,000 Hz (including the two beeps synchronous with the two visual frames). Participants completed a dual-task in which they not only made discriminations of the

Ternus display (“element motion” vs. “group motion”) but also reported the number of oddball sounds (0–2) (Figure 1).

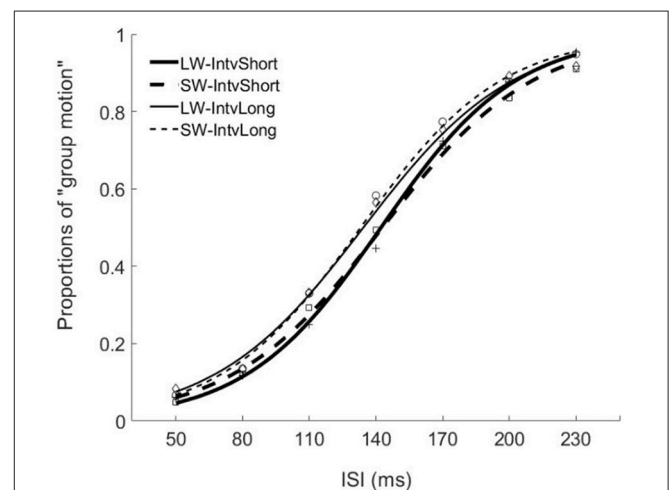
## RESULTS

### Experiment 1: The Effect of Short Temporal Window (With a Temporal Gap Between Auditory Sequence and Visual Ternus) vs. Long Temporal Window

The PSEs for the short window and long window were 149.4 ( $\pm 5.6$ , standard error) ms and 141.2 ( $\pm 4.8$ ) ms. The main effect of temporal window was significant,  $F_{(1, 13)} = 6.878$ ,  $p = 0.021$ ,  $\eta_g^2 = 0.346$ . The PSEs for short interval and long interval were 145.5 ( $\pm 5.2$ ) ms and 145.0 ( $\pm 4.8$ ) ms, the main effect of mean interval was not significant,  $F_{(1, 13)} = 0.120$ ,  $p = 0.735$ ,  $\eta_g^2 = 0.009$ . The interaction effect between factors of window and interval was not significant,  $F_{(1, 13)} = 1.033$ ,  $p = 0.328$ ,  $\eta_g^2 = 0.074$ . For the JNDs, both the main effects of temporal window and mean interval were not significant,  $F_{(1, 13)} = 3.419$ ,  $p = 0.087$ ,  $\eta_g^2 = 0.208$  and  $F_{(1, 13)} = 0.089$ ,  $p = 0.770$ ,  $\eta_g^2 = 0.007$ . And the interaction effect between the two factors was not significant,  $F_{(1, 13)} = 2.863$ ,  $p = 0.114$ ,  $\eta_g^2 = 0.180$  (Figures 2, 4).

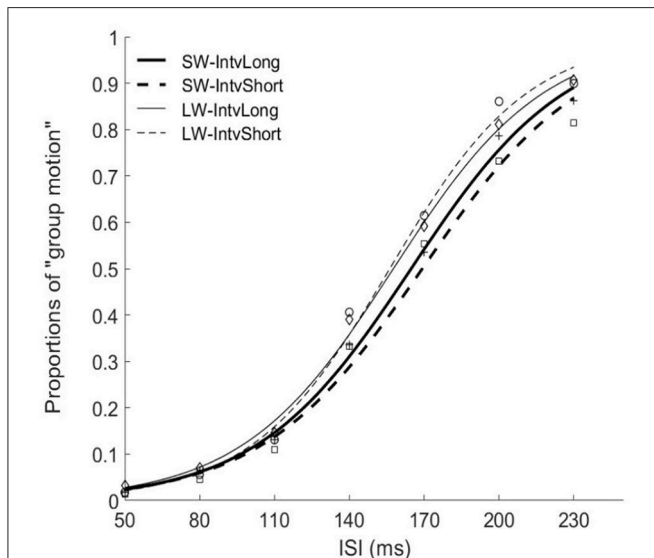
### Experiment 2: The Effect of Short Temporal Window (Without a Gap Between Auditory Sequence and Visual Ternus) vs. Long Temporal Window

The PSEs for the short window and long window were 168.7 ( $\pm 6.2$ ) ms and 156.2 ( $\pm 5.7$ ). The PSE for short window was larger



**FIGURE 2 |** Psychometric curves for Experiment 1. Mean proportions of group-motion responses were plotted as a function of the probe visual interval ( $ISI_V$ ), and fitted psychometric curves, were plotted for the auditory sequences with the different lengths of temporal windows and with different (geometric) mean intervals relative to the individual transition thresholds. SW-IntvLong, Short window with long mean auditory inter-interval; SW-IntvShort, Short window with short mean auditory inter-interval; LW-IntvLong, Long window with long mean auditory inter-interval. LW-IntvShort, long window with short mean auditory inter-interval.

than the one in long window,  $F_{(1, 12)} = 20.860$ ,  $p = 0.001$ ,  $\eta_g^2 = 0.635$ . The PSEs for short interval and long interval were  $163.8 (\pm 6.0)$  ms and  $161.0 (\pm 5.8)$ , the main effect of mean interval



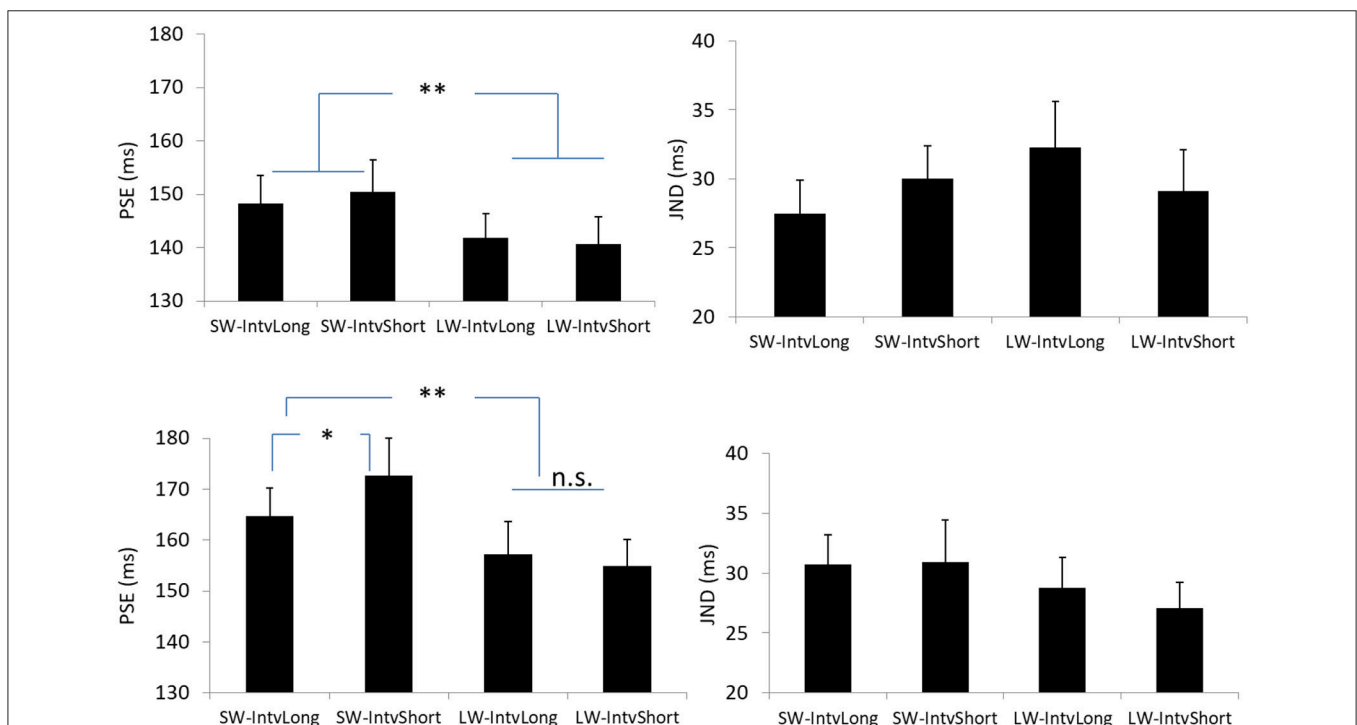
**FIGURE 3 |** Psychometric curves for Experiment 2. SW-IntvLong, Short window with long mean auditory inter-interval; SW-IntvShort, Short window with short mean auditory inter-interval; LW-IntvLong, Long window with long mean auditory inter-interval. LW-IntvShort, long window with short mean auditory inter-interval.

was not significant,  $F_{(1, 12)} = 1.869$ ,  $p = 0.197$ ,  $\eta_g^2 = 0.135$ . Importantly, the interaction effect between factors of window and interval was significant,  $F_{(1, 12)} = 5.090$ ,  $p = 0.044$ ,  $\eta_g^2 = 0.298$ . Further simple effect analyses showed that for short interval, the PSE in short window ( $172.7 \pm 7.3$  ms) was larger than the one ( $154.9 \pm 5.3$  ms) in long window,  $p = 0.001$ . For long interval, the PSE in short window ( $164.7 \pm 5.5$  ms) was larger than the one ( $157.3 \pm 6.4$  ms) in long window,  $p = 0.034$ . On the other hand, for the short window, the PSE in short interval ( $172.7 \pm 7.3$  ms) was larger than the one in long interval ( $164.7 \pm 5.5$  ms),  $p = 0.044$ . However, for the long window, the PSEs are equal in both intervals ( $154.9$  vs.  $157.3$  ms for short and long intervals),  $p = 0.377$ .

For the JNDs, both the main effects of temporal window and mean interval were not significant [ $F_{(1, 12)} = 2.479$ ,  $p = 0.141$ ,  $\eta_g^2 = 0.171$  and  $F_{(1, 12)} = 0.282$ ,  $p = 0.605$ ,  $\eta_g^2 = 0.023$ ]. The interaction effect between the two factors was not significant,  $F_{(1, 12)} = 0.408$ ,  $p = 0.535$ ,  $\eta_g^2 = 0.033$  (Figures 3, 4).

### Experiment 3: Central Tendency Effect vs. Last Interval

The PSEs for the short mean interval and long mean interval were  $143.2 (\pm 7.4)$  ms and  $135.3 (\pm 9.5)$ . The main effect of mean interval was significant,  $F_{(1, 6)} = 9.070$ ,  $p = 0.024$ ,  $\eta_g^2 = 0.602$ . The PSEs for short last interval and long last interval were  $155.8 (\pm 9.7)$  ms and  $122.6 (\pm 7.5)$  ms, respectively. The main effect of last interval was significant,  $F_{(1, 6)} = 65.970$ ,  $p = 0.000$ ,  $\eta_g^2 = 0.917$ . The interaction effect between factors of mean interval



**FIGURE 4 |** Plotted bars for PSE (point of subjective equality) and JND (just noticeable difference) for Experiments 1 (Upper) and 2 (Down). \* $p < 0.05$ , \*\* $p < 0.01$ , n.s. not significant.

and last interval was not significant,  $F_{(1,6)} = 0.195$ ,  $p = 0.674$ ,  $\eta_g^2 = 0.031$ . For the JNDs, the JND in short last interval ( $24.8 \pm 1.3$  ms) was larger than the one in long last interval ( $21.5 \pm 1.6$  ms),  $F_{(1,6)} = 11.590$ ,  $p = 0.014$ ,  $\eta_g^2 = 0.659$ . However, the main effect of mean interval was not significant,  $F_{(1,6)} = 0.762$ ,  $p = 0.416$ ,  $\eta_g^2 = 0.113$ . The interaction effect between the two factors was also not significant,  $F_{(1,6)} = 0.109$ ,  $p = 0.753$ ,  $\eta_g^2 = 0.018$ . (Figures 5, 6).

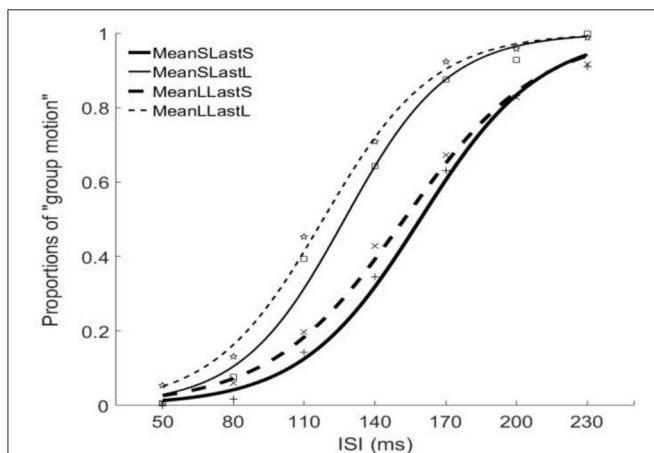
#### Experiment 4: Central Tendency Effect but With the Last Interval Fixed

Here we made formal manipulation by keeping the last interval fixed for the “Short” and “Long” auditory sequences. Figure 7 depicts the responses from a typical participant. The PSEs were  $153.1 (\pm 7.3)$ ,  $137.9 (\pm 9.1)$  for the “Short” and “Long” conditions,  $t_{(11)} = 3.640$ ,  $p < 0.01$ . Participants perceived more dominant percept of Element motion in the “Short” condition than in the “Long” condition, consistent with the findings of the

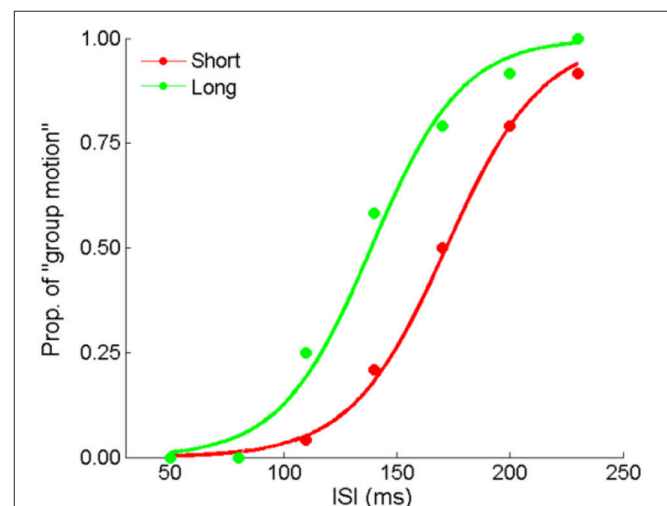
previous experiments. That is, the auditory ensemble mean still assimilated visual Ternus apparent motion when the last interval of the auditory sequence was fixed. Therefore, the audiovisual interactions we found were unlikely only due to the recency effect.

#### Experiment 5: Central Tendency Effect With Attentional Modulation

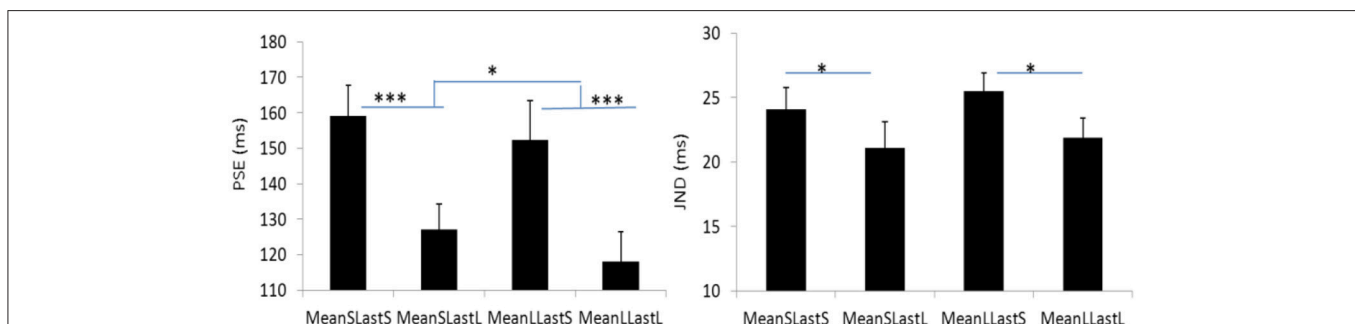
The PSEs for the *baseline*, *short*, *equal*, and *long* intervals were  $135.9 (\pm 3.3)$ ,  $171.1 (\pm 8.9)$ ,  $151.5 (\pm 9.5)$ , and  $142.1 (\pm 7.4)$  ms, the main effect of mean interval was significant,  $F_{(2,39)} = 9.020$ ,  $p < 0.001$ ,  $\eta_g^2 = 0.410$ . Bonferroni corrected comparison showed that the PSE for baseline was smaller than the one in short condition,  $p = 0.014$ . PSE for short interval condition was larger than the one in equal condition,  $p = 0.01$ ; and the PSE for short interval was also larger than the ones in the equal and long intervals,  $p = 0.019$  and  $p = 0.010$ . However, the PSEs were equal for both



**FIGURE 5** | Psychometric curves for Experiment 3. MeanSLastS (bold solid line), Mean short interval with long last auditory interval; MeanSLastL (thin solid line), Mean short interval with short last auditory interval; MeanLLastS (bold dashed line), Mean long interval with short last auditory interval; MeanLLastL (thin dashed line), Mean long interval with long last auditory interval.



**FIGURE 7** | Mean proportions of group-motion responses from a typical participant are plotted against the probe visual interval (ISlv), and fitted psychometric curves for the two geometric mean conditions: the “Short” sequence (with the smaller geometric mean) and “Long” sequence (with the larger geometric mean) in Experiment 4.



**FIGURE 6** | Plotted bars for PSE (point of subjective equality) and JND (just noticeable difference) for Experiment 3. \* $p < 0.05$ , \*\* $p < 0.01$ , and \*\*\* $p < 0.001$ .



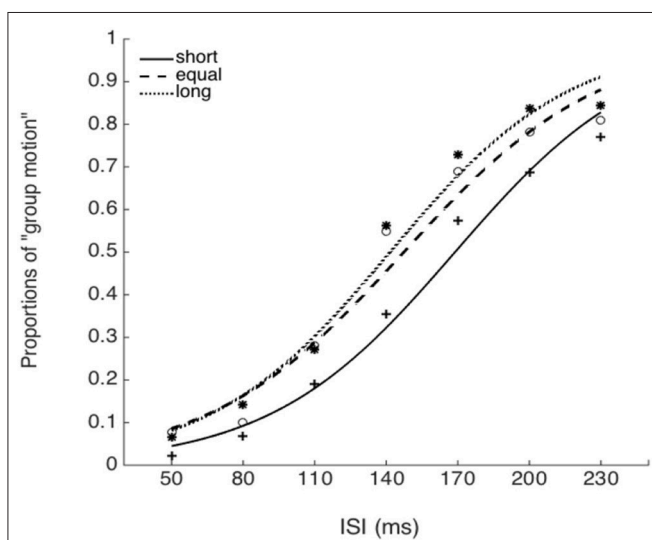
“equal” and “long” conditions,  $p = 0.411$ . The PSEs were equal for both “baseline” and “equal” condition,  $p = 0.603$ , and were equal between “baseline” and “long” conditions,  $p = 1$ .

The JNDs for the *baseline*, *short*, *equal*, and *long* intervals were  $32.2 (\pm 3.7)$ ,  $39.3 (\pm 5.1)$ ,  $44.9 (\pm 7.0)$ , and  $40.0 (\pm 4.4)$  ms, respectively. The main effect of mean interval was not significant,  $F_{(3, 39)} = 2.741$ ,  $p = 0.056$ ,  $\eta_g^2 = 0.174$  (Figures 8, 9).

The mean correct rate for reporting the number of oddball sounds was  $83.0 \pm 3.1\%$ , one sample  $T$ -test with comparison of 50% showed the correct rate was above the chance level [ $t_{(13)} = 10.518$ ,  $p = 9.984 \times 10^{-8}$ ].

## DISCUSSION

Central tendency, the tendency of judgments of quantitative properties (lengths, durations etc) for given stimuli to gravitate toward their mean, is one of the most robust perceptual effects. The present study has shown that perceptual averaging of temporal property—auditory intervals, assimilates the visual interval between the two Ternus-display frames, and biases the perception of Ternus apparent motion (either to be dominant “element motion” or dominant “group motion”). This finding is consistent with the large body of literature on temporal-context and central tendency effects, within the broader framework of Bayesian optimization (Jazayeri and Shadlen, 2010; Shi et al., 2013; Roach et al., 2017), whereby incorporating the mean of the statistical distribution in the estimation would assimilate the estimates toward the mean—known as “central tendency effect” (Jazayeri and Shadlen, 2010; Burr et al., 2013; Karaminis et al., 2016).



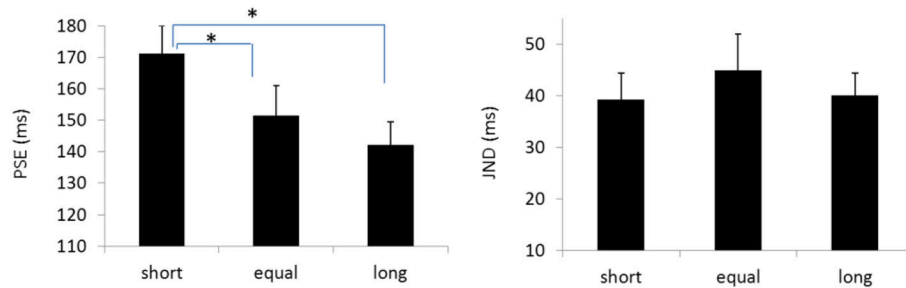
**FIGURE 8 |** Psychometric curves for Experiment 5. Short (solid line), the mean auditory inter-interval is shorter than the PSE for visual Ternus motion; Equal (dashed line), the mean auditory inter-interval is equal to the PSE for visual Ternus motion; Long (dotted line), the mean auditory inter-interval is longer than the PSE for visual Ternus motion. The PSE (“transitional threshold”) of Ternus motion was established by a pre-test for each individual.

By using the paradigm of temporal ventriloquism and the probe of visual Ternus display (Chen et al., 2010; Shi et al., 2010; Chen and Vroomen, 2013), we have previously shown that the auditory capture effect upon the visual events, in which the perceived visual interval was biased by concurrently presented auditory events. Observers tended to report the illusory visual (apparent motion) percepts with the concurrent presence of auditory beeps. However, the visual-auditory integration effect is subject to the temporal reference, i.e., the time interval between the critical visual probe and the sound sequence, the mean auditory interval and the critical interval between the last auditory stimulus and the onset of visual events. In our current setting, when the total time interval between the onset of auditory signal and the onset of visual events was above 3 s (3.2 s), it gave rise to a diminished central tendency effect. On the contrary, when this time interval was less than 2.4 s, the shortened time reference increased the likelihood of central tendency effect—materialized in the effect of “geometric” perceptual averaging for auditory intervals upon the visual Ternus motion. These findings indicate a general temporal framework of crossmodal integration. As stated in a theoretical construct of temporal perception, known as the “subjective present”—a mechanism of temporal integration binds successive events into perceptual units of 3 s duration (Pöppel, 1997). Such a temporal integration, which is automatic and pre-semantic, is also operative in movement control and other cognitive activities. In this hierarchical temporal model, the temporal reference for temporal binding could be extended but limited within 3 s, together with a memory store (Pöppel, 1997; Pöppel and Bao, 2014). When the framework exceeds 3 s, the integration of the preceding auditory interval information could be decayed, which hence makes the auditory assimilation effect reduced.

Interestingly, even with the presumed short temporal window (within 2.4 s), by inserting a short temporal gap (150 ms) between the offset of the very last beep and the onset of the first visual frame, we found the central tendency effect was reduced, and the effect was similar to the results in long temporal window condition (3.2 s). This finding suggests that the “imminent” and most recent (“immediate”) temporal gap before the target visual event is critical for the development of the central tendency effect. This inference is further substantiated by the results from Experiments 2 and 3. In Experiment 2, with the configuration of “short window,” we eliminated the short gap (150 ms) between the offset of the last beep and the onset of the visual frames. We found that the central tendency effect (short mean interval. vs. long mean interval) reappeared, though it still remains absent in the condition of “long window.” Moreover, in Experiment 3, we further found that the assimilation effect of the last interval dominates that of the mean auditory interval. This indicates that the last auditory interval wins the competition over the mean interval in driving the crossmodal assimilation.

However, the central tendency effect was less dependent on attentional modulation. Using the dual-tasks of reporting the percept of visual Ternus motion and the number of oddball stimuli [i.e., identifying the number of 500 Hz beep(s) within a sound sequence], we again found the central tendency effect was robust. The observers have invested large attentional resources to





**FIGURE 9 |** Plotted bars for PSE (point of subjective equality) and JND (just noticeable difference) for Experiments 5, \* $p < 0.05$ .

obtain the decent performance of counting the oddball sounds. Nevertheless, the performance of crossmodal assimilation effect still survived. Therefore, the central tendency effect as shown in the present study, has demonstrated its automatic and attentional-less demanding nature during crossmodal interaction (Vroomen et al., 2001; Wahn and König, 2015).

The current study has some limitations. Indeed, the temporal reference before the target visual Ternus display includes intervals composed by stimuli with different configurations. The auditory sequence was organized by filled-durations with multiple beeps, and there was a transition of intra-modal perceptual grouping (with sounds) to cross-modal grouping when the last beep was followed by the onset of the first visual Ternus frame (with audiovisual events) (Burr et al., 2013). However, the “critical” time window for multisensory integration was presented as an “empty interval” between the two visual frames. Therefore, the visual probe we adopted in current experimental paradigm might restrict the manifestation of assimilation effect, which was probably due to the differential timing sensitivities to the “filled-duration” in auditory sequence vs. “empty-duration” in the visual probe (Rammsayer and Lima, 1991; Grondin, 1993; Rammsayer, 2010). Moreover, the temporal window, as shown in the auditory sequence, covaried with the mean ISIs (mean auditory intervals). This potential confound remains even although we have manipulated the comparisons of durations between the mean ISIs and the critical interval between the two visual frames (Experiments 1, 2, 3, and 5), and tried to tease apart the “central tendency effect” vs. “recency effect” by fixing the last intervals. Further research is needed to elucidate this point.

Taken together, the current study has shown that crossmodal assimilation in temporal domain is shaped by the temporal reference, in which the observers use the temporal information by dynamically averaging the intervals (as they unfold in time sequence) and exploiting the last interval before the target events. The central tendency effect in temporal domain, similar to the central effect associated with other sensory properties

such as weights and hues, is adaptively subject to the frame of reference (Hollingworth, 1910; Helson, 1947, 1948; Helson and Himelstein, 1955; Sherif et al., 1958; Thomas and Jones, 1962; Helson and Avant, 1967; Thomas et al., 1973; Hébert et al., 1974; Thomas and Strub, 1974; Newlin et al., 1978; Burr et al., 2013; Karaminis et al., 2016). Importantly, the temporal information near the target event is critical for crossmodal assimilation, wherein the recency effect prevails over the central tendency effect during the assimilation process (Burr et al., 2013; Karaminis et al., 2016). Crossmodal assimilation is more dependent on the temporal duration which entails the integration of task-relevant (temporal) information to be efficient within a short window (3 s) in addition to efficient working memory functions (Pöppel, 1997; Block and Gruber, 2014; Pöppel and Bao, 2014). However, the crossmodal assimilation is less subject to another process—attentional modulation (Talsma et al., 2010).

## AUTHOR CONTRIBUTIONS

YW conducted Experiment 1 and analyzed data. LC conducted Experiments 2–4, analyzed data and wrote the manuscript.

## ACKNOWLEDGMENTS

This work is funded by the Natural Science Foundation of China (NSFC61527804, 81371206) and was partially funded by NSFC and the German Research Foundation (DFG) in Project Crossmodal Learning, NSFC 61621136008 / DFG TRR-169.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2018.00039/full#supplementary-material>

**Supplement 1:** Demo of Ternus Display.

## 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
- Allan, L. G. (1974). Psychophysical theories of duration discrimination. *Percept. Psychophys.* 16, 26–34. doi: 10.3758/BF03203244
- Alsius, A., Mottonen, R., Sams, M. E., Soto-Faraco, S., and Tiippana, K. (2014). Effect of attentional load on audiovisual speech perception: evidence from ERPs. *Front. Psychol.* 5:727. doi: 10.3389/fpsyg.2014.00727
- Alsius, A., Navarra, J., Campbell, R., and Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Curr. Biol.* 15, 839–843. doi: 10.1016/j.cub.2005.03.046
- Block, R. A., and Gruber, R. P. (2014). Time perception, attention, and memory: a selective review. *Acta Psychol.* 149, 129–133. doi: 10.1016/j.actpsy.2013.11.003
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Bruns, P., and Röder, B. (2015). Sensory recalibration integrates information from the immediate and the cumulative past. *Sci. Rep.* 5:12739. doi: 10.1038/srep12739
- Buchan, J. N., and Munhall, K. G. (2011). The influence of selective attention to auditory and visual speech on the integration of audiovisual speech information. *Perception* 40, 1164–1182. doi: 10.1068/p6939
- Burr, D., Della Rocca, E., and Morrone, M. C. (2013). Contextual effects in interval-duration judgements in vision, audition and touch. *Exp. Brain Res.* 230, 87–98. doi: 10.1007/s00221-013-3632-z
- Cardinal, L. J. (2015). Central tendency and variability in biological systems: part 2. *J. Commun. Hosp. Intern. Med. Perspect.* 5:28972. doi: 10.3402/jchimp.v5.28972
- Chen, L., Shi, Z., and Müller, H. J. (2010). Influences of intra- and crossmodal grouping on visual and tactile Ternus apparent motion. *Brain Res.* 1354, 152–162. doi: 10.1016/j.brainres.2010.07.064
- Chen, L., and Vroomen, J. (2013). Intersensory binding across space and time: a tutorial review. *Atten. Percept. Psychophys.* 75, 790–811. doi: 10.3758/s13414-013-0475-4
- Chen, L., Zhang, M., Ai, F., Xie, W., and Meng, X. (2016). Crossmodal synesthetic congruency improves visual timing in dyslexic children. *Res. Dev. Disabil.* 55, 14–26. doi: 10.1016/j.ridd.2016.03.010
- Chen, Y. C., and Spence, C. (2017). Assessing the role of the ‘unity assumption’ on multisensory integration: a review. *Front. Psychol.* 8:445. doi: 10.3389/fpsyg.2017.00445
- Chuen, L., and Schutz, M. (2016). The unity assumption facilitates cross-modal binding of musical, non-speech stimuli: the role of spectral and amplitude envelope cues. *Atten. Percept. Psychophys.* 78, 1512–1528. doi: 10.3758/s13414-016-1088-5
- Cohen, M. A., Dennett, D. C., and Kanwisher, N. (2016). What is the bandwidth of perceptual experience? *Trends Cogn. Sci.* 20, 324–335. doi: 10.1016/j.tics.2016.03.006
- Cowan, N. (2001). Metatheory of storage capacity limits. *Behav. Brain Sci.* 24, 154–176. doi: 10.1017/S0140525X0161392X
- Donohue, S. E., Green, J. J., and Woldorff, M. G. (2015). The effects of attention on the temporal integration of multisensory stimuli. *Front. Integr. Neurosci.* 9:32. doi: 10.3389/fnint.2015.00032
- Donohue, S. E., Roberts, K. C., Grent-’t-Jong, T., and Woldorff, M. G. (2011). The cross-modal spread of attention reveals differential constraints for the temporal and spatial linking of visual and auditory stimulus events. *J. Neurosci.* 31, 7982–7990. doi: 10.1523/JNEUROSCI.5298-10.2011
- Duncan, J., Martens, S., and Ward, R. (1997). Restricted attentional capacity within but not between sensory modalities. *Nature* 387, 808–810.
- Ernst, M., and Di Luca, M. (2011). “Multisensory perception: from integration to remapping,” in *Sensory Cue Integration*, ed J. Trommershäuser (New York, NY: Oxford University Press), 225–250.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787
- Getty, D. J. (1975). Discrimination of short temporal intervals: a comparison of two models. *Percept. Psychophys.* 18, 1–8. doi: 10.3758/BF03199358
- Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Percept. Psychophys.* 54, 383–394. doi: 10.3758/BF03205274
- Gupta, D. S., and Chen, L. (2016). Brain oscillations in perception, timing and action. *Curr. Opin. Behav. Sci.* 8, 161–166. doi: 10.1016/j.cobeha.2016.02.021
- Habets, B., Bruns, P., and Roder, B. (2017). Experience with crossmodal statistics reduces the sensitivity for audio-visual temporal asynchrony. *Sci. Rep.* 7:1486. doi: 10.1038/s41598-017-01252-y
- Hébert, J. A., Bullock, M., Levitt, L., Woodward, K. G., and McGuirk, F. D. (1974). Context and frequency effects in generalization of a human voluntary response. *J. Exp. Psychol.* 102, 456–462. doi: 10.1037/h0035877
- Helson, H. (1947). Adaptation-level as frame of reference for prediction of psychophysical data. *Am. J. Psychol.* 60, 1–29. doi: 10.2307/1417326
- Helson, H. (1948). Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.* 55, 297–313. doi: 10.1037/h0056721
- Helson, H., and Avant, L. L. (1967). Stimulus generalization as a function of contextual stimuli. *J. Exp. Psychol.* 73, 565–567. doi: 10.1037/h0024319
- Helson, H., and Himelstein, P. (1955). A short method for calculating the adaptation-level for absolute and comparative rating judgments. *Am. J. Psychol.* 68, 631–637. doi: 10.2307/1418791
- Hillock, A. R., Powers, A. R., and Wallace, M. T. (2011). Binding of sights and sounds: age-related changes in multisensory temporal processing. *Neuropsychologia* 49, 461–467. doi: 10.1016/j.neuropsychologia.2010.11.041
- Hillock-Dunn, A., Grantham, D. W., and Wallace, M. T. (2016). The temporal binding window for audiovisual speech: children are like little adults. *Neuropsychologia* 88, 74–82. doi: 10.1016/j.neuropsychologia.2016.02.017
- Hollingworth, H. L. (1910). The central tendency of judgment. *J. Philos. Psychol. Sci. Methods* 7, 461–469. doi: 10.2307/2012819
- Jazayeri, M., and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nat. Neurosci.* 13, 1020–1026. doi: 10.1038/nn.2590
- Jones, M. R., and McAuley, J. D. (2005). Time judgments in global temporal contexts. *Percept. Psychophys.* 67, 398–417. doi: 10.3758/BF03193320
- Karaminis, T., Cicchini, G. M., Neil, L., Cappagli, G., Aagten-Murphy, D., Burr, D., et al. (2016). Central tendency effects in time interval reproduction in autism. *Sci. Rep.* 6:28570. doi: 10.1038/srep28570
- Keetels, M., and Vroomen, J. (2008). Tactile-visual temporal ventriloquism: no effect of spatial disparity. *Percept. Psychophys.* 70, 765–771. doi: 10.3758/PP.70.5.765
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., and Broussard, C. (2007). What’s new in psychtoolbox-3. *Perception* 36, 1–16. doi: 10.1068/v070821
- Klemen, J., Buchel, C., and Rose, M. (2009). Perceptual load interacts with stimulus processing across sensory modalities. *Eur. J. Neurosci.* 29, 2426–2434. doi: 10.1111/j.1460-9568.2009.06774.x
- Klemen, J., and Chambers, C. D. (2012). Current perspectives and methods in studying neural mechanisms of multisensory interactions. *Neurosci. Biobehav. Rev.* 36, 111–133. doi: 10.1016/j.neubiorev.2011.04.015
- Lewkowicz, D. J., and Flom, R. (2014). The audiovisual temporal binding window narrows in early childhood. *Child Dev.* 85, 685–694. doi: 10.1111/cdev.12142
- Li, Y., Long, J., Huang, B., Yu, T., Wu, W., Li, P., et al. (2016). Selective audiovisual semantic integration enabled by feature-selective attention. *Sci. Rep.* 6:18914. doi: 10.1038/srep18914
- Mégevand, P., Molholm, S., Nayak, A., and Foxe, J. J. (2013). Recalibration of the multisensory temporal window of integration results from changing task demands. *PLoS ONE* 8:e71608. doi: 10.1371/journal.pone.0071608
- Meredith, M. A., Nemitz, J. W., and Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal Factors. *J. Neurosci.* 7, 3215–3229.
- Miller, J. E., Carlson, L. A., and McAuley, J. D. (2013). When what you hear influences when you see: listening to an auditory rhythm influences the temporal allocation of visual attention. *Psychol. Sci.* 24, 11–18. doi: 10.1177/0956797612446707
- Misceo, G. F., and Taylor, N. J. (2011). A “unity assumption” does not promote intersensory integration. *Exp. Psychol.* 58, 385–390. doi: 10.1027/1618-3169/a000107
- Nakajima, Y., ten Hoopen, G., Hilkuysen, G., and Sasaki, T. (1992). Time-shrinking: a discontinuity in the perception of auditory temporal patterns. *Percept. Psychophys.* 51, 504–507. doi: 10.3758/BF03211646
- Newlin, R. J., Rodgers, J. P., Dickson, J. F., Strub, H., and Thomas, D. R. (1978). The central tendency effect in stimulus generalization: effects of establishing

- a “preexperimental” frame of reference. *Percept. Psychophys.* 24, 161–167. doi: 10.3758/BF03199543
- Parise, C. V., Spence, C., and Ernst, M. O. (2012). When correlation implies causation in multisensory integration. *Curr. Biol.* 22, 46–49. doi: 10.1016/j.cub.2011.11.039
- Penney, T. B., Gibbon, J., and Meck, W. H. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 1770–1787. doi: 10.1037/0096-1523.26.6.1770
- 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., and Bao, Y. (2014). “Temporal windows as a bridge from objective to subjective time,” in *Subjective Time*, eds D. Lloyd and V. Arstila (Cambridge, MA: MIT Press), 241–261.
- Powers, A. R., Hevey, M. A., and Wallace, M. T. (2012). Neural correlates of multisensory perceptual learning. *J. Neurosci.* 32, 6263–6274. doi: 10.1523/JNEUROSCI.6138-11.2012
- 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
- Rammsayer, T. H. (2010). Differences in duration discrimination of filled and empty auditory intervals as a function of base duration. *Atten. Percept. Psychophys.* 72, 1591–1600. doi: 10.3758/APP.72.6.1591
- Rammsayer, T. H., and Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: cognitive and perceptual factors. *Percept. Psychophys.* 50, 565–574. doi: 10.3758/BF03207541
- Roach, N. W., McGraw, P. V., Whitaker, D. J., and Heron, J. (2017). Generalization of prior information for rapid Bayesian time estimation. *Proc. Natl. Acad. Sci. U.S.A.* 114, 412–417. doi: 10.1073/pnas.1610706114
- Sherif, M., Taub, D., and Hovland, C. I. (1958). Assimilation and contrast effects of anchoring stimuli on judgments. *J. Exp. Psychol.* 55, 150–155. doi: 10.1037/h0048784
- Shi, Z., and Burr, D. (2016). Predictive coding of multisensory timing. *Curr. Opin. Behav. Sci.* 8, 200–206. doi: 10.1016/j.cobeha.2016.02.014
- Shi, Z., Chen, L., and Muller, H. J. (2010). Auditory temporal modulation of the visual Ternus effect: the influence of time interval. *Exp. Brain Res.* 203, 723–735. doi: 10.1007/s00221-010-2286-3
- Shi, Z., Church, R. M., and Meck, W. H. (2013). Bayesian optimization of time perception. *Trends Cogn. Sci.* 17, 556–564. doi: 10.1016/j.tics.2013.09.009
- Stevenson, R. A., Siemann, J. K., Schneider, B. C., Eberly, H. E., Woynarowski, T. G., Camarata, S. M., et al. (2014). Multisensory temporal integration in autism spectrum disorders. *J. Neurosci.* 34, 691–697. doi: 10.1523/JNEUROSCI.3615-13.2014
- Stevenson, R. A., Zemtsov, R. K., and Wallace, M. T. (2012). Individual differences in the multisensory temporal binding window predict susceptibility to audiovisual illusions. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 1517–1529. doi: 10.1037/a0027339
- Sugano, Y., Keetels, M., and Vroomen, J. (2010). Adaptation to motor-visual and motor-auditory temporal lags transfer across modalities. *Exp. Brain Res.* 201, 393–399. doi: 10.1007/s00221-009-2047-3
- Sugano, Y., Keetels, M., and Vroomen, J. (2012). The Build-up and transfer of sensorimotor temporal recalibration measured via a synchronization task. *Front. Psychol.* 3:246. doi: 10.3389/fpsyg.2012.00246
- Sugano, Y., Keetels, M., and Vroomen, J. (2016). Auditory dominance in motor-sensory temporal recalibration. *Exp. Brain Res.* 234, 1249–1262. doi: 10.1007/s00221-015-4497-0
- Talsma, D., Doty, T. J., and Woldorff, M. G. (2007). Selective attention and audiovisual integration: is attending to both modalities a prerequisite for early integration? *Cereb. Cortex* 17, 679–690. doi: 10.1093/cercor/bhk016
- Talsma, D., Senkowski, D., Soto-Faraco, S., and Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends Cogn. Sci.* 14, 400–410. doi: 10.1016/j.tics.2010.06.008
- Tang, X., Wu, J., and Shen, Y. (2016). The interactions of multisensory integration with endogenous and exogenous attention. *Neurosci. Biobehav. Rev.* 61, 208–224. doi: 10.1016/j.neubiorev.2015.11.002
- Ternus, J. (1926). Experimentelle untersuchungen über phänomenale identität. *Psychol. Forsch.* 7, 81–136. doi: 10.1007/BF02424350
- Thomas, D. R., and Jones, C. G. (1962). Stimulus generalization as a function of the frame of reference. *J. Exp. Psychol.* 64, 77–80. doi: 10.1037/h0043304
- Thomas, D. R., and Strub, H. (1974). Adaptation level and the central tendency effect in stimulus generalization. *J. Exp. Psychol.* 103, 466–474. doi: 10.1037/h0037184
- Thomas, D. R., Svinicki, M. D., and Vogt, J. (1973). Adaptation level as a factor in human discrimination learning and stimulus generalization. *J. Exp. Psychol.* 97, 210–219. doi: 10.1037/h0033908
- Treutwein, B., and Strasburger, H. (1999). Fitting the psychometric function. *Percept. Psychophys.* 61, 87–106. doi: 10.3758/BF03211951
- Vatakis, A., and Spence, C. (2007). Crossmodal binding: evaluating the “unity assumption” using audiovisual speech stimuli. *Percept. Psychophys.* 69, 744–756. doi: 10.3758/BF03193776
- Vatakis, A., and Spence, C. (2008). Evaluating the influence of the ‘unity assumption’ on the temporal perception of realistic audiovisual stimuli. *Acta Psychol.* 127, 12–23. doi: 10.1016/j.actpsy.2006.12.002
- Vroomen, J., Bertelson, P., and de Gelder, B. (2001). The ventriloquist effect does not depend on the direction of automatic visual attention. *Percept. Psychophys.* 63, 651–659. doi: 10.3758/BF03194427
- Vroomen, J., and Keetels, M. (2010). Perception of intersensory synchrony: a tutorial review. *Atten. Percept. Psychophys.* 72, 871–884. doi: 10.3758/APP.72.4.871
- Wahn, B., and König, P. (2015). Audition and vision share spatial attentional resources, yet attentional load does not disrupt audiovisual integration. *Front. Psychol.* 6:1084. doi: 10.3389/fpsyg.2015.01084
- Wallace, M. T., and Stevenson, R. A. (2014). The construct of the multisensory temporal binding window and its dysregulation in developmental disabilities. *Neuropsychologia* 64, 105–123. doi: 10.1016/j.neuropsychologia.2014.08.005
- 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

**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 © 2018 Wan and Chen. 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) and the copyright owner 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.



# Timing Deficits in ADHD: Insights From the Neuroscience of Musical Rhythm

Jessica L. Slater<sup>1\*</sup> and Matthew C. Tate<sup>1,2</sup>

<sup>1</sup> Department of Neurological Surgery, Northwestern University, Chicago, IL, United States, <sup>2</sup> Department of Neurology, Northwestern University, Chicago, IL, United States

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

Joachim Hass,  
Zentralinstitut für Seelische  
Gesundheit (ZI), Germany  
Michael H. Thaut,  
University of Toronto, Canada  
Maurizio Mattia,  
Istituto Superiore di Sanità, Italy

### \*Correspondence:

Jessica L. Slater  
j-slater@northwestern.edu

**Received:** 02 April 2018

**Accepted:** 18 June 2018

**Published:** 06 July 2018

### Citation:

Slater JL and Tate MC (2018) Timing Deficits in ADHD: Insights From the Neuroscience of Musical Rhythm. *Front. Comput. Neurosci.* 12:51. doi: 10.3389/fncom.2018.00051

Everyday human behavior relies upon extraordinary feats of coordination within the brain. In this perspective paper, we argue that the rich temporal structure of music provides an informative context in which to investigate how the brain coordinates its complex activities in time, and how that coordination can be disrupted. We bring insights from the neuroscience of musical rhythm to considerations of timing deficits in Attention Deficit/Hyperactivity Disorder (ADHD), highlighting the significant overlap between neural systems involved in processing musical rhythm and those implicated in ADHD. We suggest that timing deficits warrant closer investigation since they could lead to the identification of potentially informative phenotypes, tied to neurobiological and genetic factors. Our novel interdisciplinary approach builds upon recent trends in both fields of research: in the neuroscience of rhythm, an increasingly nuanced understanding of the specific contributions of neural systems to rhythm processing, and in ADHD, an increasing focus on differentiating phenotypes and identifying distinct etiological pathways associated with the disorder. Finally, we consider the impact of musical experience on rhythm processing and the potential value of musical rhythm in therapeutic interventions.

**Keywords:** music, rhythm, attention deficit hyperactivity disorder, ADHD, cognitive control, motor timing, neuroplasticity, musical expertise

## INTRODUCTION

Music is pervasive across cultures and plays an important role in human interaction, development and social bonding (Cross, 2001). The temporal structure of music is integral to its functions, and the experience of music relies upon a precisely-timed orchestration of activity across the brain's sensory, cognitive, motor, and reward systems. Musical rhythms inspire us to move (Keller and Rieger, 2009; Dalla Bella et al., 2013), and movement can, in turn, shape our perception of rhythmic patterns (Phillips-Silver and Trainor, 2005, 2007). Music also facilitates interpersonal synchrony, increasing pro-social behavior (Cirelli et al., 2012, 2014) and breaking down perceived barriers between self and other by coordinating shared emotional experiences (Tarr et al., 2014). Several studies suggest that interaction with music promotes synchronous neural activity not only across brain regions, but between the brains of individuals, for example during music listening (Abrams et al., 2013) and improvisation (Müller et al., 2013).

The rewarding qualities of music are also intrinsically linked to its temporal structure, through the creation and manipulation of expectations over time (Cooper and Meyer, 1960; Huron, 2006).



Within this temporal framework, the fulfillment and violation of expectations provides a rich palette of emotional expression, mediated by the reward transmitter dopamine (Schultz, 1998; Salimpoor and Zatorre, 2013). The rhythmic patterns found across a range of musical styles have been shown to exhibit an optimal balance of predictability and surprise, even in their written form (Levitin et al., 2012), and the subtle timing variations found in live musical performance further contribute to the emotional expression perceived by a listener (Repp, 1995; Palmer, 1997; Ashley, 2002; Bhatara et al., 2011). As these examples highlight, the influence of music on human experience is closely tied to its temporal structure and the coordinated neural activity it induces, both within and between individuals.

The dynamic interplay between predictive (top-down) and reactive (bottom-up) processing, exemplified in how the brain responds to musical rhythm, is also a necessary foundation for cognitive functions, such as attention (Engel et al., 2001; Raichle, 2010). For example, the ability to anticipate what is likely to happen next and streamline the allocation of neural resources accordingly must be balanced with the ability to respond to unexpected salient events in the environment. In disorders such as ADHD, this balance is disrupted, resulting in impaired attentional control and difficulties inhibiting irrelevant inputs. We have chosen to consider ADHD in particular because in addition to the core symptoms of inattention and/or hyperactivity/impulsivity, ADHD is also characterized by deficits in motor and perceptual timing (Smith et al., 2002; Fair et al., 2012; Zelaznik et al., 2012; Demers et al., 2013; Noreika et al., 2013). Recent studies have revealed rhythm-related deficits in ADHD (Hove et al., 2017; Puyjarinet et al., 2017), and much of the same neural infrastructure that supports the processing of musical rhythm is implicated in ADHD, from brain circuitry (Silk et al., 2009; Silberstein et al., 2016; Mueller et al., 2017) and neural dynamics (Başar and Güntekin, 2008; Mazaheri et al., 2014; Loo et al., 2017) to dopamine signaling, with leading genetic risk factors for ADHD including dopamine gene variants (Swanson et al., 2000; DiMaio et al., 2003). Here, we propose that insights from research on musical rhythm could offer a more nuanced understanding of timing deficits in ADHD, and potentially lead to the identification of informative phenotypes, linked to neurobiological and genetic factors.

## THE NEURAL INFRASTRUCTURE OF MUSICAL RHYTHM

In this section we highlight key components of the neural infrastructure involved in processing musical rhythm. Although this is by no means an exhaustive review, some basic definitions of terms may prove useful. We will use the term “rhythm” to refer to temporal patterns formed from sequences of durations or onsets, whereas “beat” refers to a periodic pulse. In a piece of music, the beat typically defines the basic unit of timing, and “meter” refers to the grouping of beats into a recurring pattern of stresses or accents, such as would differentiate the feel of a waltz vs. a march.

## Sensory-Motor Integration

Studies with non-human primates and even zebrafish have shown that neural ensembles can entrain to a rhythmic stimulus (Quintana and Fuster, 1999; Sumbre et al., 2008), and it is likely that human interaction with musical rhythm is founded upon these basic entrainment mechanisms. However, it is notable that the natural human tendency to move to music, for example by tapping a foot to the beat, has proven surprisingly elusive in the animal kingdom (Patel et al., 2009).

Imaging studies have revealed that in humans, rhythm perception is associated with activation not only in auditory cortices but in frontal, parietal and motor regions, including the supplementary motor area (SMA), basal ganglia and cerebellum (Grahn and Brett, 2007; Grahn, 2012; Large et al., 2015; Merchant et al., 2015). It has been suggested that the close sensory-motor coupling necessary for synchronization of movement to music may be unique to vocal learning species (including parrots and songbirds, as well as humans), in which it is a necessary basis for learning and producing complex communication signals (Patel and Iversen, 2014). Recent evidence of successful entrainment to the musical beat in non-vocal-learning species, for example a California sea lion (Cook et al., 2013), have cast doubt on this theory. Nonetheless, it is well established that close interaction between sensory and motor systems provides a sophisticated mechanism of temporal prediction and feedback (Schroeder et al., 2010), and that this plays an important role in how humans process musical rhythm.

The extensive activation of motor areas during rhythm perception, even in the absence of overt movement (Zatorre et al., 2007; Chen et al., 2008; Grahn and Rowe, 2009), is consistent with accumulating evidence that these systems serve a broader role in temporal processing and cognition. For example, fronto-striatal and fronto-cerebellar pathways are increasingly viewed as contributing to more general pattern-detection, predictive and cognitive functions (Akshoomoff and Courchesne, 1992; Graybiel, 1997; Schubotz, 2007). It has been proposed that striatal pathways are particularly involved in generating internal representations of beat and metrical structure (Grahn and Brett, 2007; Schwartz and Kotz, 2013). On the other hand, cerebellar circuits are more involved in the precise encoding of complex sequences, fast timing features and durations (Grube et al., 2010; Schwartz and Kotz, 2013). Together, these pathways create a system that can generate complex temporal predictions while also adapting to incoming information.

## Models of Rhythm Perception

In constructing computational models of rhythm perception, a major challenge is to capture not only the individual components of temporal processing that are involved, but how those mechanisms interact in real time to maintain the ongoing balance between predictive (top-down) and reactive (bottom-up) processing, discussed above (see McAuley, 2010; Grahn, 2012, for review). For example, several rule-based models have been proposed in which the regular beat and metrical structure inferred by a rhythmic pattern are maintained by an internal clock (Longuet-Higgins and Lee, 1982; Povel and Essens, 1985; Desain and Honing, 1999). However, these models do not

generally account for adaptive, online predictions and instead determine a “best fit” pattern of regular intervals based on the rhythm sequence as a whole (summarized in Grahn, 2012).

Models based on the entrainment of multiple oscillators have had greater success in accounting for online prediction that is tolerant to more complex rhythmic structure while remaining sensitive to natural variations in performance (Large and Kolen, 1994; Large and Palmer, 2002; Angelis et al., 2013). Indeed, there is evidence to suggest that natural, non-random patterns of timing variability (i.e., those exhibiting fractal scaling and long-range correlations) may actually improve the accuracy of listeners’ temporal predictions (Rankin et al., 2009, 2014), and this was also demonstrated by the model (Large and Palmer, 2002).

In their theory of neural resonance, Large and Snyder extend these computational models to propose that entrainment is performed in the brain by neural oscillators (Large and Snyder, 2009), and this theory is supported by evidence from imaging and EEG studies (Large and Snyder, 2009; Nozaradan et al., 2012; Tierney and Kraus, 2015). Interestingly, individual variation in the temporal characteristics of neural activity (including long-range correlations) has been shown to predict variability in motor timing behavior (Linkenkaer-Hansen et al., 2001; Smit et al., 2013). A recent paper also linked these temporal characteristics of neural activity to fluctuations in attention, and it was proposed that the typical increase in long-range correlations over the course of development may be delayed or disrupted in ADHD (Smit and Anokhin, 2017). This represents a fascinating area for future study, and a further potential link between ADHD and the temporal dynamics of brain and behavior.

Within entrainment models, different frequencies of neural oscillations serve distinct functions. For example, Large and Snyder suggest that bursts of high frequency oscillatory activity facilitate coordination across motor and sensory systems. Peaks in beta (13–30 Hz) and gamma (30–100 Hz) power were observed as an anticipatory response to rhythmic patterns (Snyder and Large, 2005; Fujioka et al., 2009), and persisted even when the sound stimulus stopped, supporting their role as self-sustaining timekeepers. Further, temporal modulations in beta activity were altered by the specific metrical structure imposed by the listener onto an ambiguous rhythm pattern, suggesting top-down modulation of oscillatory dynamics (Iversen et al., 2009). Given the association between beta oscillations and motor coordination, the modulation of beta power may provide another indication of the influence of motor systems on rhythm processing (Large et al., 2015).

Neural responses to musical rhythm may also take the form of entrainment to specific frequencies actually present in the stimulus, for example the frequency of the musical beat. Neural entrainment to the beat has been observed in a number of EEG studies in the form of increased spectral power at the frequency corresponding to the tempo of the musical beat, typically within the delta range (1–4 Hz), and even to harmonics and subharmonics of that frequency (Nozaradan et al., 2012; Tierney and Kraus, 2013, 2015; Nozaradan, 2014). The influence of motor systems on this form of neural beat entrainment was investigated in a recent lesion study (Nozaradan et al., 2017). Both cerebellar and basal ganglia patients showed reduced neural activity aligned

with the beat compared with controls, with cerebellar patients showing reductions specifically with faster tempo rhythms, and basal ganglia patients showing a greater deficit with complex rhythm patterns, which the authors interpreted as relying more heavily on the internal generation of a beat. These findings suggest that variation in cerebellar and striatal function (such as observed in ADHD) may be associated with distinct rhythm processing deficits. This study therefore provides compelling evidence for distinct specializations of these two motor areas in the coordination of neural entrainment to musical rhythm, linked with dissociable deficits.

## PARSING HETEROGENEITY IN ADHD: THE SEARCH FOR PHENOTYPES

ADHD is a highly prevalent and heterogenous disorder. Despite significant research efforts, characterization of the neurobiological basis of ADHD has proven elusive: diagnosis still relies heavily on self-report questionnaires, and treatment typically takes the form of a trial-and-error pharmacological approach. It has been difficult to identify biomarkers of the disorder because there has been no clear mapping between neural measures and clinical subtypes (i.e., predominantly inattentive, predominantly hyperactive/impulsive and combined type).

Although ADHD is associated with structural and functional abnormalities, including within frontal, striatal and cerebellar pathways, these findings have generally been small, and have not always been replicated (see Rubia, 2016, for review). Similarly, profiles of oscillatory dynamics have not been consistent enough to provide a clear neural “signature” of ADHD. EEG studies reveal abnormal patterns of oscillatory activity (Başar and Güntekin, 2008; Mazaheri et al., 2014; Loo et al., 2017), including reduced power in the beta frequency range. Indeed, a clinical diagnostic device assessing the ratio between theta and beta activity was developed and approved by the FDA (USDHHS, 2013). However, a subsequent meta-analysis suggested the theta-beta ratio is only elevated within a subgroup of individuals with ADHD, and is therefore not a reliable basis for diagnosis (Arns et al., 2013). A more nuanced understanding of distinct phenotypes of ADHD could help to increase diagnostic accuracy, and improve the development of clinical tools to aid in the evaluation and monitoring of treatment.

Research in the field is shifting toward the identification of distinct phenotypes and multiple etiologies (Castellanos and Tannock, 2002; Nigg et al., 2005; Durston et al., 2011). There is evidence from neuropsychological (Rommelse et al., 2008; Fair et al., 2012; Nikolas and Nigg, 2015), electrophysiological (Başar and Güntekin, 2008; Mazaheri et al., 2014; Loo et al., 2017) and genetic studies (Shaw et al., 2007; Giedd et al., 2008; Kebir and Joobor, 2011) to suggest the presence of distinct subgroups within ADHD, beyond the clinical subtypes. However, these subgroups have yet to be reconciled across methodologies to provide full characterization of etiological pathways.

Although motor and timing deficits are not included within the diagnostic criteria for ADHD, they are increasingly recognized as common symptoms (Toplak et al., 2006; Demers et al., 2013; Kaiser et al., 2015; Dahan et al., 2016), and have

been identified as a promising area for future study (Rubia, 2016). Consistent with the presence of multiple phenotypes, a recent study identifying rhythm deficits in children and adults with ADHD noted significant variation in performance within the ADHD group (Puyjarinet et al., 2017). Based on neuropsychological studies, it has been suggested that deficits in temporal information processing (e.g., duration discrimination) and increased response variability may represent distinct phenotypes, linked to dysfunction in cerebellar and basal ganglia pathways, respectively (Durstun et al., 2011; Fair et al., 2012). Given the distinct roles of fronto-cerebellar and fronto-striatal pathways in rhythm processing (Grahn and Brett, 2009; Grahn, 2012; Merchant et al., 2015; Nozaradan et al., 2017), including their separate influence on neural entrainment discussed in the previous section, we argue that further examination of rhythm-related deficits in ADHD could help to characterize phenotypes of ADHD, and to shed light on the different ways in which the dynamics within associated neural systems may be disrupted.

Further, genetic risk factors for ADHD include genes affecting dopaminergic transmission, which may influence timing behavior (Valera et al., 2010). This is supported by pharmacological studies in which timing deficits in ADHD are reduced by methylphenidate (which increases levels of dopamine) (Noreika et al., 2013) as well as a study in which dopamine manipulation in healthy controls was associated with impaired timing skills (Coull et al., 2012). As mentioned in the introduction, dopamine indexes temporal expectation within the context of musical rhythm. More broadly, dopamine supports neural communication within reward, motor and cognitive pathways and is involved in a wide range of functions including reward-based learning, motor coordination and cognitive control. It has been proposed that a common theme across its various functions is that dopamine coordinates neural systems to optimize responsiveness at different timescales, matching the timescales of activity in the environment (Schultz, 2007). In other words, dopamine helps to keep the brain “in sync” with the world around it. This is accomplished via multiple dopamine release mechanisms with distinct kinetic properties (Schultz, 2007). Therefore, we speculate that genetic variation in specific components of the dopaminergic system could lead to distinct deficits in neural and behavioral timing. This is consistent with evidence from animal studies, in which different genetic modifications affecting dopamine transmission in mice were associated with distinct behavioral timing deficits (Cevik, 2003; Drew et al., 2007; Balci et al., 2009, 2010), as well as evidence of dissociable timing deficits in humans linked to dopamine gene variants (Wiener et al., 2011).

Dopamine also helps to mediate the balance between inhibitory and excitatory neural activity that sustains neural oscillations, therefore genetic variations in dopaminergic signaling at different timescales may also influence temporal characteristics of oscillatory dynamics, such as the long-range correlations discussed above. Disrupted neural dynamics may in turn influence the development of cortical networks (Uhlhaas et al., 2010). Indeed, longitudinal studies have demonstrated distinct trajectories of structural brain development associated with different dopamine gene polymorphisms in ADHD (Shaw

TABLE 1 | How the infrastructure of musical rhythm processing is influenced by ADHD and musical expertise.

	Cognitive function	Sensorimotor timing	Rhythm perception	Neural dynamics	Neural pathways	Neuromodulatory systems
ADHD	Deficits in attention, inhibitory control, working memory.	Increased motor timing variability, linked to poor inhibitory control.	Difficulties with beat perception and duration estimation.	Abnormal patterns of oscillatory activity across multiple frequency bands.	Decreased volumes in frontal, parietal and motor regions, including cerebellum and basal ganglia. Decreased connectivity within motor and cognitive control networks.	Disrupted dopaminergic signaling, linked to genetic variation in dopamine receptors and transporters.
Musicians	Enhanced attention, inhibitory control, working memory.	More consistent sensorimotor timing, correlated with enhanced inhibitory control.	Improved accuracy in beat perception and duration discrimination tasks.	Functional changes in oscillatory activity linked to music training, including increased coherence between frequencies.	Increased cerebellar and basal ganglia volumes. Increased connectivity within motor and cognitive control networks.	Preliminary evidence for increased dopamine receptor expression in musicians (potentially indicating genetic predisposition to music).

et al., 2007; Giedd et al., 2008), however the potential role of neural dynamics in mediating these developmental differences remains to be explored. Recent research indicates that ADHD, neural dynamics and timing-related behaviors are all heritable (Tye et al., 2011; Agostino and Cheng, 2016), suggesting that a “genes to behavior” approach may prove fruitful.

## EFFECTS OF EXPERTISE

Several aspects of rhythm processing that are implicated in ADHD are also strengthened in expert musicians (summarized in **Table 1**), suggesting the potential for these systems to be shaped by experience. Behaviorally, musicians are better than controls at rhythm perception and temporal discrimination tasks (Rammsayer and Altenmüller, 2006; Wallentin et al., 2010) and have more consistent sensorimotor timing (Repp and Su, 2013). They also demonstrate enhanced cognitive function, including attention, inhibitory control and working memory (see Benz et al., 2015, for recent review), with enhanced inhibitory control linked to more consistent sensorimotor timing (Slater et al., 2017, 2018). Researchers found that musicians had larger volumes in motor areas including the cerebellum and basal ganglia, as well as frontal and parietal regions associated with cognitive control (see Schlaug, 2015, for review), and music training has been associated with functional changes to oscillatory dynamics (Bhattacharya and Petsche, 2005; Trainor et al., 2009).

It is possible that group comparisons reflect innate differences in those drawn to pursue music rather than causal effects of training, in fact there is some preliminary evidence showing increased expression of dopamine receptors in musicians compared with controls, suggesting a potential genetic tendency toward musicianship (Emanuele et al., 2010). However, evidence from longitudinal studies (Moreno et al., 2011; Roden et al., 2014) as well as links between behavioral enhancements, extent of expertise (Slater et al., 2018) and specific instrument played (Krause et al., 2010) suggest that experience plays at least some role in observed differences. Further, therapies focusing on motor timing or rhythm have shown some success in ameliorating the broader symptoms of ADHD (Shaffer et al., 2001; Leisman and Melillo, 2010; Dahan et al., 2016), although more intervention studies are needed. Taken together, these findings suggest that common underlying mechanisms involved in both cognitive and motor control could potentially be strengthened by music-based interventions, building on the established use of music-based therapies in the treatment of a variety of other disorders. With a clearer understanding of

distinct phenotypes, the efficacy of such interventions for ADHD could be greatly improved.

## CONCLUSIONS

By considering how the brain processes musical rhythm, we force ourselves to take an integrated approach to how the brain coordinates its activities in time. Here, we argue that it is exactly this kind of integrated approach that is needed to advance understanding of a complex, heterogeneous disorder such as ADHD.

Whereas a great deal of neuroscientific research has focused on the spatial dimension—within perception itself, as well as in the localization of functions to particular brain regions—the inherently temporal nature of musical sound helps to bring mechanisms of neural coordination to the forefront. In this review, we have explored common neural infrastructure that is involved in processing musical rhythm, and implicated in ADHD. We have discussed how the heterogeneity of ADHD has hampered progress toward the identification of biomarkers and objective diagnostic tools. We suggest that further investigation of the basis of rhythm and timing deficits could ultimately help to form a more integrated view of the etiologies of ADHD, bridging the gap between genetic factors (e.g., variation in dopaminergic signaling), neural dynamics and the development of cortical networks, and the behavioral control of cognition and movement. We have also highlighted that the same neural systems are strengthened in expert musicians, suggesting the potential for neuroplasticity to have remediating effects. This novel, interdisciplinary approach could inform therapeutic strategies, harnessing the rewarding properties of music to strengthen coordination within the brain.

## AUTHOR CONTRIBUTIONS

JS conceived and wrote the article, MT contributed to the writing of the article.

## FUNDING

Research reported in this publication was supported, in part, by the National Institutes of Health's National Center for Advancing Translational Sciences, Grant Number KL2TR001424. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## REFERENCES

- Abrams, D. A., Ryali, S., Chen, T., Chordia, P., Khouzam, A., Levitin, D. J., et al. (2013). Inter-subject synchronization of brain responses during natural music listening. *Eur. J. Neurosci.* 37, 1458–1469. doi: 10.1111/ejn.12173
- Agostino, P. V., and Cheng, R.-K. (2016). Contributions of dopaminergic signaling to timing accuracy and precision. *Curr. Opin. Behav. Sci.* 8, 153–160. doi: 10.1016/j.cobeha.2016.02.013
- Akshoomoff, N. A., and Courchesne, E. (1992). A new role for the cerebellum in cognitive operations. *Behav. Neurosci.* 106, 731. doi: 10.1037/0735-7044.106.5.731
- Angelis, V., Holland, S., Upton, P. J., and Clayton, M. (2013). Testing a computational model of rhythm perception using polyrhythmic stimuli. *J. New Music Res.* 42, 47–60. doi: 10.1080/09298215.2012.718791
- Arns, M., Conners, C. K., and Kraemer, H. C. (2013). A decade of EEG theta/beta ratio research in ADHD: a meta-analysis. *J. Atten. Disord.* 17, 374–383. doi: 10.1177/1087054712460087



- Ashley, R. (2002). Do [n't] change a hair for me: the art of jazz rubato. *Music Percept.* 19, 311–332. doi: 10.1525/mp.2002.19.3.311
- Balci, F., Day, M., Rooney, A., and Brunner, D. (2009). Disrupted temporal control in the R6/2 mouse model of Huntington's disease. *Behav. Neurosci.* 123, 1353. doi: 10.1037/a0017650
- Balci, F., Ludvig, E. A., Abner, R., Zhuang, X., Poon, P., and Brunner, D. (2010). Motivational effects on interval timing in dopamine transporter (DAT) knockdown mice. *Brain Res.* 1325, 89–99. doi: 10.1016/j.brainres.2010.02.034
- Başar, E., and Güntekin, B. (2008). A review of brain oscillations in cognitive disorders and the role of neurotransmitters. *Brain Res.* 1235, 172–193. doi: 10.1016/j.brainres.2008.06.103
- Benz, S., Sellaro, R., Hommel, B., and Colzato, L. S. (2015). Music makes the world go round: the impact of musical training on non-musical cognitive functions—a review. *Front. Psychol.* 6:2023. doi: 10.3389/fpsyg.2015.02023
- Bhatara, A., Tirovolas, A. K., Duan, L. M., Levy, B., and Levitin, D. J. (2011). Perception of emotional expression in musical performance. *J. Exp. Psychol. Hum. Percept. Perform.* 37, 921. doi: 10.1037/a0021922
- Bhattacharya, J., and Petsche, H. (2005). Phase synchrony analysis of EEG during music perception reveals changes in functional connectivity due to musical expertise. *Signal Process.* 85, 2161–2177. doi: 10.1016/j.sigpro.2005.07.007
- Castellanos, F. X., and Tannock, R. (2002). Neuroscience of attention-deficit/hyperactivity disorder: the search for endophenotypes. *Nature Rev. Neurosci.* 3, 617–628. doi: 10.1038/nrn896
- Cevik, M. (2003). “Neurogenetics of interval timing,” in *Functional and Neural Mechanisms of Interval Timing*, ed W. H. Meck (Boca Raton, FL: CRC Press), 297–316.
- Chen, J. L., Penhune, V. B., and Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cereb. Cortex* 18, 2844–2854. doi: 10.1093/cercor/bhn042
- Cirelli, L. K., Einarson, K., and Trainor, L. J. (2012). “Bouncing babies to the beat: Music and helping behaviour in infancy,” in 12th International Conference on Music Perception and Cognition (Thessaloniki).
- Cirelli, L. K., Wan, S. J., and Trainor, L. J. (2014). Fourteen-month-old infants use interpersonal synchrony as a cue to direct helpfulness. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20130400. doi: 10.1098/rstb.2013.0400
- Cook, P., Rouse, A., Wilson, M., and Reichmuth, C. (2013). A California sea lion (*Zalophus californianus*) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *J. Comp. Psychol.* 127, 412–27. doi: 10.1037/a0032345
- Cooper, G., and Meyer, L. B. (1960). *The Rhythmic Structure of Music*. Chicago, IL: University of Chicago Press.
- Coull, J. T., Hwang, H. J., Leyton, M., and Dagher, A. (2012). Dopamine precursor depletion impairs timing in healthy volunteers by attenuating activity in putamen and supplementary motor area. *J. Neurosci.* 32, 16704–16715. doi: 10.1523/JNEUROSCI.1258-12.2012
- Cross, I. (2001). Music, cognition, culture, and evolution. *Ann. N. Y. Acad. Sci.* 930, 28–42. doi: 10.1111/j.1749-6632.2001.tb05723.x
- Dahan, A., Ryder, C. H., and Reiner, M. (2016). Components of motor deficiencies in ADHD and possible interventions. *Neuroscience* 378, 34–53. doi: 10.1016/j.neuroscience.2016.05.040
- Dalla Bella, S., Białuska, A., and Sowinski, J. (2013). Why movement is captured by music, but less by speech: role of temporal regularity. *PLoS ONE* 8:e71945. doi: 10.1371/journal.pone.0071945
- Demers, M. M., McNevin, N., and Azar, N. R. (2013). ADHD and motor control: a review of the motor control deficiencies associated with attention deficit/hyperactivity disorder and current treatment options. *Crit. Rev. Phys. Rehabil. Med.* 25, 3–4. doi: 10.1615/CritRevPhysRehabilMed.20130.09763
- Desain, P., and Honing, H. (1999). Computational models of beat induction: the rule-based approach. *J. New Music Res.* 28, 29–42. doi: 10.1076/jnmr.28.1.29.3123
- DiMaio, S., Grizenko, N., and Joobar, R. (2003). Dopamine genes and attention-deficit hyperactivity disorder: a review. *J. Psychiatry Neurosci.* 28, 27–38.
- Drew, M. R., Simpson, E. H., Kellendonk, C., Herzberg, W. G., Lipatova, O., Fairhurst, S., et al. (2007). Transient overexpression of striatal D2 receptors impairs operant motivation and interval timing. *J. Neurosci.* 27, 7731–7739. doi: 10.1523/JNEUROSCI.1736-07.2007
- Durston, S., van Belle, J., and de Zeeuw, P. (2011). Differentiating frontostriatal and fronto-cerebellar circuits in attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 69, 1178–1184. doi: 10.1016/j.biopsych.2010.07.037
- Emanuele, E., Boso, M., Cassola, F., Broglia, D., Bonoldi, I., Mancini, L., et al. (2010). Increased dopamine DRD4 receptor mRNA expression in lymphocytes of musicians and autistic individuals: bridging the music-autism connection. *Neuro Endocrinol. Lett.* 31, 122–125.
- 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
- Fair, D. A., Bathula, D., Nikolas, M. A., and Nigg, J. T. (2012). Distinct neuropsychological subgroups in typically developing youth inform heterogeneity in children with ADHD. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6769–6774. doi: 10.1073/pnas.1115365109
- 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
- Giedd, J. N., Lenroot, R. K., Shaw, P., Lalonde, F., Celano, M., White, S., et al. (2008). Trajectories of anatomic brain development as a phenotype. *Novartis Found. Symp.* 289, 101–112; discussion: 112–108; 193–105.
- Grahn, J. A. (2012). Neural mechanisms of rhythm perception: current findings and future perspectives. *Top. Cogn. Sci.* 4, 585–606. doi: 10.1111/j.1756-8765.2012.01213.x
- Grahn, J. A., and Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *J. Cogn. Neurosci.* 19, 893–906. doi: 10.1162/jocn.2007.19.5.893
- Grahn, J. A., and Brett, M. (2009). Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex* 45, 54–61. doi: 10.1016/j.cortex.2008.01.005
- Grahn, J. A., and Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.* 29, 7540–7548. doi: 10.1523/JNEUROSCI.2018-08.2009
- Graybiel, A. M. (1997). The basal ganglia and cognitive pattern generators. *Schizophrenia Bull.* 23, 459–469. doi: 10.1093/schbul/23.3.459
- Grube, M., Lee, K.-H., Griffiths, T. D., Barker, A. T., and Woodruff, P. W. (2010). Frontiers: transcranial magnetic theta-burst stimulation of the human cerebellum distinguishes absolute, duration-based from relative, beat-based perception of subsecond time intervals. *Front. Auditory Cogn. Neurosci.* 1:171. doi: 10.3389/fpsyg.2010.00171
- Hove, M. J., Gravel, N., Spencer, R. M. C., and Valera, E. M. (2017). Finger tapping and pre-attentive sensorimotor timing in adults with ADHD. *Exp. Brain Res.* 235, 3663–3672. doi: 10.1007/s00221-017-5089-y
- Huron, D. B. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT press.
- 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
- Kaiser, M.-L., Schoemaker, M., Albaret, J.-M., and Geuze, R. (2015). What is the evidence of impaired motor skills and motor control among children with attention deficit hyperactivity disorder (ADHD)? Systematic review of the literature. *Res. Dev. Disabil.* 36, 338–357. doi: 10.1016/j.ridd.2014.09.023
- Kebir, O., and Joobar, R. (2011). Neuropsychological endophenotypes in attention-deficit/hyperactivity disorder: a review of genetic association studies. *Eur. Arch. Psychiatry Clin. Neurosci.* 261, 583–594. doi: 10.1007/s00406-011-0207-5
- Keller, P. E., and Rieger, M. (2009). Special issue—Musical movement and synchronization. *Music Percept.* 26, 397–400. doi: 10.1525/mp.2009.26.5.397
- Krause, V., Schnitzler, A., and Pollok, B. (2010). Functional network interactions during sensorimotor synchronization in musicians and non-musicians. *Neuroimage* 52, 245–251. doi: 10.1016/j.neuroimage.2010.03.081
- Large, E. W., Herrera, J. A., and Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Front. Syst. Neurosci.* 9:159. doi: 10.3389/fnsys.2015.00159
- Large, E. W., and Kolen, J. F. (1994). Resonance and the perception of musical meter. *Conn. Sci.* 6, 177–208. doi: 10.1080/09540099408915723
- Large, E. W., and Palmer, C. (2002). Perceiving temporal regularity in music. *Cogn. Sci.* 26, 1–37. doi: 10.1207/s15516709cog2601\_1
- Large, E. W., and Snyder, J. S. (2009). Pulse and meter as neural resonance. *Ann. N. Y. Acad. Sci.* 1169, 46–57. doi: 10.1111/j.1749-6632.2009.04550.x

- Leisman, G., and Melillo, R. (2010). Effects of motor sequence training on attentional performance in ADHD children. *Int. J. Disabil. Hum. Dev.* 9, 275–282. doi: 10.1515/IJDHD.2010.043
- Levitin, D. J., Chordia, P., and Menon, V. (2012). Musical rhythm spectra from Bach to Joplin obey a 1/f power law. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3716–3720. doi: 10.1073/pnas.1113828109
- Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., and Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *J. Neurosci.* 21, 1370–1377. doi: 10.1523/JNEUROSCI.21-04-01370.2001
- Longuet-Higgins, H. C., and Lee, C. S. (1982). The perception of musical rhythms. *Perception* 11, 115–128. doi: 10.1068/p110115
- Loo, S. K., McGough, J. J., McCracken, J. T., and Smalley, S. L. (2017). Parsing heterogeneity in attention-deficit/hyperactivity disorder using EEG-based subgroups. *J. Child Psychol. Psychiatry* 59, 223–231. doi: 10.1111/jcpp.12814
- Mazaheri, A., Fassbender, C., Coffey-Corina, S., Hartanto, T. A., Schweitzer, J. B., and Mangun, G. R. (2014). Differential oscillatory electroencephalogram between attention-deficit/hyperactivity disorder subtypes and typically developing adolescents. *Biol. Psychiatry* 76, 422–429. doi: 10.1016/j.biopsych.2013.08.023
- McAuley, J. D. (2010). “Tempo and rhythm,” in *Music Perception. Springer Handbook of Auditory Research, Vol. 36*, eds M. R. Jones, R. R. Fay, and A. N. Popper (New York, NY: Springer Science + Business Media), 165–199.
- Merchant, H., Grahn, J. A., Trainor, L., Rohrmeier, M., and Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. *Philos. Trans. R. Soc. B Biol. Sci.* 370:20140093. doi: 10.1098/rstb.2014.0093
- Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., and Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychol. Sci.* 22, 1425–1433. doi: 10.1177/0956797611416999
- Mueller, A., Hong, D. S., Shepard, S., and Moore, T. (2017). Linking ADHD to the neural circuitry of attention. *Trends Cogn. Sci.* 21, 474–488. doi: 10.1016/j.tics.2017.03.009
- Müller, V., Sängler, J., and Lindenberger, U. (2013). Intra- and Inter-Brain Synchronization during Musical Improvisation on the Guitar. *PLoS ONE* 8:e73852. doi: 10.1371/journal.pone.0073852
- Nigg, J. T., Willcutt, E. G., Doyle, A. E., and Sonuga-Barke, E. J. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: do we need neuropsychologically impaired subtypes? *Biol. Psychiatry* 57, 1224–1230. doi: 10.1016/j.biopsych.2004.08.025
- Nikolas, M. A., and Nigg, J. T. (2015). Moderators of neuropsychological mechanism in attention-deficit hyperactivity disorder. *J. Abnorm. Child Psychol.* 43, 271–281. doi: 10.1007/s10802-014-9904-7
- Noreika, V., Falter, C. M., and Rubia, K. (2013). Timing deficits in attention-deficit/hyperactivity disorder (ADHD): evidence from neurocognitive and neuroimaging studies. *Neuropsychologia* 51, 235–266. doi: 10.1016/j.neuropsychologia.2012.09.036
- Nozaradan, S. (2014). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20130393. doi: 10.1098/rstb.2013.0393
- Nozaradan, S., Peretz, I., and Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *J. Neurosci.* 32, 17572–17581. doi: 10.1523/JNEUROSCI.3203-12.2012
- Nozaradan, S., Schwartz, M., Obermeier, C., and Kotz, S. A. (2017). Specific contributions of basal ganglia and cerebellum to the neural tracking of rhythm. *Cortex* 95, 156–168. doi: 10.1016/j.cortex.2017.08.015
- Palmer, C. (1997). Music performance. *Annu. Rev. Psychol.* 48, 115–138. doi: 10.1146/annurev.psych.48.1.115
- 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
- Phillips-Silver, J., and Trainor, L. J. (2005). Feeling the beat: movement influences infant rhythm perception. *Science* 308, 1430–1430. doi: 10.1126/science.1110922
- Phillips-Silver, J., and Trainor, L. J. (2007). Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition* 105, 533–546. doi: 10.1016/j.cognition.2006.11.006
- Povel, D.-J., and Essens, P. (1985). Perception of temporal patterns. *Music Percept.* 2, 411–440. doi: 10.2307/40285311
- Puyjarinet, F., Bégel, V., Lopez, R., Dellacherie, D., and Dalla Bella, S. (2017). Children and adults with Attention-Deficit/Hyperactivity Disorder cannot move to the beat. *Sci. Rep.* 7, 11550. doi: 10.1038/s41598-017-11295-w
- Quintana, J., and Fuster, J. M. (1999). From perception to action: temporal integrative functions of prefrontal and parietal neurons. *Cerebral Cortex* 9, 213–221. doi: 10.1093/cercor/9.3.213
- Raichle, M. E. (2010). Two views of brain function. *Trends Cogn. Sci.* 14, 180–190. doi: 10.1016/j.tics.2010.01.008
- Rammsayer, T., and Altenmüller, E. (2006). Temporal information processing in musicians and nonmusicians. *Music Percept.* 24, 37–48. doi: 10.1525/mp.2006.24.1.37
- Rankin, S. K., Fink, P. W., and Large, E. W. (2014). Fractal structure enables temporal prediction in music. *J. Acoust. Soc. Am.* 136, EL256–EL262. doi: 10.1121/1.4890198
- Rankin, S. K., Large, E. W., and Fink, P. W. (2009). Fractal tempo fluctuation and pulse prediction. *Music Percept.* 26, 401–413. doi: 10.1525/mp.2009.26.5.401
- Repp, B. H. (1995). Expressive timing in Schumann’s “Träumerei.” An analysis of performances by graduate student pianists. *J. Acoust. Soc. Am.* 98, 2413–2427.
- 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
- Roden, I., Grube, D., Bongard, S., and Kreutz, G. (2014). Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. *Psychol. Music* 42, 284–298. doi: 10.1177/0305735612471239
- Rommelse, N. N., Altink, M. E., Oosterlaan, J., Beem, L., Buschgens, C. J., Buitelaar, J., et al. (2008). Speed, variability, and timing of motor output in ADHD: which measures are useful for endophenotypic research? *Behav. Genet.* 38, 121–132. doi: 10.1007/s10519-007-9186-8
- Rubia, K. (2016). Can functional decoding elucidate meta-analytic brain dysfunctions in adult attention-deficit/hyperactivity disorder? *Biol. Psychiatry* 80, 890–892. doi: 10.1016/j.biopsych.2016.10.003
- Salimpoor, V. N., and Zatorre, R. J. (2013). Neural interactions that give rise to musical pleasure. *Psychol. Aesthet. Creat. Arts* 7, 62. doi: 10.1037/a0031819
- Schlaug, G. (2015). “Musicians and music making as a model for the study of brain plasticity,” in *Music, Neurology, and Neuroscience: Evolution, the Musical Brain, Medical Conditions, and Therapies, Progress in Brain Research, Vol. 217*, eds E. Altenmüller, S. Finger, and F. Boller (Amsterdam: Elsevier) 37–55. doi: 10.1016/bs.pbr.2014.11.020
- Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., and Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Curr. Opin. Neurobiol.* 20, 172–176. doi: 10.1016/j.conb.2010.02.010
- Schubotz, R. I. (2007). Prediction of external events with our motor system: towards a new framework. *Trends Cogn. Sci.* 11, 211–218. doi: 10.1016/j.tics.2007.02.006
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27. doi: 10.1152/jn.1998.80.1.1
- Schultz, W. (2007). Multiple Dopamine Functions at Different Time Courses. *Annu. Rev. Neurosci.* 30, 259–288. doi: 10.1146/annurev.neuro.28.061604.135722
- 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
- Shaffer, R. J., Jacokes, L. E., Cassily, J. F., Greenspan, S. I., Tuchman, R. F., and Stemmer, P. J. (2001). Effect of Interactive Metronome® training on children with ADHD. *Am. J. Occupat. Ther.* 55, 155–162. doi: 10.5014/ajot.55.2.155
- Shaw, P., Gornick, M., Lerch, J., Addington, A., Seal, J., Greenstein, D., et al. (2007). Polymorphisms of the dopamine D4 receptor, clinical outcome, and cortical structure in attention-deficit/hyperactivity disorder. *Arch. Gen. Psychiatry* 64, 921–931. doi: 10.1001/archpsyc.64.8.921
- Silberstein, R. B., Pipingas, A., Farrow, M., Levy, F., Stough, C. K., and Camfield, D. A. (2016). Brain functional connectivity abnormalities in attention-deficit hyperactivity disorder. *Brain Behav.* 6:e00583. doi: 10.1002/brb3.583

- Silk, T. J., Vance, A., Rinehart, N., Bradshaw, J. L., and Cunnington, R. (2009). White-matter abnormalities in attention deficit hyperactivity disorder: a diffusion tensor imaging study. *Hum. Brain Mapp.* 30, 2757–2765. doi: 10.1002/hbm.20703
- Slater, J., Ashley, R., Tierney, A., and Kraus, N. (2018). Got rhythm? Better inhibitory control is linked with more consistent drumming and enhanced neural tracking of the musical beat in adult percussionists and nonpercussionists. *J. Cogn. Neurosci.* 30, 14–24. doi: 10.1162/jocn\_a\_01189
- Slater, J., Azem, A., Nicol, T., Swedenborg, B., and Kraus, N. (2017). Variations on the theme of musical expertise: cognitive and sensory processing in percussionists, vocalists and non-musicians. *Eur. J. Neurosci.* 45, 952–963. doi: 10.1111/ejn.13535
- Smit, D. J., Linkenkaer-Hansen, K., and de Geus, E. J. (2013). Long-range temporal correlations in resting-state alpha oscillations predict human timing-error dynamics. *J. Neurosci.* 33, 11212–11220. doi: 10.1523/jneurosci.2816-12.2013
- Smit, D. J. A., and Anokhin, A. P. (2017). Development and genetics of brain temporal stability related to attention problems in adolescent twins. *Int. J. Psychophysiol.* 115, 86–97. doi: 10.1016/j.ijpsycho.2016.07.498
- Smith, A., Taylor, E., Rogers, J. W., Newman, S., and Rubia, K. (2002). Evidence for a pure time perception deficit in children with ADHD. *J. Child Psychol. Psychiatry* 43, 529–542. doi: 10.1111/1469-7610.00043
- 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
- Sumbre, G., Muto, A., Baier, H., and Poo, M. M. (2008). Entrained rhythmic activities of neuronal ensembles as perceptual memory of time interval. *Nature* 456, 102. doi: 10.1038/nature07351
- Swanson, J., Flodman, P., Kennedy, J., Spence, M. A., Moyzis, R., Schuck, S., et al. (2000). Dopamine genes and ADHD. *Neurosci. Biobehav. Rev.* 24, 21–25. doi: 10.1016/S0149-7634(99)00062-7
- Tarr, B., Launay, J., and Dunbar, R. I. (2014). Music and social bonding: “self-other” merging and neurohormonal mechanisms. *Front. Psychol.* 5:1096. doi: 10.3389/fpsyg.2014.01096
- Tierney, A., and Kraus, N. (2013). Neural responses to sounds presented on and off the beat of ecologically valid music. *Front. Syst. Neurosci.* 7:14. doi: 10.3389/fnsys.2013.00014
- Tierney, A., and Kraus, N. (2015). Neural entrainment to the rhythmic structure of music. *J. Cogn. Neurosci.* 27, 400–408. doi: 10.1162/jocn\_a\_00704
- Toplak, M. E., Dockstader, C., and Tannock, R. (2006). Temporal information processing in ADHD: findings to date and new methods. *J. Neurosci. Methods* 151, 15–29. doi: 10.1016/j.jneumeth.2005.09.018
- Trainor, L. J., Shahin, A. J., and Roberts, L. E. (2009). Understanding the benefits of musical training: effects on oscillatory brain activity. *Ann. N. Y. Acad. Sci.* 1169, 133–142. doi: 10.1111/j.1749-6632.2009.04589.x
- Tye, C., McLoughlin, G., Kuntsi, J., and Asherson, P. (2011). Electrophysiological markers of genetic risk for attention deficit hyperactivity disorder. *Expert Rev. Mol. Med.* 13, e9–e9. doi: 10.1017/S1462399411001797
- Uhlhaas, P. J., Roux, F., Rodriguez, E., Rotarska-Jagiela, A., and Singer, W. (2010). Neural synchrony and the development of cortical networks. *Trends Cogn. Sci.* 14, 72–80. doi: 10.1016/j.tics.2009.12.002
- USDHHS (2013). *US Department of Health and Human Services. FDA approval letter for NEBA System.* Available online at: [http://www.accessdata.fda.gov/cdrh\\_docs/pdf11/K112711.pdf](http://www.accessdata.fda.gov/cdrh_docs/pdf11/K112711.pdf)
- Valera, E. M., Spencer, R. M., Zeffiro, T. A., Makris, N., Spencer, T. J., Faraone, S. V., et al. (2010). Neural substrates of impaired sensorimotor timing in adult attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 68, 359–367. doi: 10.1016/j.biopsych.2010.05.012
- Wallentin, M., Nielsen, A. H., Friis-Olivarius, M., Vuust, C., and Vuust, P. (2010). The musical ear test, a new reliable test for measuring musical competence. *Learn. Individ. Differ.* 20, 188–196. doi: 10.1016/j.lindif.2010.02.004
- 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
- Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8, 547–558. doi: 10.1038/nrn2152
- Zelaznik, H. N., Vaughn, A. J., Green, J. T., Smith, A. L., Hoza, B., and Linnea, K. (2012). Motor timing deficits in children with attention-deficit/hyperactivity disorder. *Hum. Mov. Sci.* 31, 255–265. doi: 10.1016/j.humov.2011.05.003

**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 © 2018 Slater and Tate. 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) and the copyright owner(s) 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.



# An Oscillatory Neural Autoencoder Based on Frequency Modulation and Multiplexing

Karthik Soman<sup>1</sup>, Vignesh Muralidharan<sup>2</sup> and V. Srinivasa Chakravarthy<sup>1\*</sup>

<sup>1</sup> Bhupat and Jyoti Mehta School of Biosciences, Department of Biotechnology, Indian Institute of Technology Madras, Chennai, India, <sup>2</sup> Department of Psychology, University of California, San Diego, San Diego, CA, United States

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

Andrea Soltoggio,  
Loughborough University,  
United Kingdom  
Ana-Maria Cebolla,  
Free University of Brussels, Belgium

### \*Correspondence:

V. Srinivasa Chakravarthy  
schakra@iitm.ac.in

**Received:** 02 April 2018

**Accepted:** 19 June 2018

**Published:** 10 July 2018

### Citation:

Soman K, Muralidharan V and  
Chakravarthy VS (2018) An Oscillatory  
Neural Autoencoder Based on  
Frequency Modulation and  
Multiplexing.  
*Front. Comput. Neurosci.* 12:52.  
doi: 10.3389/fncom.2018.00052

Oscillatory phenomena are ubiquitous in the brain. Although there are oscillator-based models of brain dynamics, their universal computational properties have not been explored much unlike in the case of rate-coded and spiking neuron network models. Use of oscillator-based models is often limited to special phenomena like locomotor rhythms and oscillatory attractor-based memories. If neuronal ensembles are taken to be the basic functional units of brain dynamics, it is desirable to develop oscillator-based models that can explain a wide variety of neural phenomena. Autoencoders are a special type of feed forward networks that have been used for construction of large-scale deep networks. Although autoencoders based on rate-coded and spiking neuron networks have been proposed, there are no autoencoders based on oscillators. We propose here an oscillatory neural network model that performs the function of an autoencoder. The model is a hybrid of rate-coded neurons and neural oscillators. Input signals modulate the frequency of the neural encoder oscillators. These signals are then multiplexed using a network of rate-code neurons that has afferent Hebbian and lateral anti-Hebbian connectivity, termed as Lateral Anti Hebbian Network (LAHN). Finally the LAHN output is de-multiplexed using an output neural layer which is a combination of adaptive Hopf and Kuramoto oscillators for the signal reconstruction. The Kuramoto-Hopf combination performing demodulation is a novel way of describing a neural phase-locked loop. The proposed model is tested using both synthetic signals and real world EEG signals. The proposed model arises out of the general motivation to construct biologically inspired, oscillatory versions of some of the standard neural network models, and presents itself as an autoencoder network based on oscillatory neurons applicable to time series signals. As a demonstration, the model is applied to compression of EEG signals.

**Keywords:** oscillatory autoencoder, Kuramoto oscillator, adaptive Hopf oscillator, frequency modulation, multiplexing, phase synchronization, EEG

## INTRODUCTION

Despite decades of research, the question of neural code is still controversial. Currently there are two well-accepted approaches to the problem: the spike rate code and the spike timing code. The former assumes that the neural code lies in the spike rate and has given rise to large class of rate-coded neural networks (Lippmann, 1989; Ruck et al., 1990; Lawrence et al., 1997). The latter holds that the code lies in the spike timing and has led to creation of a large class of spiking



neuron networks (Maass, 1997b; Izhikevich, 2003, 2004; Ghosh-Dastidar and Adeli, 2009). Both rate-coded and spiking neuron networks are endowed with universal computational properties (Maass, 1997a; Auer et al., 2008). However the basic functional unit of the brain seems to be, not a single neuron, but a “cell assembly” (Buzsáki et al., 2012), a cortical column being an example of such a unit (Buzsáki and Draguhn, 2004). The collective activity of a cell assembly is not a spike train but a smoother signal called the local field potential (LFP) (Buzsáki et al., 2012). Most of the functional neuro-imaging data including the electroencephalogram (EEG) and functional Magnetic Resonance Imaging (fMRI) encompass the description of the neural activity at this level (Logothetis et al., 2001; David and Friston, 2003; Whittingstall and Logothetis, 2009). Thus when it comes to the description of neural activity at the level of cell assemblies the standard tools and concepts of signal processing could be deployed. The activity of a single cell assembly can then be described in terms of amplitude, frequency, and phase. Communication between two cell assemblies can be described in terms of phase difference at a given frequency. Hence observed neuro physiological phenomena may be explained in terms of oscillator entrainment and phase synchronization.

It is then natural to envisage neural models of three broad classes—rate code based, spike-based, and oscillator based. There are indeed neural models of oscillators (Wang and Terman, 1995; Campbell et al., 1999; Ijspeert, 2008) but they seem to be often applied to specialized purposes and do not seem to enjoy the universality of both rate coded and spiking neuron network models. Oscillatory neuron models are used to model extensively oscillatory phenomena of the brain like building generative models of cortical oscillations to understand brain rhythms and neuronal synchronization (Cumin and Unsworth, 2007; Breakspear et al., 2010). Furthermore when it comes to modeling behavior, they are also restricted to those behaviors that are intrinsically rhythmic like the locomotor movements, rhythmic hand movements, or swimming movements (Ijspeert et al., 2005; Ijspeert, 2008). Such restricted use of oscillator models is untenable since the very same brain oscillations which drive the hand when making rhythmic tapping movements also enable it to perform non-rhythmic point-to-point reaching movements. Although there are exceptions to this case (see Hoppensteadt and Izhikevich, 2000; Heitmann et al., 2015) there exist only a minimal literature on using oscillatory dynamics to explain non-oscillatory behavior. Therefore it is important to investigate if oscillatory neural network models possess the property of universal computation that forms the core strength of its rival models: rate-coded and spiking neural network models.

The strength of the rate coded and spiking neuron networks lies in the fact that they have been designed to solve a wide of range of useful information processing problems: to construct transformations from one space to another (Lippmann, 1989; Schmidhuber, 2015), to map high dimensional information onto bounded two-dimensional spaces (Kohonen, 1998), to process sequences (Frasconi et al., 1995), to store patterns as attractors (Hopfield, 1982; Trappenberg, 2003), to construct dimensionality reduced representations by autoencoding (Oja, 1989; Sanger, 1989; Hinton and Salakhutdinov, 2006) and so on. In this

realm of applications, in most cases, equivalent oscillatory neural network models have not been designed which, when realized, could form another dimension for understanding standard neural network theory.

Apart from the aforementioned research on neural codes, in the realm of neural signal processing, it becomes natural to link the brain signals arising from EEG and MEG to an underlying oscillatory process which connects to the mechanistic underpinnings of brain circuitry. Utilizing these ideas, a large body of literature exists in the domain of EEG related applications like Brain Computed Interfaces (BCIs). Often in these studies motor imagery EEG signals are recorded, classified and the results of classification are used to drive a machine like the wheelchair (Leeb et al., 2007a,b). The dependence on the stationarity of signals is very important for current methods, including optimal spatial filtering (Ramoser et al., 2000) to solve these class of problems posing difficulty in reliable processing of EEG. The stochastic and non-linear nature of EEG signal thus poses critical challenges in its processing such as feature extraction and further classification (Pfurtscheller and Neuper, 2001). As of now, there exists no benchmark method to decipher this problem of EEG processing. We believe that a better understanding of the oscillatory neural network models, mimicking the underlying neural process, could pave way to a novel class of algorithms for processing EEG signals.

Although the objective of the proposed model is to shed light on the oscillatory neural code, we would also like to briefly cite literature on time series data mining and time series representations. Time series data mining is apparently a challenging one because of the unique characteristic features of the time series data such as presence of noise, and non-linear relation of the data elements (Wilson, 2017). A problem that often arises in time series data processing is to form an optimal representation of the data either by reducing or approximating it, but making sure that the approximated version of the data still carries the local/global features of the original version. For the ease and efficient use of the data, the main challenge is to choose an optimal representation of the same. Time series data representation is a well-studied area where methods such as Discrete Fourier Transform (DFT) (Faloutsos et al., 1994), Discrete Wavelet Transform (DWT) (Percival and Walden, 2006), time series Piecewise approximation (Keogh et al., 2001a,b) have been proposed. Due to the current trends in the use of “big data” processing, other novel methods such as transformation of the time series data to discrete variables or symbols has become popular (Lin et al., 2007). The main idea behind this type of methodology is to transform time series data to a sequential data of symbols by initially discretizing the time series using methods like Piecewise Aggregate Approximation (PAA) (Keogh et al., 2001b). This can be treated as a way to reduce the number of points in the time series data and this is followed by converting the approximated numerical data to corresponding symbols using popular algorithms like SAX (Lin et al., 2007). The advantage of converting the time series to symbolic sequences is that, once the transformation is made, standard pattern matching algorithms can be applied to the sequences for further processing. The aforementioned methods are successful in data mining area,

but carry little information on the neural processing of time series data. This is not a flaw of the aforementioned methods because they are not intended to provide any neural perspective on time series data processing. However, the real brain is adept at time series processing since most of the sensory inputs coming from different sensory modalities such as vision, proprioception, auditory, vestibular, tactile, and olfactory stimulus are dynamic in nature. Hence, the objective of this study is to propose a computational model that implements the autoencoding of time series data using biologically plausible neural principles. The very next sub section named as “background” gives the impetus behind the proposed modeling architecture.

## Background

In response to the aforementioned general motivation, we now present a network of neural oscillators that serves as an oscillatory autoencoder. The reason why we choose the autoencoder architecture is due to the function it serves i.e., encoding the high dimensional input to a low dimensional abstract representation and further decoding it back to the original input signal. From a neural perspective this can be broadly viewed as different stages of neural information transfer. The first stage starts with the encoding of high dimensional sensory stimulus coming from multiple sensory modalities to a more compatible abstract representation in the subcortical structures. For example, visual information fetched by ~125 million retinal photoreceptors converge to ~1 million neurons of the lateral geniculate nucleus in the thalamus (Hubel, 1995). This is one of the instances (among many) of huge dimensionality reduction that takes place in the real brain. The decoder can be viewed as the stage in which the information is transferred from the subcortical structures to other cortical structures with more number neurons i.e., transfer of information from lower dimension to higher dimension (Guillery and Sherman, 2002). Standard autoencoder networks use static neurons that have limitations in capturing the temporal features of the input in a naturalistic fashion. The proposed model uses the dynamics of oscillatory system such as phase synchronization, frequency tuning, and also uses the signal processing concepts such as frequency modulation (FM) and multiplexing (MUX) to shed light on the possible information transfer mechanisms in the brain.

We brief out here the methods that are adopted to accomplish the aforementioned objective (this is explained in detail in the following methods section). In this model, a set of band-limited signals are frequency modulated by a layer of neural oscillators, multiplexed by a layer of rate-coded neurons, and subsequently demultiplexed and demodulated by oscillatory neurons. The network is a hybrid model consisting of two kinds of oscillator models (Kuramoto and Hopf oscillators) and rate-coded dynamic neurons. The signals obtained at the output of the MUX stage may be considered as a reduced-dimensional representation of the input signals. Finally we test the model on actual EEG signals (real world data). The paper is outlined as follows. Section II presents the methods and the model equations, followed by the results in Section III and finally the discussion in Section IV.

## METHODS

Here we propose the architecture of an autoencoder using oscillatory neurons. The motivation for an oscillatory autoencoder is explained above in the introduction section. The model architecture described here consists of Encoder and Decoder modules as shown in **Figure 1**. The encoder process the input signals and makes a lower dimensional representation of the same. The decoder module reconstructs back the original input signal from this abstract representation.

The encoder receives inputs as an array of  $N$  band limited signals,  $s_1(t), \dots, s_N(t)$ . These signals are frequency modulated and multiplexed by the encoder. The multiplexed signals are demultiplexed and demodulated by the decoder. Both the encoder and the decoder are networks of oscillators. The networks are hybrids of Hopf and Kuramoto oscillators (Kuramoto, 1984; Righetti et al., 2006). The motivation for choosing two different phase oscillators is described in the decoder section. The encoder and decoder modules are modeled as follows.

### A. Encoder

The encoder has two stages viz. Frequency Modulation (FM) stage and MUX stage.

#### FM Stage

FM stage has  $N$  phase oscillators each with different intrinsic frequencies.  $N$  is equal to the dimension of the input. Each of the input signals is connected to one of these oscillators. Input is encoded by the phase dynamics as given in (1). This phase dynamics is equivalent to FM (Haykin et al., 1989) and hence the name FM stage.

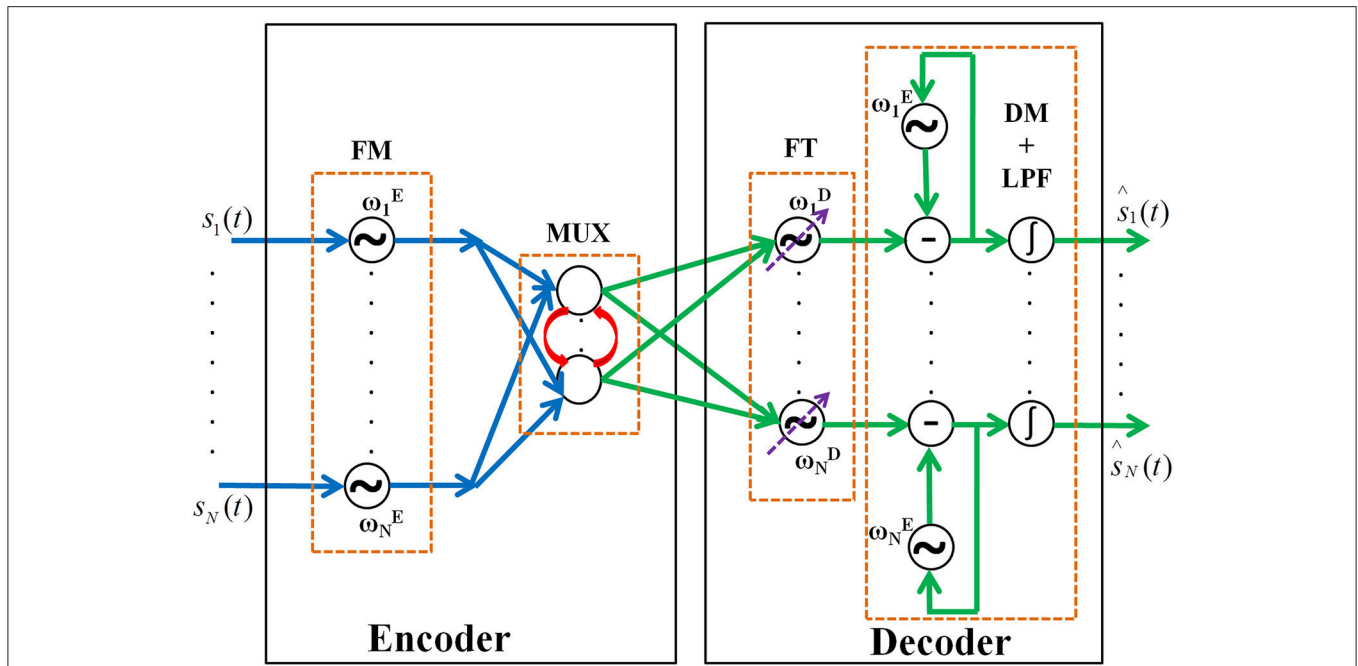
$$\dot{\theta}_i = \omega_i^E + s_i(t) \quad (1)$$

$\theta_i$  is the phase of the  $i$ th oscillator in the encoder layer.  $\omega_i^E$  is the intrinsic angular frequency of the  $i$ th oscillator in the encoder layer. (Note: The superscript E stands for Encoder layer).

#### MUX Stage

A classical MUX in electronics literature ensures harmonious transfer of information between the sender and receiver by acting like a multiple switch (Omotayo, 1985). Hence, a MUX usually has  $n$  number of input lines and 1 output line. However, in the proposed model we do not use this strict definition of MUX instead we take the idea of compressing the  $n$  input signals to  $m$  dimensions where  $m < n$ . This is what is exactly achieved through the hidden layers of a traditional autoencoder. The reason why we named it MUX is to bring about a direct comparison of neural information transfer to the radio FM communication principles.

The MUX stage is implemented by a neural network architecture known as Lateral Anti-Hebbian Network (LAHN) that has Hebbian (excitatory) afferent and anti-Hebbian (inhibitory) lateral connections (Földiák, 1990). The dynamics of a neuron in LAHN is given by Equation (2). Hebbian learning applied to the afferent weight connections (Equation 5) brings the afferent weight vector close to the input data ensuring



**FIGURE 1 |** The network architecture of the oscillatory autoencoder network. In the encoder module, the incoming message signals ( $s_1, \dots, s_N(t)$ ) are encoded, via FM, onto carrier signals with intrinsic frequencies  $\omega_1^E, \dots, \omega_N^E$ . A lateral anti-Hebbian layer is used for frequency multiplexing (MUX) of the modulated signals. In the decoder module, frequency tracking (FT) is performed by a series of adaptive Hopf oscillators which can tune their intrinsic frequency dynamically, followed by a layer of Kuramoto oscillators which synchronize to the frequency of the Hopf oscillators to form a basic unit which extracts the embedded input signals (DM). This is further passed to a leaky integrator (LPF) to get smoothened output of the network [ $\hat{s}_1, \dots, \hat{s}_N(t)$ ].

feature selection by that particular neuron. The anti-Hebbian rule applied to the lateral connections induces competition among the LAHN neurons. Hence each LAHN neuron learns different features from the input data. This network was shown to extract optimal features from the input data by converging transformation weight vectors to the subspace of the principal components of the input data (Földiák, 1990). Since this network maximizes the variance of the output (Földiák, 1990), it extracts optimal features from the input data. The information required for the unsupervised learning of LAHN neuron is available locally at its synaptic connections (Equations 4, 5) and this makes the network biologically plausible.

This LAHN layer acts as the hidden layer for the oscillatory autoencoder. The low-dimensional representations constructed by the hidden layer of a traditional autoencoder are constructed by this MUX stage in the proposed model. Hence, the number of inputs going to the MUX layer is same as the number of encoder oscillators in the FM stage and the number of outputs from the MUX should be essentially lesser in number than its input to achieve a dimensionality reduction.

The dynamics of a neuron in the MUX stage is given in (2) and (3).

$$Y_i(t) = \sum_{j=1}^N q_{ij} O_j(t) + \sum_{k=1}^n w_{ik} Y_k(t-1) \quad (2)$$

$$O_j = \sin(\theta_j) \quad (3)$$

$Y_i$  is the output of  $i$ th neuron,  $q$  and  $w$  are the afferent and lateral weight connections of MUX respectively,  $N$  is the dimension of the input,  $n$  is the total number of neurons in the LAHN.  $O_j$  is the state of the  $j$ th input oscillator. In MUX, lateral weights are updated using anti-Hebbian learning and afferent weights are updated using Hebbian learning (Földiák, 1990) as given in (4) and (5).

$$\Delta w_{ik} = -\eta_L Y_i(t) Y_k(t-1) \quad (4)$$

$$\Delta q_{ij} = \eta_F [O_j(t) Y_i(t) - q_{ij} Y_i^2(t)] \quad (5)$$

$\eta_L$  and  $\eta_F$  are the learning rates for lateral and feed forward weights respectively. MUX with  $n$  nodes trained using (4) and (5) mixes the input FM signals with a minimal overlap in their frequency spectrums which further decreases the reconstruction error.

## B. Decoder

The decoder has three stages such as Frequency Tracking (FT) performed by adaptive Hopf oscillators, Demodulation (DM) using Kuramoto oscillators, and final smoothening of signal by low-pass filtering (LPF) using leaky integrator neurons stages respectively. Each section is explained in detail below.

### FT Stage

Initially the responses of MUX are passed onwards to the FT stage. The purpose of this stage is to tease out the individual frequencies which are mixed by the MUX stage. This

frequency tracking is achieved by using Hopf oscillators with adaptive frequency dynamics. Hopf oscillators were successfully implemented as an adaptive frequency system that updates its intrinsic frequency in an iterative way until it converges to one of the frequencies of the input data (Righetti et al., 2006). This system of Hopf oscillators was previously shown by Righetti et al. (2006) to learn the frequency components of its input signals. This was achieved by adding a frequency adaptation variable to the classical two variable Hopf oscillator dynamics (Righetti et al., 2006). This was shown in phase oscillators having unit circle phase plane limit cycles i.e., using Hopf oscillators. They have further explained similar frequency adaptation dynamics for relaxation oscillators too. However, in this model we are using harmonic phase oscillators for the frequency tracking stage as explained below.

Here, we wanted to achieve the aforementioned phenomena of tracking the frequency of input data. The adaptive frequency Hopf oscillators act like band-pass filters and filter out different frequency bands from the mixed input signal. The adaptive frequency dynamics is accomplished using the following equations:

$$\dot{r}_i = r_i(\mu - r_i^2) \quad (6)$$

$$\dot{\phi}_i = \omega_i^D - \frac{\varepsilon}{r_i} Y \sin(\phi_i) \quad (7)$$

$$\dot{\omega}_i^D = -\varepsilon Y \sin(\phi_i) \quad (8)$$

$r$ ,  $\phi$  and  $\omega_D$  are the radius, phase and angular frequency variables of a Hopf oscillator respectively (Note: the superscript  $D$  stands for Decoder module).  $\mu$  is the parameter that controls the radius of the limit cycle. For  $\mu = 1$ , it produces a unit circle limit cycle.  $\varepsilon$  is the coupling factor between the MUX and the Hopf oscillators (Righetti et al., 2006). Because of linearity of the MUX,  $\varepsilon$  can be computed directly using (9).

$$\varepsilon = P^+ \quad (9)$$

$$P = (I - W)^{-1}Q \quad (10)$$

$P$  is the transformation matrix of the MUX and  $P^+$  is the pseudo inverse of matrix  $P$ .  $I$  is the identity matrix.  $W$  and  $Q$  are the lateral and afferent weight matrices of LAHN respectively.  $P$  can be derived by virtue of the linearity of LAHN as given in (2).

## DM Stage

The purpose of the DM stage is to extract the low-frequency, band limited message signals from the outputs of the FT stage. The DM consists of a layer of Kuramoto oscillators. This shift from Hopf oscillator (in FT stage) to Kuramoto oscillator (in DM stage) is to implement the process of phase synchronization. Kuramoto oscillatory dynamics have been previously implemented to achieve phase synchronization (Kuramoto, 1984). This synchronization in the phase of two oscillators is essential for extracting the message from the FM signal (Haykin et al., 1989) (see Supplementary Material). Each Hopf oscillator in the FT stage is coupled in a one-to-one fashion to a Kuramoto oscillator in the DM stage. The pairs of oscillators

(the Kuramoto oscillators of DM and the Hopf oscillators of FT stage) are coupled through their respective phase variables as shown in (11) and (12).

$$\dot{\gamma}_i = \omega_i^E + K D_i \quad (11)$$

$$D_i = \sin(\phi_i - \gamma_i) \quad (12)$$

$\gamma_i$  is the phase variable of  $i$ th Kuramoto oscillator. It has the same intrinsic frequency,  $\omega_i^E$ , as that of the encoder oscillators (Equation 1) and  $K$  is a positive coupling factor (Kuramoto, 1984). This stage is crucial since phase synchronization occurs at this stage and the synchronization dynamics further decodes the low frequency message signal embedded in the output of the Hopf Oscillator (see Supplementary Material).

## LPF Stage

$D_i$  shown in (12) is the output of the decoder which is further passed through a leaky integrator to smoothen the outputs, i.e., low pass filtering (LPF stage). Leaky integrator acts as a low pass filter which further smoothen out the decoded signal, and eliminates any high frequency components present. Dynamics of leaky integrator is given in (13).

$$\frac{d\hat{s}_i}{dt} = -A\hat{s}_i + D_i(t) \quad (13)$$

$s_i$  is the state of  $i$ th leaky integrator which is the reconstructed version of the input signal  $s_i(t)$ ;  $A$  is the leakage factor which is a positive constant.

Hence the proposed model is a hybrid one consisting of oscillatory layers sandwiching a rate coded layer. Hopf oscillators are used in the model for FM. A layer of linear neurons with lateral connections is used for frequency multiplexing which essentially mixes the FM signals. Hopf oscillators with adaptive frequency are used to track the carrier frequencies of the FM signals. Finally, Kuramoto oscillators are used to demodulate the FM signal and extract the message signal. Parameter values used for the simulation is given in **Table 1**.

## RESULTS

We now test the model described in the previous section on an array of synthetic signals and also on real world EEG signals.

**TABLE 1** | Parameter values.

Parameter	Value
$\mu$	1
$\eta_F$	0.01
$\eta_L$	0.01
$dt$	0.01 sec
$\epsilon$	3
$K$	1



## A. Simulation of the Model on Synthetic Signals

The synthetic signals used for the simulation are of the general form  $s(t) = A_1 \sin(\omega_1 t) + A_2 \sin(\omega_2 t)$ .

Specifically, we consider 4 signals shown in (14), (15), (16) and (17) (Figure 2).

$$s_1(t) = \sin(10\pi t) + 0.5 \sin(12\pi t) \quad (14)$$

$$s_2(t) = \sin(20\pi t) + 0.5 \sin(28\pi t) \quad (15)$$

$$s_3(t) = \sin(50\pi t) + 0.5 \sin(56\pi t) \quad (16)$$

$$s_4(t) = \sin(70\pi t) + 0.5 \sin(80\pi t) \quad (17)$$

The initial intrinsic angular frequencies of the FM oscillators are taken as  $\omega^E = [200 \text{ Hz}, 350 \text{ Hz}, 850 \text{ Hz}, 1000 \text{ Hz}]$ . The input signals, as given by Equations (14)–(17), are used to modulate the encoder oscillators as per Equation (1). Let the resultant frequency modulated signals be  $O_1, O_2, O_3, O_4$  respectively as given by Equation (3). **Figures 2A–D** shows the waveforms of the input signals (for a short duration) as given by Equations (14)–(17). **Figures 3A–D** shows the corresponding frequency spectra. All the frequency spectra are obtained using the Fourier Transformation on the input signals. **Figures 3A–D** clearly show that the input signals are modulated to the higher frequency regime corresponding to the respective carrier waves.

The modulated signals are passed through a MUX which has two neurons ( $n = 2$ ). The outputs of MUX neurons (MUX composite signal) are  $Y_1$  and  $Y_2$  as per Equation (2). The spectra of  $Y_1$  and  $Y_2$  are depicted in **Figures 3E,F**. It is evident from **Figures 3E,F** that the MUX selectively picks and mixes the frequency components of the input signals in such a way that their frequency spectra have minimum overlap. In **Figure 3E**, one neuron of the MUX was more biased to frequency spectra of  $O_1$  and  $O_4$ . In **Figure 3F**, the second neuron of the MUX was more biased to frequency spectra of  $O_2$  and  $O_3$ . The tendency of the hidden layer in

a traditional autoencoder to decorrelate the input signals, is manifesting in the present context as a tendency to remix the input signals so that there is minimal overlap in the spectrum (Földiák, 1990).

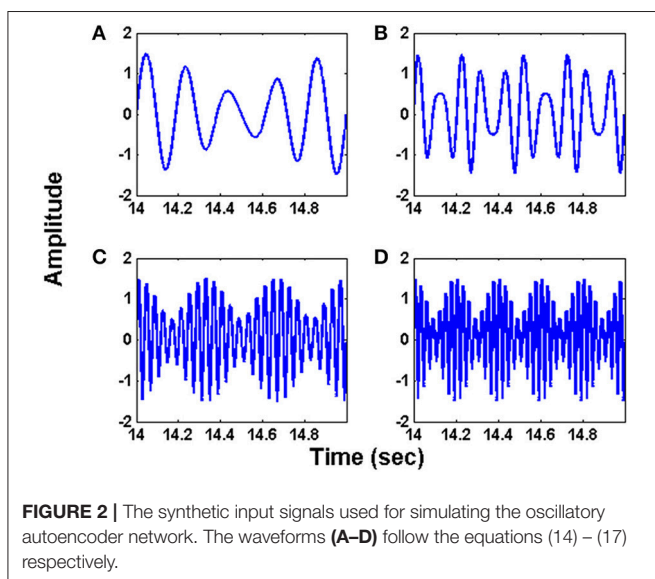
The FT stage has four Hopf oscillators, which are intended to track the four modulating frequencies. Tracking the frequency is similar to tuning the intrinsic oscillations to that particular channel frequency to fetch the information passed through that respective channel. **Figure 4** depicts the frequency adaptation of Hopf oscillators at the FT stage. The intrinsic frequencies of Hopf oscillators are initialized randomly and during the course of time their frequencies get entrained to a specific modulator frequency. Through this adaptation, oscillators are able to select a specific channel of information from a mixture of MUX signals.

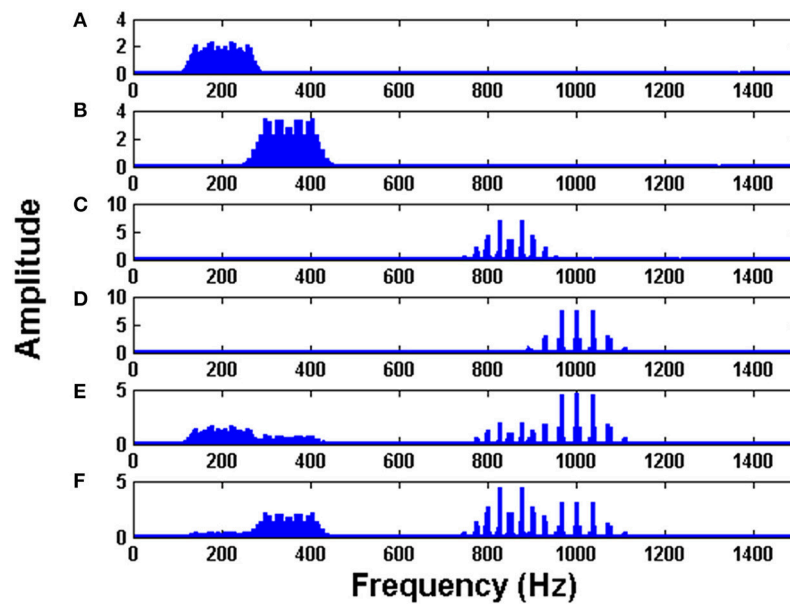
**Figure 5** shows the FFT of the four Hopf oscillators' responses. It is evident from the spectrum that each Hopf oscillator is able to pick individual channel that carries the message signal and hence implements the demodulation of the frequency modulated signals. This is an interesting phenomenon which is also observed in the real brain where two cortical regions get entrained to a similar LFP frequency for information transfer or feature binding (Singer and Gray, 1995; Fell and Axmacher, 2011). Synchronization phenomena also circumvent the need for any training between the cortical structures to learn the transmitted information. That is, simply by tuning to a common frequency, two neural structures can communicate over a temporary channel, without any retraining of connections. This is discussed further in detail in the discussion section.

**Figures 6A–D** shows the original and the reconstructed signals (shown for a short duration). The demodulated signals are of lower amplitude and phase shifted compared to the original signal. This can be further corrected using proper amplification and lag shift operation on the output signals. To quantify the accuracy of reconstructed signal, we compute the reconstruction error. This gives an idea on how good the system is with regard to its function as an autoencoder. It is not advisable to directly compare the input and the raw output signal because of the phase shift present in the output signal. To this end, we first corrected the phase shift in the output signal, by computing cross correlation between the input and the output signal. Next, we computed the lag corresponding to the maximum correlation value and circular shifted the output signal using the previously found maximum lag value to correct the phase shift. The percentage (%) reconstruction error is then computed as the deviation of the Pearson correlation coefficient between the input and the phase corrected output signal from unity.

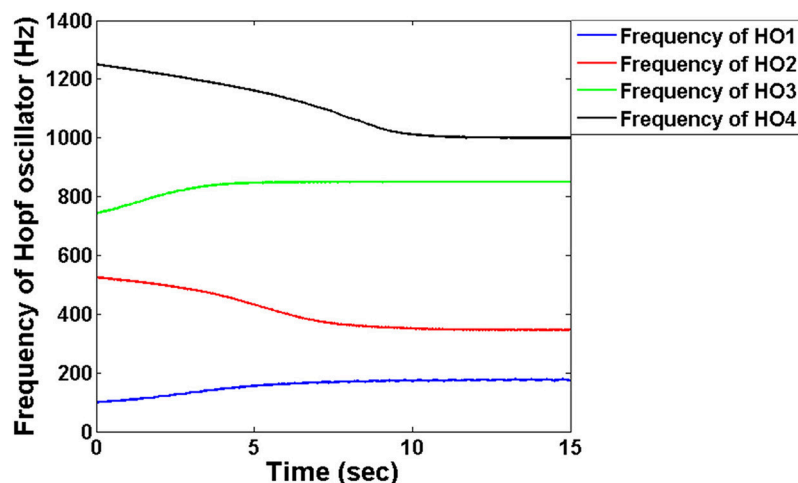
$$\%error = [1 - \text{corr}(x, y)] \times 100$$

where  $x$  is the input signal and  $y$  is the amplitude and phase corrected reconstructed output signal. The percentage (%) reconstruction error with respect to the number of nodes in the LAHN network shows a decreasing trend indicating a better recovery of signal with increasing number of nodes in the LAHN layer (**Figure 6E**). The reconstruction error is computed after





**FIGURE 3 |** The Fourier transform (FFT) of: **(A)** Modulated signal  $O_1$  **(B)** Modulated signal  $O_2$  **(C)** Modulated signal  $O_3$  **(D)** Modulated signal  $O_4$  **(E)** MUX composite signal  $Y_1$  **(F)** MUX composite signal  $Y_2$ . To train the MUX network (LAHN), afferent weights were initialized using random values from uniform distribution  $[0, 1]$  and lateral weights were initialized to zero.  $\eta_L$  and  $\eta_F$  were taken as  $10^{-4}$ .



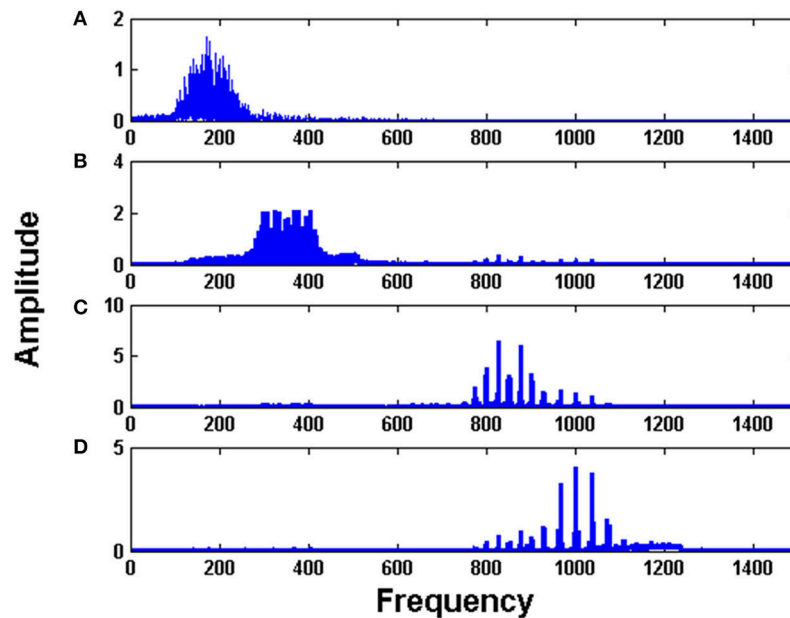
**FIGURE 4 |** The adaptation of frequencies at the level of the Hopf oscillators in the decoder module: The frequencies are initialized randomly close to the encoder frequencies which via adaptation entrain into each of the four carrier signal frequency. We chose  $\mu = 1$  for this simulation.

phase correction of the output signals as explained above. This result shows that choosing an optimal number of neurons in the hidden layer (based on the reconstruction error), it is possible to form a more efficient abstract representation of the input signal.

## B. Simulation of the Model on Real World Signals (EEG Signals)

This section explains the simulation results of the oscillatory autoencoder model on real world data (i.e., data obtained

through empirical ways). For this, we considered empirically recorded EEG signals obtained from BCI Competition 2008 - Graz data set B (Leeb et al., 2008). The dataset essentially consists of two class motor imagery EEG signals recorded from three channels (C3, Cz, and C4) (Leeb et al., 2008) at a sampling rate of 250 Hz. **Figure 7** shows the 1 s duration EEG signals from the aforementioned channels. These EEG signals were recorded during a motor imagery task. For further information on the experimental protocol the readers may refer (Leeb et al., 2008).



**FIGURE 5 |** The Fourier spectrum of the four Hopf oscillators: (A–D) in the decoder layer show tuning of each oscillator to each channel.

These three EEG signals form the input to the model, which is further used to modulate the frequency of the phase oscillators with intrinsic frequencies (500, 600, and 750 Hz). **Figures 8A–C** shows the frequency spectrum of the frequency modulated signals (EEG FM signals). These signals were further forward passed to the LAHN layer (with two nodes) to get the low dimensional representation of the same and to perform MUX operation. **Figures 8D,E** shows the frequency spectrum of the MUX-LAHN signals. It is vivid from the figure that each LAHN neuron captures the frequency information of the EEG FM signals and hence forms a low dimensional representation of the raw EEG signals.

Composite MUX signals are further forward passed to the adaptive Hopf oscillators where each Hopf oscillator tunes its intrinsic frequency to each channel frequency. Adaptive Hopf oscillators thus separate the signals from the composite MUX signal (as shown in **Figures 9A–B**) and this is evident from the frequency spectrum of each Hopf oscillator (**Figures 9C–E**).

The adaptive Hopf oscillator outputs are further passed to the demodulator Kuramoto oscillators for phase locking and extracts out the embedded EEG signals. **Figure 10** shows the reconstructed EEG signal from three channels along with the original signal. The reconstructed EEG signals from the oscillatory autoencoder are smoother than the original EEG signals. This smoothing could be due to the large time scales that govern learning in LAHN and adaptive Hopf oscillatory stage. The subtle changes in the reconstructed signal are due to the lower dimensional representation of the LAHN hidden layer. However, the hidden layer serves as a reliable low dimensional representation of the EEG signals which is further delineated in the following discussion section.

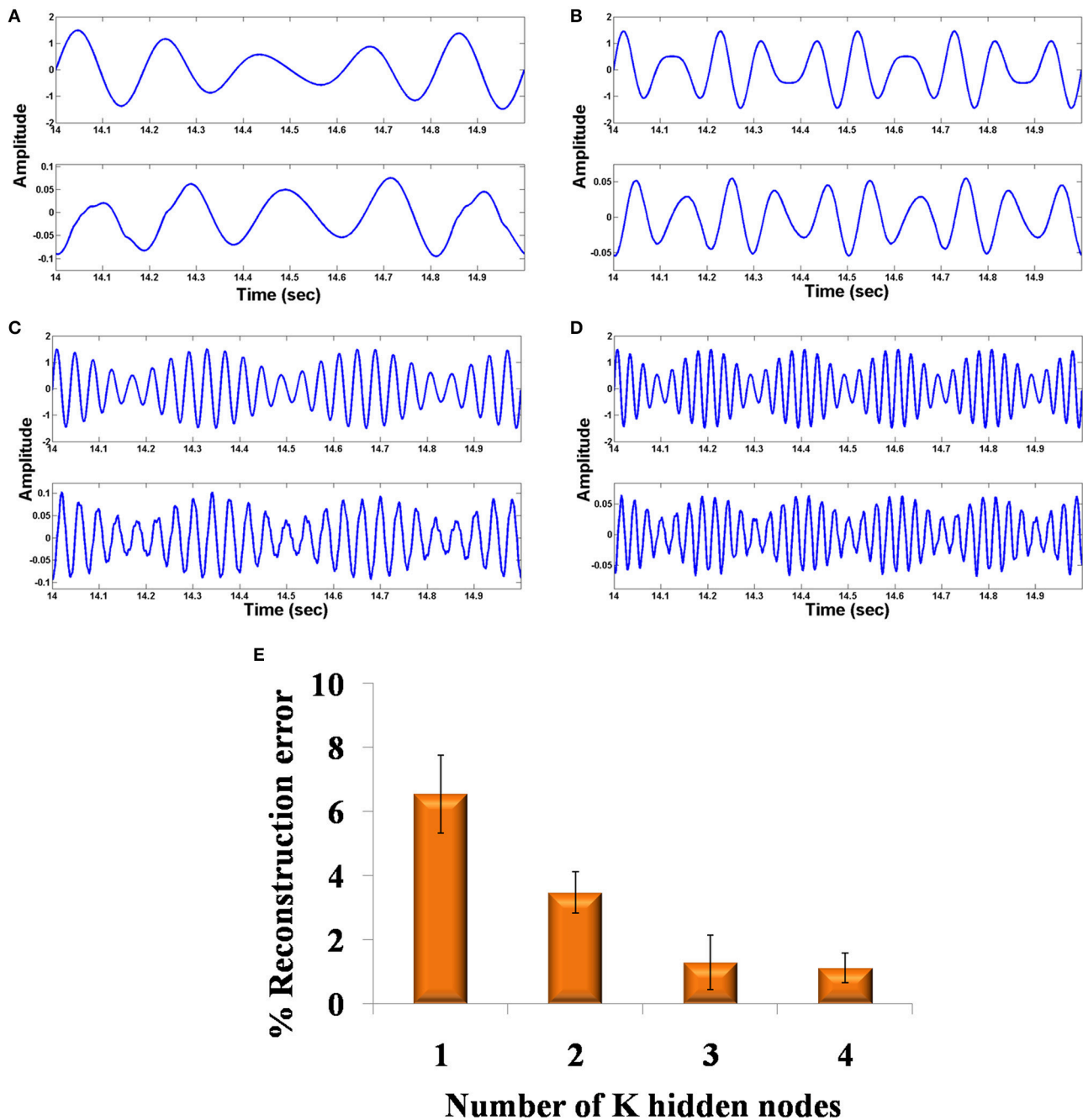
## Comparison of the Model Result With the Benchmark Method for Dimensionality Reduction

In the case of aforementioned EEG result, apart from computing the % reconstruction error, we compare the obtained values with a benchmark dimensionality reduction and reconstruction method to check the goodness of the proposed model. To accomplish this we performed standard Principal Component Analysis (PCA) on the input EEG data to reduce its dimension and further reconstructed back the signal to compute the % reconstruction error. After computing the reconstruction error for each signal, an average reconstruction error is computed to compare it with that of the proposed oscillatory network model. The average reconstruction error of PCA is obtained as 5.68% (computed using MATLAB custom code) considering the first two principal components (because LAHN layer in the model has 2 neurons). The average reconstruction error of the oscillatory autoencoder model is obtained as 5.31%. The average reconstruction error of the proposed oscillatory network model is slightly lower than the standard PCA method by 0.31%. Apart from the decrease in the reconstruction error, the neural attributes of the proposed model and also the theory that the model embodies on the mechanisms of neural information transfer in the brain enhances the significance of the proposed model.

## DISCUSSION

### Summary of the Work

We propose here an oscillatory autoencoder that reconstructs the input signal using a well defined encoder and decoder

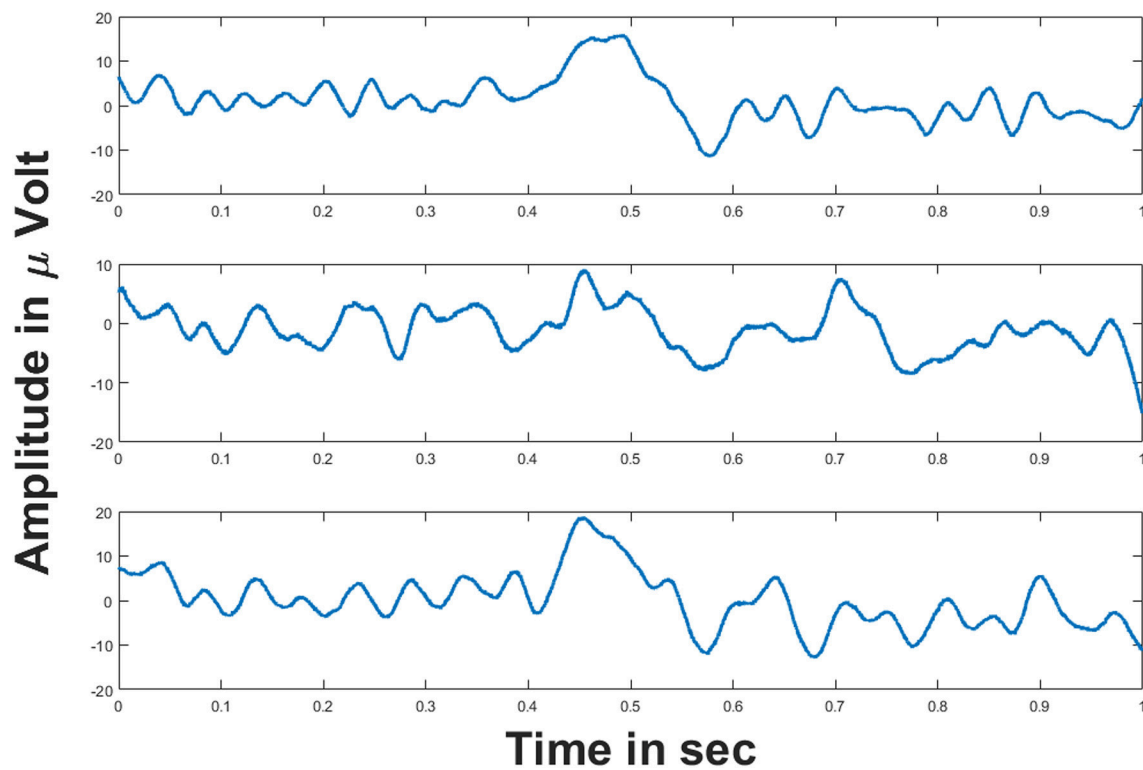


**FIGURE 6 |** Comparison of original and reconstructed signal from the oscillatory autoencoder: (A–D) show the original input (top) to the encoder and the reconstructed signal (bottom) from the decoder using  $n = 2$  in the MUX layer. We used  $A = 0.9$  for the LPF stage. (E) shows the % of reconstruction error with respect to the number of nodes ( $n = 1, 2, 3, 4$ ) in the MUX layer.

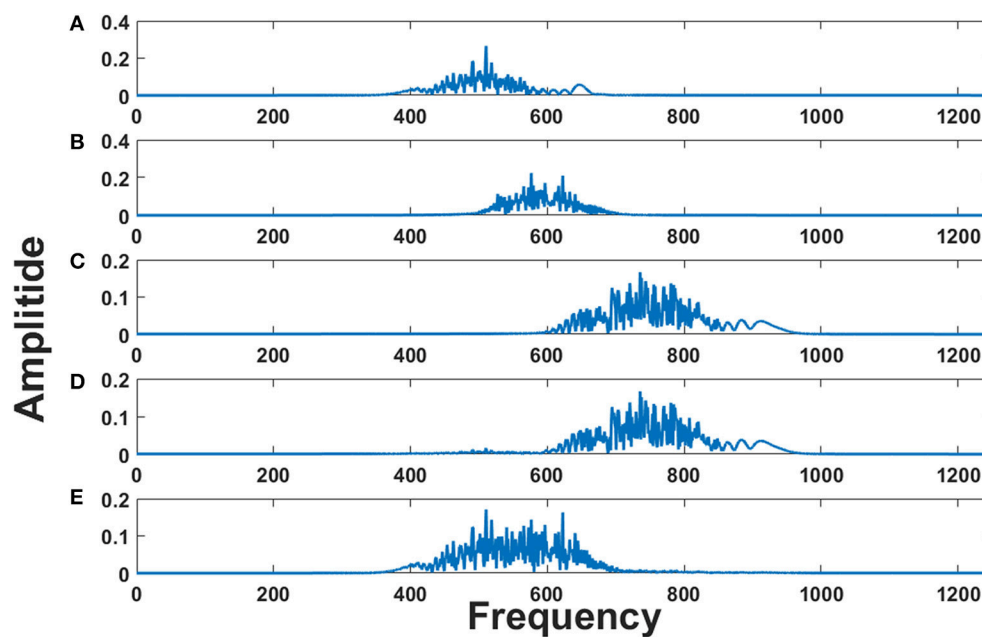
using the principles of FM, MUX, adaptive frequency dynamics, and phase synchronization. We simulated the model using synthetic (linear combination of sinusoids) and also real world EEG signals, thus showing the robustness of the model. The proposed study gives a proof of principle for the potentiality of the oscillatory neural networks in

non-trivial applications where oscillations are seldom used such as autoencoder (problem addressed in this paper), feature extraction, clustering, classification etc where mostly rate coded networks are used. The criticality of oscillations in neurobiology, as mentioned below, is the motivation of this work.

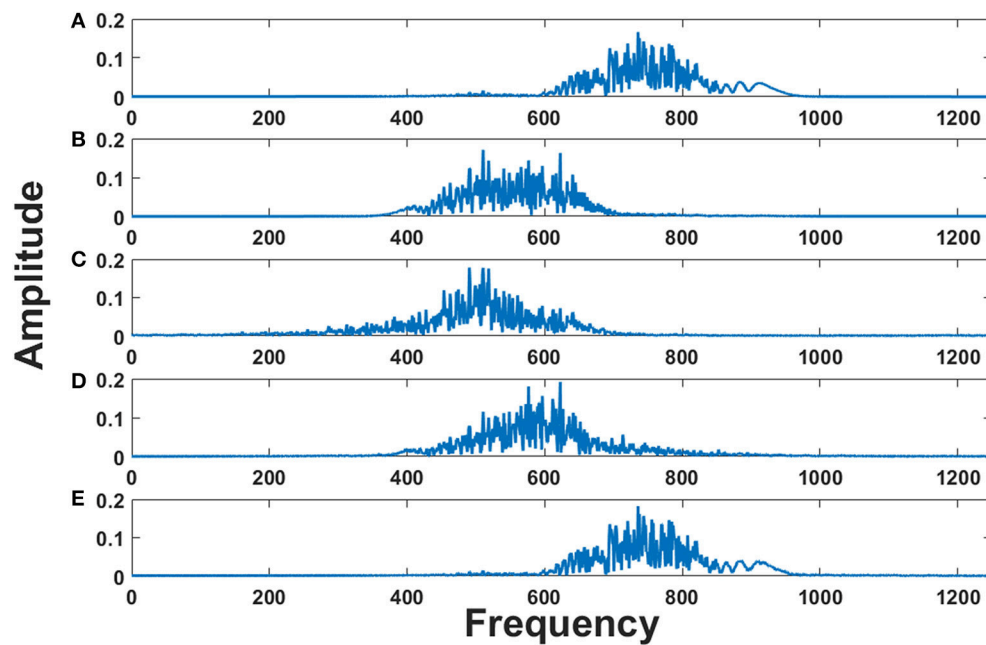




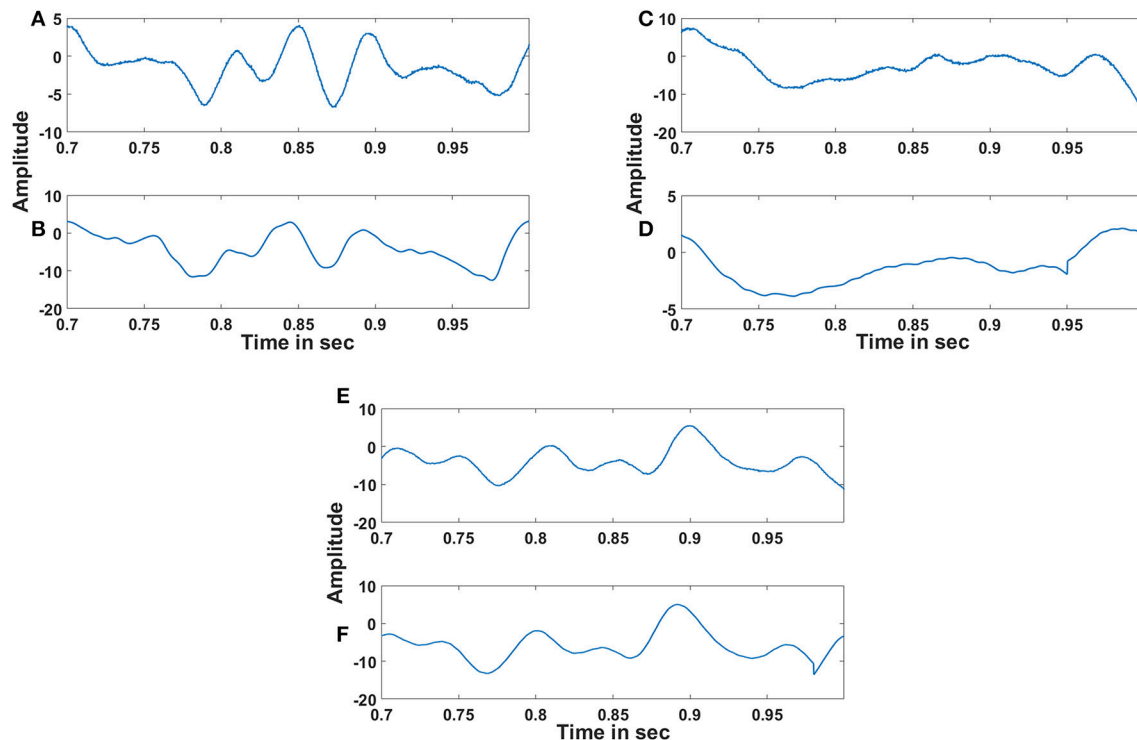
**FIGURE 7** | EEG signals from three channels: (Top–Bottom) figures show the EEG signals from C3, Cz and C4 channels respectively for 1 s duration. Voltages on y-axis are given in  $\mu$  volt.



**FIGURE 8** | Frequency spectrum of EEG FM signals and EEG MUX signal: (A–C) show the frequency spectrum of the EEG frequency modulated signals. (D,E) show the frequency spectrum of the composite MUX signal obtained from LAHN. It is vivid from the figures that the MUX signals cover the spectral information of the EEG FM signals and form a low dimensional representation of the same.



**FIGURE 9 |** Frequency spectrum of adaptive Hopf oscillators: (A,B) show the frequency spectrum of the two MUX signals. (C–E) show the frequency spectrum of three adaptive Hopf oscillators and it is evident from the figure that the adaptive Hopf oscillators separate out the frequency spectrum of the EEG FM signals.



**FIGURE 10 |** Reconstruction of EEG signals: (A), (C) and (E) show the original EEG signals recorded from C3, Cz and C4 respectively and (B), (D) and (F) show the reconstructed EEG signals of the respective channels by the oscillatory autoencoder model. The % of reconstruction error for each reconstruction is 5.8, 6.23, and 3.9% respectively. The absence of noise in the reconstructed signal and also the dimensionality reduction can influence the reconstruction error.

## Criticality of the Oscillations

Although in computational neuroscience literature, oscillatory neurons are not as common as rate-coded or spiking neuron models, oscillations figure prominently in experimental neurobiology. There exists a large corpus of experimental literature that correlates animal behavior with the aspects of neural oscillations (Buzsáki, 2002; Lisman and Buzsáki, 2008; Adhikari et al., 2010; Fell and Axmacher, 2011). Instances can be found from experimental neurobiology wherein all the major components of neural information processing viz., communication, representation and learning are implemented by neural oscillations. Colgin et al. (2009) reported that CA1 region of hippocampus communicates with Medial Entorhinal Cortex (MEC) via fast gamma synchronization (65–140 Hz) and with CA3 region via slow gamma synchronization (25–50 Hz) (Colgin et al., 2009). That is, by changing the frequency of the signal, it is possible to select the route by which communication takes place. Spatially distributed neurons can encode for several individual features of an object by synchronizing the neural discharges of the features, a phenomenon known as feature binding (Singer and Gray, 1995). For instance, the presentation of an optimally oriented bar gives rise to synchronized spiking of neurons, which are spatially distributed, in the area 17 of the visual cortex (Gray and Singer, 1987). Synchronization in the neural discharge is mirrored in the phase of the corresponding oscillatory LFP activity too. Hence there is a high correlation between the spike timing with the phase of the LFP oscillations. In case of feature binding, synchronization may not sometimes be evident from the spiking activity of the neurons, but the LFP activity shows robust phase synchronization (Alonso and Garcia-Austt, 1987; Buzsáki et al., 1992). Thus understanding the system dynamics in terms of oscillations becomes crucial. In the perspective of learning, a volley of high frequency pre-synaptic pulses with simultaneous depolarization at the postsynaptic side leads to Long Term Potentiation (LTP) (Bliss and Lomo, 1973; Lüscher and Malenka, 2012). These high frequency spikes can be correlated with the corresponding LFP oscillations. Hence the same LTP defined in terms of the spikes can be redefined using oscillatory LFP (Chauvette et al., 2012). Oscillations also have a pivotal role in cognition in both normal and pathological conditions. For example, the disconnectivity hypothesis of schizophrenia relates the disease symptoms to the dysfunction in the communication between different brain regions (Williams and Boksa, 2010). Gamma rhythm has been reported to have a role in the information transfer between the brain regions (Gray et al., 1989). In the early onset schizophrenic patients there is a reduction in the power of the gamma oscillation in the Prefrontal Cortex (PFC) a reason accounted for impaired working memory (Haenschel et al., 2009). Longer time scale oscillations like circadian rhythms are also known to play a critical role in major psychological disorders like bipolar disorder, depression, addiction (McClung, 2007; Alloy et al., 2015). Thus, from circadian to high gamma rhythms, oscillator models can be used to describe brain dynamics over a wide range of frequencies.

## Relation of Neural Information Transfer to Radio Communication Principles

The proposed work reinforces the hypothesis of information transfer between the brain regions to FM radio principles proposed by Hoppensteadt and Izhikevich (1998), that cortical areas communicate each other by making sure that their oscillations satisfy a resonant condition (Hoppensteadt and Izhikevich, 1998). They hypothesized that cortical oscillations are frequency modulated (FM) and, when the frequencies of two cortical areas match, they communicate by phase modulation. Thus, cortical communication is proposed to operate on the lines similar to FM radio. Although their paper proposed that signals can be frequency modulated and demodulated, it does not present how these concepts could actually be exploited to perform autoencoding, i.e., the input messages getting frequency modulated, multiplexed, demultiplexed and frequency demodulated. The proposed oscillatory autoencoder model realizes this concept by invoking the adaptive frequency and phase synchronization dynamics which take care of the frequency tuning to the incoming FM signal and hence offers a neurally plausible mechanism for the signal transmission and reconstruction (autoencoding) in the brain. This is achieved by the use of Hopf and Kuramoto dynamics. Both Kuramoto and Hopf oscillators have been previously used as models of neural oscillations in many instances (Cumin and Unsworth, 2007; Righetti et al., 2009). The Kuramoto model has been used to explain neuronal synchronization in large connected networks (Cumin and Unsworth, 2007) and especially building generative models of cortical oscillations (Breakspear et al., 2010). On the other hand, adaptive Hopf oscillators have been used for the generation of rhythmic output patterns such as central pattern generators involved in locomotion (Ijspeert et al., 2005). However, we have not come across any literature exploiting the phase synchronization properties of Kuramoto oscillators and adaptive frequency aspects of Hopf oscillator to model frequency multiplexing and demultiplexing. One of the interesting achievements of the proposed model is to show that Kuramoto–Hopf oscillator combination could act as a neural phase locked loop (PLL) which can be used to decode information from a given cortical region.

## Possible Applications of the Model

Autoencoder networks are usually constructed out of rate coded neurons, though in the recent times autoencoder networks with spiking neurons have also been proposed (Burbank, 2015). In its simplest form, a rate-coded autoencoder is a feedforward network with a single hidden layer and is trained such that the target output is the same as the input; the hidden layer has fewer neurons than the input or the output layer. Then the hidden layer learns to represent the input using fewer dimensions and therefore achieves dimensionality reduction of the input space (Hinton and Salakhutdinov, 2006). A similar reduction is achieved in the proposed oscillatory model since the hidden layer, LAHN, is of lower dimension than the input layer.

The connection between Hebbian learning rule and Principal Component Analysis (PCA) is not a new idea since Oja has previously shown how a linear neuron adapting its synaptic weight connections using Hebbian learning rule can converge to the first principal component of the input data (Oja, 1989). This was further extended by Sanger using an asymmetric Generalized Hebbian Algorithm (GHA) learning rule that makes the network to learn the first  $n$  principal components (Sanger, 1989) instead of just one principal component. Hebbian/anti-Hebbian network also comes under the category of subspace learning network. This type of network, reduces the input data dimension by learning the principal subspace of the input data (Földiák, 1990; Hu et al., 2015; Pehlevan et al., 2015). Other neural networks in this line are subspace network, Rubner's network (Rubner and Tavan, 1989; Rubner and Schulten, 1990) etc. Although these networks were initially modeled to explain the computations behind the processing of streaming sensory inputs, the synaptic plasticity rules based on the local activity of the neurons neuronal activity were postulated rather than derived from a cost function (Földiák, 1990). This gap was further bridged by computing the local learning rules from a principled cost function (Hu et al., 2015; Pehlevan et al., 2015). Changing the non-linearity of the neuronal activation function explained the potentiality of these networks in extracting the higher order moments of the input data and hence qualified them as the neural architectures for Independent Component Analysis (ICA) (Oja, 1997). The aforementioned studies prove the criticality of this type of network in various applications that include subspace learning, source separation problem, dimensionality reduction etc.

This dimensionality reduction has further implications especially in EEG processing. The model reconstructs the original EEG signals from their lower dimensional LAHN representations. This means that these LAHN signals can serve as the reliable representations especially for high channel EEG signals. These representations could potentially be useful in BCI related processing such as classification of EEG signals, feature clustering, movement signature detection etc. The EEG signals used for the model simulation are two class motor imagery signals which are of particular interest in BCI application. Hence the proposed model not only provides a biologically plausible explanation for the information transfer in the brain but also shows its possible potential application in BCI related EEG processing. Another important feature that makes the current model suitable for EEG processing is its ability to average out the noise present in the input signal. As shown in the results section, the input EEG signals have high frequency ripples in its original form which is further averaged out to produce a smooth reconstructed signal as the output from the model (Figure 10). This could possibly be due to the large temporal scale Hebbian learning that happens in the hidden LAHN layer which could thus average out the noise present in the input.

### Future Extensions of the Proposed Model

A possible extension of the current model could be to add additional circuitry that will enable routing of the signal from

the  $i$ th input channel to the  $j$ th output channel. It must be possible to choose the input/output channels to be coupled through another layer that projects to the current LAHN layer that performs multiplexing of FM signals. In such an extended model, the LAHN layer and the additional circuitry for route selection can be compared to the functions of the thalamus with respect to cortico-thalamic information processing. Hence the proposed model serves as the proof of principle for the potentiality of the oscillatory neural networks in information transfer i.e., encoding and decoding of real world signals using the principles of modulation, MUX, frequency adaptation and phase synchronization and also shows its possible potential role in EEG related applications. Another direction that the proposed model could possibly take is to pick the brain components from the EEG signals. By *brain component* we mean the sources inside the brain responsible for the generation of the EEG signal. Current approaches like Independent Component Analysis (ICA) require manual selection of components which has a source inside the brain for further analysis. We envisage that using a hierarchical network of LAHN, we could possibly isolate the brain components better due to its inherent ability to filter out noise. As a future work, we envisage that the current model could be possibly used (may be by invoking minor changes) to study the EEG related phenomena like mu band Event Related Desynchronization (ERD), Visual Evoked Event Related Potential (ERP) etc which can possibly shed light on the neural principles behind the occurrence of these phenomena.

## CONCLUSION

We propose a hybrid oscillatory network model that performs the function of an autoencoder. Using this network, we are able to encode the information onto oscillations, reduce the dimensionality of information and effectively decode them using a neural phase locked loop. The model was successfully applied to both synthetic as well as real world EEG signals. Hence the proposed model shows an oscillatory neural framework in describing information transfer in the brain. By reconstructing the EEG signals from its abstract representations in the hidden layer we have shown the model's ability in better feature extraction of EEG signal which is a critical part in EEG processing. Finally, we conclude that exploring the universality of oscillator networks would open avenues for developing an entirely new class of neural network models that describe brain function in terms of oscillatory properties— amplitude, frequency, and phase. The whole motivation of this work was to show a proof of principle for the potentiality of the oscillatory networks in other domains where usually rate coded or spiking neurons were used. In the future, we plan to apply the model to a wider variety of real world time series signals.

## DATA AVAILABILITY STATEMENT

All the simulations are done in MATLAB R2016a and the code is made available in Model DB repository. URL is



<http://senselab.med.yale.edu/ModelDB/showModel.cshtml?model=243595>

## AUTHOR CONTRIBUTIONS

All authors contributed equally to the work. KS performed designing, coding, analysis of the model, and manuscript preparation. VM performed analysis of the model, and

manuscript preparation. VC performed designing the model, analysis of the model and manuscript preparation.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2018.00052/full#supplementary-material>

## REFERENCES

- Adhikari, A., Topiwala, M. A., and Gordon, J. A. (2010). Synchronized activity between the ventral hippocampus and the medial prefrontal cortex during anxiety. *Neuron* 65, 257–269. doi: 10.1016/j.neuron.2009.12.002
- Alloy, L. B., Nusslock, R., and Boland, E. M. (2015). The development and course of bipolar spectrum disorders: an integrated reward and circadian rhythm dysregulation model. *Annu. Rev. Clin. Psychol.* 11, 213–250. doi: 10.1146/annurev-clinpsy-032814-112902
- Alonso, A., and Garcia-Austt, E. (1987). Neuronal sources of theta rhythm in the entorhinal cortex of the rat. *Exp. Brain Res.* 67, 493–501. doi: 10.1007/BF00247282
- Auer, P., Burgsteiner, H., and Maass, W. (2008). A learning rule for very simple universal approximators consisting of a single layer of perceptrons. *Neural Netw.* 21, 786–795. doi: 10.1016/j.neunet.2007.12.036
- Bliss, T. V., and Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J. Physiol.* 232, 331–356. doi: 10.1113/jphysiol.1973.sp010273
- Breakspear, M., Heitmann, S., and Daffertshofer, A. (2010). Generative models of cortical oscillations: neurobiological implications of the kuramoto model. *Front. Hum. Neurosci.* 4:190. doi: 10.3389/fnhum.2010.00190
- Burbank, K. S. (2015). Mirrored STDP Implements Autoencoder Learning in a Network of Spiking Neurons. *PLoS Comput. Biol.* 11:e1004566. doi: 10.1371/journal.pcbi.1004566
- Buzsáki, G. (2002). Theta oscillations in the hippocampus. *Neuron* 33, 325–340. doi: 10.1016/S0896-6273(02)00586-X
- Buzsáki, G., Anastassiou, C. A., and Koch, C. (2012). The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nat. Rev. Neurosci.* 13, 407–420. doi: 10.1038/nrn3241
- Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745
- Buzsáki, G., Horváth, Z., Urioste, R., Hetke, J., and Wise, K. (1992). High-frequency network oscillation in the hippocampus. *Science* 256, 1025–1027.
- Campbell, S. R., Wang, D. L., and Jayaprakash, C. (1999). Synchrony and desynchrony in integrate-and-fire oscillators. *Neural Comput.* 11, 1595–1619.
- Chauvette, S., Seigneur, J., and Timofeev, I. (2012). Sleep oscillations in the thalamocortical system induce long-term neuronal plasticity. *Neuron* 75, 1105–1113. doi: 10.1016/j.neuron.2012.08.034
- Colgin, L. L., Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., et al. (2009). Frequency of gamma oscillations routes flow of information in the hippocampus. *Nature* 462, 353–357. doi: 10.1038/nature08573
- Cumin, D., and Unsworth, C. P. (2007). Generalising the Kuramoto model for the study of neuronal synchronisation in the brain. *Phys. D Nonlin. Phenomena* 226, 181–196. doi: 10.1016/j.physd.2006.12.004
- David, O., and Friston, K. J. (2003). A neural mass model for MEG/EEG: coupling and neuronal dynamics. *NeuroImage* 20, 1743–1755. doi: 10.1016/j.neuroimage.2003.07.015
- Faloutsos, C., Ranganathan, M., and Manolopoulos, Y. (1994). “Fast subsequence matching in time-series databases,” in *SIGMOD '94 Proceedings of the 1994 ACM SIGMOD International Conference on Management of Data* (New York, NY).
- Fell, J., and Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nat. Rev. Neurosci.* 12, 105–118. doi: 10.1038/nrn2979
- Földiák, P. (1990). Forming sparse representations by local anti-Hebbian learning. *Biol. Cybern.* 64, 165–170. doi: 10.1007/BF02331346
- Frasconi, P., Gori, M., and Soda, G. (1995). Recurrent neural networks and prior knowledge for sequence processing: a constrained nondeterministic approach. *Knowledge Based Syst.* 8, 313–332. doi: 10.1016/0950-7051(96)81916-2
- Ghosh-Dastidar, S., and Adeli, H. (2009). Spiking neural networks. *Int. J. Neural Syst.* 19, 295–308. doi: 10.1142/S0129065709002002
- Gray, C. M., König, P., Engel, A. K., and Singer, W. (1989). Oscillatory responses in cat visual-cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature* 338, 334–337. doi: 10.1038/338334a0
- Gray, C., and Singer, W. (1987). Stimulus-dependent neuronal oscillations in the cat visual cortex area. *Neuroscience* 17 (Suppl.) 22, 434.
- Guillery, R. W., and Sherman, S. M. (2002). Thalamic relay functions and their role in corticocortical communication: generalizations from the visual system. *Neuron* 33, 163–175. doi: 10.1016/S0896-6273(01)00582-7
- Haenschel, C., Bittner, R. A., Waltz, J., Haertling, F., Wibral, M., Singer, W., et al. (2009). Cortical oscillatory activity is critical for working memory as revealed by deficits in early-onset schizophrenia. *J. Neurosci.* 29, 9481–9489. doi: 10.1523/JNEUROSCI.1428-09.2009
- Haykin, S. S., Moher, M., and Song, T. (1989). *An Introduction to Analog and Digital Communications*. New York, NY: Wiley.
- Heitmann, S., Boonstra, T., Gong, P. L., Breakspear, M., and Ermentrout, B. (2015). The rhythms of steady posture: motor commands as spatially organized oscillation patterns. *Neurocomputing* 170, 3–14. doi: 10.1016/j.neucom.2015.01.088
- Hinton, G. E., and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science* 313, 504–507. doi: 10.1126/science.1127647
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558.
- Hoppensteadt, F. C., and Izhikevich, E. M. (1998). Thalamo-cortical interactions modeled by weakly connected oscillators: could the brain use FM radio principles? *BioSystems* 48, 85–94.
- Hoppensteadt, F. C., and Izhikevich, E. M. (2000). Pattern recognition via synchronization in phase-locked loop neural networks. *IEEE Trans. Neural Netw.* 11, 734–738. doi: 10.1109/72.846744
- Hu, T., Pehlevan, C., and Chklovskii, D. B. (2015). “A hebbian/anti-hebbian network for online sparse dictionary learning derived from symmetric matrix factorization,” in *2014 48th Asilomar Conference on Signals, Systems and Computers* (Pacific Grove, CA: IEEE), 613–619.
- Hubel, D. H. (1995). *Eye, Brain, and Vision*. New York, NY: Scientific American Library; Scientific American Books.
- Ijspeert, A. J. (2008). Central pattern generators for locomotion control in animals and robots: a review. *Neural Netw.* 21, 642–653. doi: 10.1016/j.neunet.2008.03.014
- Ijspeert, A. J., Crespi, A., and Cabelguen, J. M. (2005). Simulation and robotics studies of salamander locomotion: applying neurobiological principles to the control of locomotion in robots. *Neuroinformatics* 3, 171–195. doi: 10.1385/NI:3:3:171
- Izhikevich, E. M. (2003). Simple model of spiking neurons. *IEEE Trans. Neural Netw.* 14, 1569–1572. doi: 10.1109/TNN.2003.820440
- Izhikevich, E. M. (2004). Which model to use for cortical spiking neurons? *IEEE Trans. Neural Netw.* 15, 1063–1070. doi: 10.1109/TNN.2004.832719

- Keogh, E., Chakrabarti, K., Mehrotra, S., and Pazzani, M. (2001a). Locally adaptive dimensionality reduction for indexing large time series databases. *Sigmod. Record*. 30, 151–162. doi: 10.1145/568518.568520
- Keogh, E., Chakrabarti, K., Pazzani, M., and Mehrotra, S. (2001b). Dimensionality reduction for fast similarity search in large time series databases. *Knowl. Inf. Syst.* 3, 263–286. doi: 10.1007/PL00011669
- Kohonen, T. (1998). The self-organizing map. *Neurocomputing* 21, 1–6. doi: 10.1016/S0925-2312(98)00030-7
- Kuramoto, Y. (1984). *Chemical Oscillations, Waves, and Turbulence*. Tokyo: Springer.
- Lawrence, S., Giles, C. L., Tsoi, A. C., and Back, A. D. (1997). Face recognition: a convolutional neural-network approach. *IEEE Trans. Neural Netw.* 8, 98–113. doi: 10.1109/72.554195
- Leeb, R., Brunner, C., Müller-Putz, G., Schlögl, A., and Pfurtscheller, G. (2008). *BCI Competition 2008–Graz data Set B*. Berlin: Graz University of Technology.
- Leeb, R., Friedman, D., Müller-Putz, G. R., Scherer, R., Slater, M., and Pfurtscheller, G. (2007a). Self-paced (asynchronous) BCI control of a wheelchair in virtual environments: a case study with a tetraplegic. *Comput. Intell. Neurosci.* 2007:79642. doi: 10.1155/2007/79642
- Leeb, R., Lee, F., Keinrath, C., Scherer, R., Bischof, H., and Pfurtscheller, G. (2007b). Brain-computer communication: Motivation, aim, and impact of exploring a virtual apartment. *IEEE Trans. Neural Syst. Rehabil. Eng.* 15, 473–482. doi: 10.1109/TNSRE.2007.906956
- Lin, J., Keogh, E., Wei, L., and Lonardi, S. (2007). Experiencing SAX: a novel symbolic representation of time series. *Data Min. Knowl. Discov.* 15, 107–144. doi: 10.1007/s10618-007-0064-z
- Lippmann, R. P. (1989). Review of neural networks for speech recognition. *Neural Comput.* 1, 1–38. doi: 10.1162/neco.1989.1.1.1
- Lisman, J., and Buzsáki, G. (2008). A neural coding scheme formed by the combined function of gamma and theta oscillations. *Schizophr. Bull.* 34, 974–980. doi: 10.1093/schbul/sbn060
- Logothetis, N. K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412, 150–157. doi: 10.1038/35084005
- Lüscher, C., and Malenka, R. C. (2012). NMDA receptor-dependent long-term potentiation and long-term depression (LTP/LTD). *Cold Spring Harb. Perspect. Biol.* 4:a005710. doi: 10.1101/cshperspect.a005710
- Maass, W. (1997a). Fast sigmoidal networks via spiking neurons. *Neural Comput.* 9, 279–304.
- Maass, W. (1997b). Networks of spiking neurons: the third generation of neural network models. *Neural Netw.* 10, 1659–1671.
- McClung, C. A. (2007). Circadian rhythms, the mesolimbic dopaminergic circuit, and drug addiction. *Sci. World J.* 7, 194–202. doi: 10.1100/tsw.2007.213
- Oja, E. (1989). Neural networks, principal components, and subspaces. *Int. J. Neural Syst.* 1, 61–68. doi: 10.1142/S0129065789000475
- Oja, E. (1997). The nonlinear PCA learning rule in independent component analysis. *Neurocomputing* 17, 25–45.
- Omotayo, O. (1985). Effective transfer of data through multiplexing. *Data Process* 27, 26–28.
- Pehlevan, C., Hu, T., and Chklovskii, D. B. (2015). A hebbian/anti-hebbian neural network for linear subspace learning: a derivation from multidimensional scaling of streaming data. *Neural Comput.* 27, 1461–1495. doi: 10.1162/NECO\_a\_00745
- Percival, D. B., and Walden, A. T. (2006). *Wavelet Methods for Time Series Analysis*. Cambridge: Cambridge University Press.
- Pfurtscheller, G., and Neuper, C. (2001). Motor imagery and direct brain-computer communication. *Proc. IEEE* 89, 1123–1134. doi: 10.1109/5.939829
- Ramoser, H., Müller-Gerking, J., and Pfurtscheller, G. (2000). Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. Rehabil. Eng.* 8, 441–446. doi: 10.1109/86.895946
- Righetti, L., Buchli, J., and Ijspeert, A. J. (2006). Dynamic Hebbian learning in adaptive frequency oscillators. *Phys. D Nonlin. Phenomena* 216, 269–281. doi: 10.1016/j.physd.2006.02.009
- Righetti, L., Buchli, J., and Ijspeert, A. J. (2009). Adaptive frequency oscillators and applications. *Open Cybern. Syst. J.* 3, 64–69. doi: 10.2174/1874110X00903010064
- Ruck, D. W., Rogers, S. K., and Kabrisky, M. (1990). Feature selection using a multilayer perceptron. *J. Neural Netw. Comput.* 2, 40–48.
- Rubner, J., and Schulten, K. (1990). Development of feature detectors by self-organization. A network model. *Biol. Cybern.* 62, 193–199.
- Rubner, J., and Tavan, P. (1989). A self-organizing network for principal-component analysis. *EPL Europhys. Lett.* 10, 693.
- Sanger, T. D. (1989). Optimal unsupervised learning in a single-layer linear feedforward neural network. *Neural Netw.* 2, 459–473. doi: 10.1016/0893-6080(89)90044-0
- Schmidhuber, J. (2015). Deep learning in neural networks: an overview. *Neural Netw.* 61, 85–117. doi: 10.1016/j.neunet.2014.09.003
- Singer, W., and Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annu. Rev. Neurosci.* 18, 555–586. doi: 10.1146/annurev.ne.18.030195.003011
- Trappenberg, T. (2003). Continuous attractor neural networks. *Rev. Dev. Biol. Inspir. Comput.* 398–425.
- Wang, D., and Terman, D. (1995). Locally excitatory globally inhibitory oscillator networks. *IEEE Trans. Neural Netw.* 6, 283–286. doi: 10.1109/72.363423
- Whittingstall, K., and Logothetis, N. K. (2009). Frequency-band coupling in surface EEG reflects spiking activity in monkey visual cortex. *Neuron* 64, 281–289. doi: 10.1016/j.neuron.2009.08.016
- Williams, S., and Boksa, P. (2010). Gamma oscillations and schizophrenia. *J. Psychiatry Neurosci.* 35, 75–77. doi: 10.1503/jpn.100021
- Wilson, S. J. (2017). Data representation for time series data mining: time domain approaches. *WIRs Comput. Statist.* 9:e1392. doi: 10.1002/wics.1392

**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 © 2018 Soman, Muralidharan and Chakravarthy. 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) and the copyright owner(s) 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.



# Sensorimotor Synchronization With Auditory and Visual Modalities: Behavioral and Neural Differences

Daniel C. Comstock<sup>1</sup>, Michael J. Hove<sup>2</sup> and Ramesh Balasubramaniam<sup>1\*</sup>

<sup>1</sup> Cognitive and Information Sciences, University of California, Merced, Merced, CA, United States, <sup>2</sup> Department of Psychological Science, Fitchburg State University, Fitchburg, MA, United States

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

Edward W. Large,  
University of Connecticut,  
United States

David Ian Anderson,  
San Francisco State University,  
United States

Todd Troyer,  
University of Texas at San Antonio,  
United States

### \*Correspondence:

Ramesh Balasubramaniam  
ramesh@ucmerced.edu

**Received:** 05 March 2018

**Accepted:** 19 June 2018

**Published:** 18 July 2018

### Citation:

Comstock DC, Hove MJ and  
Balasubramaniam R (2018)  
Sensorimotor Synchronization With  
Auditory and Visual Modalities:  
Behavioral and Neural Differences.  
*Front. Comput. Neurosci.* 12:53.  
doi: 10.3389/fncom.2018.00053

It has long been known that the auditory system is better suited to guide temporally precise behaviors like sensorimotor synchronization (SMS) than the visual system. Although this phenomenon has been studied for many years, the underlying neural and computational mechanisms remain unclear. Growing consensus suggests the existence of multiple, interacting, context-dependent systems, and that reduced precision in visuo-motor timing might be due to the way experimental tasks have been conceived. Indeed, the appropriateness of the stimulus for a given task greatly influences timing performance. In this review, we examine timing differences for sensorimotor synchronization and error correction with auditory and visual sequences, to inspect the underlying neural mechanisms that contribute to modality differences in timing. The disparity between auditory and visual timing likely relates to differences in the processing specialization between auditory and visual modalities (temporal vs. spatial). We propose this difference could offer potential explanation for the differing temporal abilities between modalities. We also offer suggestions as to how these sensory systems interface with motor and timing systems.

**Keywords:** sensorimotor synchronization, timing, rhythm, visual perception, auditory perception

## INTRODUCTION

Many behavioral studies have examined human timing ability in tasks of sensorimotor synchronization (SMS) where subjects synchronize their movements to an external rhythm.

Comparisons between auditory metronomes and visual flashing metronomes reveal that movement synchronization is less variable and can occur at faster rates with auditory metronomes (Chen et al., 2002; Repp, 2003; Repp and Penel, 2004; Lorås et al., 2012). However, visuo-motor synchronization greatly improves when synchronizing with a *moving* periodic visual metronome (Hove et al., 2010). Adding a changing velocity profile to the moving visual metronome further reduces variability in SMS tapping (Hove et al., 2013a; Iversen et al., 2015), and Gan et al. (2015) suggests that a more realistic velocity profile can bring visual SMS to be as temporally precise as auditory SMS, at moderate but not fast tempi. While most studies of SMS look at finger tapping, others have included synchronized circle drawing, gait, dancing, and eye movements in the context of modality-specific timing effects (e.g., Repp and Su, 2013).

Studies on auditory and visual interference also suggest auditory timing is more prominent. When concurrent auditory metronomes and visual flashing metronomes are presented out-of-phase, the auditory sequences interfere with visuomotor timing, but not vice versa (Repp and Penel, 2002, 2004). The interference effect is considerably reduced with moving visual metronomes and is tied to training and experience as the auditory dominance is stronger in musicians and weaker in video gamers (Hove et al., 2013a). Similarly, auditory cues can improve visual temporal discrimination (Morein-Zamir et al., 2003; Parise and Spence, 2008). This effect only holds for the temporal domain however, as the visual system dominates when auditory and visual stimuli conflict in the spatial domain; spatial dominance in the visual modality is apparent in the well-known “ventriloquist effect” (Vroomen et al., 2001).

## ROLE OF ERROR CORRECTION IN TIMING

Error correction is a crucial component of any SMS task. By inducing perturbations and errors in SMS, we can gain insight into the underlying timing mechanisms. A common method to induce errors in a SMS task is to occasionally perturb an otherwise isochronous metronome (Repp, 2000, 2001a,b; Praamstra et al., 2003; Repp and Keller, 2004; Jang et al., 2016; Jantzen et al., 2018). Error correction in SMS can be broken down into two distinct mechanisms: a phase-correction mechanism for correcting errors in relative phase, and a period-correction mechanism that corrects changes to the internal timekeeper period (Repp, 2001b; Repp and Keller, 2004). Period corrections require conscious awareness of the error as it involves a conscious updating of the internal rhythm; while a phase correction can happen even with errors too small for conscious awareness and does not involve updating the central timekeeper period and so is considered a more peripheral process than period correction (Repp, 2001b, 2005). An error corrected under the phase-correction mechanism is typically a gradual adjustment that occurs over several beats, while an error corrected under the period-correction mechanism will be evidenced by a pronounced correction, usually followed by a more gradual phase-correction-like pattern after the initial large correction (Repp, 2001b).

While error correction has been well documented in auditory SMS, relatively little work has investigated error correction in visual SMS. In a recent study comparing error correction for auditory and flashing visual sequences, we observed error corrections for perturbations in the auditory condition that were modulated by the direction of the perturbations, but no such modulation was found for perturbations in the visual condition (Comstock and Balasubramaniam, 2017a). This suggests the visual system may not engage in the same SMS timing mechanisms as the auditory system. Additional evidence for a discrepancy in error correction for auditory and visual sequences can be gleaned from the autocorrelation structure of adjacent taps: unlike auditory SMS, tapping with visual flashes does not produce a negative lag1 autocorrelation that can indicate of the

presence of a robust central timekeeping and error-correction mechanism (Hove and Keller, 2010). However, visuomotor synchronization with moving and apparent-motion metronomes do produce a negative lag1 autocorrelation, suggesting that a moving visual metronome may engage error correction (Hove and Keller, 2010; Hove et al., 2010); note that negative lag1 autocorrelation does not necessarily stem from error correction and can arise from other timing factors (e.g., Wing and Kristofferson, 1973). It remains unclear if error correction will occur with perturbations in moving visual metronomes or with larger phase perturbations in a flashing visual metronome.

## UNDERLYING PHYSIOLOGY OF THE AUDITORY AND VISUAL TIMING SYSTEM

### Brain Networks Involved in Timing Activity

Investigating the neural underpinnings in auditory and visual timing is a massive undertaking due to the many different timing subprocesses and tasks, including: SMS, interval timing, rhythm perception, timing recall, time perception, etc.. Excellent reviews of the brain mechanisms involved in various timing activities include: a review of neural activity in music production (Zatorre et al., 2007); a review of neural activity involved in time perception (Wiener et al., 2010); and an overview of neural activation in SMS as part of a larger review of SMS (Repp and Su, 2013). This body of work consistently demonstrates that temporal processing across tasks and sensory modalities relies heavily on the motor system. This motor network includes the supplemental motor area (SMA), primary motor cortex, lateral premotor cortex, anterior cingulate, basal ganglia, and cerebellum (Repp and Su, 2013). Auditory rhythm perception activates the motor system and is closely linked to movement (Janata et al., 2012; Iversen and Balasubramaniam, 2016; Ross et al., 2016a,b). The SMA is also strongly implicated in motor timing (Coull et al., 2016; Merchant and Yarrow, 2016), and along with the pre-SMA could be a hub of motor timing (Schwartz et al., 2012). Subcortical regions are especially active during sub-second time perception (Wiener et al., 2010), sub-second interval timing (Repp and Su, 2013), and rhythm timing (Grahn and Rowe, 2009; Wiener et al., 2010; Coull et al., 2011; Teki et al., 2011; Hove et al., 2013b). There is evidence of a dorsal auditory stream connecting the auditory cortex to the motor cortex through the posterior parietal cortex that plays a role in rhythm perception (Patel and Iversen, 2014; Ross et al., 2018). Interestingly this dorsal stream is also implicated in visual and tactile rhythm perception (Araneda et al., 2017; Rauschecker, 2017), adding to the idea of a common timing system tied to the motor system. Further evidence of the common timing system is found in a study of auditory and visual synchronization that dissociated modality and tapping stability – putamen activation was highest when synchronizing to auditory beeps, moderate with a frequency-modulated siren and with a moving visual metronome, and lowest with a flashing visual metronome, closely paralleling behavioral performance (Hove et al., 2013b).



While visual SMS activates many of the same motor regions as auditory SMS (Hove et al., 2013b; Araneda et al., 2017), some activations are specific to the visual system. The visual cortex shows activity related to interval timing that follows the expected scalar property, such that size of timing errors measured in the visual cortex scale in proportion to size of the interval being timed as predicted by Weber's law (Shuler, 2016). Additionally, Zhou et al. (2014) found evidence that visual feature processing in the early visual cortex can contribute to duration perception, furthering the notion that at least some timing information is processed independently within the visual cortex. Additionally, in visual rhythm perception, the visual cortex plays a role predicting rhythmic onsets (Comstock and Balasubramaniam, 2017b, 2018). The additional activations with visual timing tasks, taken together with behavioral results, suggest the timing accuracy in visual processing may be compared to the auditory system due to the additional computational demands of processing the higher complexity of visual spatial information along with temporal information.

## Role of Cortical Oscillations in Timing Encoding and Spreading Information Across the Brain

In addition to looking at the networks and regions involved in temporal processing, a growing body of work shows the role of cortical oscillations in encoding timing across multiple frequency bands. Cortical oscillations play a role in connecting regions across the brain, with higher frequencies utilized for localized interaction and lower frequencies for longer range interaction (Sarnthein et al., 1998; Von Stein and Sarnthein, 2000). This pattern of oscillations is used to connect and calibrate disparate timing systems in the brain (Gupta and Chen, 2016). Oscillations relating to timing appear to arise from multiple context-specific timing systems in the brain (Wiener and Kanai, 2016). The question is then how these functionally and anatomically disparate systems integrate and interact. It appears that oscillations from different timing systems are coordinated within the striatum (Matell and Meck, 2004; Gu et al., 2015).

Beta band activity ( $\sim 20$  Hz) is tied to the motor system and several studies indicate beta's role in predicting timing of auditory rhythms (Fujioka et al., 2009, 2012, 2015). Additionally, beta activity reflects top-down imposition of metrical structure on auditory rhythms (Iversen et al., 2009). Recently, beta activity has also been linked to timing predictions within the visual system in response to visual rhythms (Comstock and Balasubramaniam, 2017b).

With rhythm perception, evidence shows that internal oscillations arise to match the fundamental frequency of the rhythm, and frequency of the meter (Nozaradan et al., 2011), as well as to the frequency of imagined rhythms (Okawa et al., 2017). These findings align with the Neural Resonance Theory that posits neural rhythms synchronize to auditory rhythms, and these neural rhythms can influence attention, expectancy, and motor planning (Large and Snyder, 2009). As of yet, it is unclear

if this same neural resonance to meter would arise with visual stimuli.

## Neural Underpinnings of Error Correction

The neural correlates of error correction reveal more evidence for multiple interacting and overlapping timing mechanisms. Error detection of timing perturbations in auditory SMS tasks modulates the P1, N1, and N2 auditory ERP components depending on both the size and direction of the perturbation (Praagstra et al., 2003; Jang et al., 2016). Jantzen et al. (2018) also found a theta response stemming from the Pre-SMA and anterior cingulate for error detection, an increase in theta coupling between the SMA and the motor cortex for late perturbations. In visual error detection, the visual P1 component is reduced in latency only for large late perturbations (Comstock and Balasubramaniam, 2017a). Each of these instances show cortical activation specific to a type of perturbation, although these effects are generally limited to larger perturbations.

Smaller perturbations that elicit a phase-correction response are believed to be driven primarily by subcortical mechanisms. Applying repetitive TMS to downregulate motor and premotor cortices produced no effect on phase correction (Dumas et al., 2005), whereas phase-correction was impaired by repetitive TMS to the cerebellum (Bijsterbosch et al., 2011). This fits with the suggestion that phase-correction is primarily subcortical based on evidence from how rapidly the movement trajectory changes after a perturbation (Hove et al., 2014). A possible network that exhibits the rapid timing required for the phase-correction response is a cortico-striatal circuit connecting the cerebellum to the SMA-striatal network via the thalamus (Kotz et al., 2016).

The data on the neural underpinnings of error correction suggest multiple timing systems, each with specific roles, yet able to coordinate for rapid response. Commensurate with this idea is work suggesting the basal ganglia integrates various timing systems through oscillation comparators (Matell and Meck, 2004; Gu et al., 2015). The limited data on visual error correction, however, leave open how well this network can interface with the visual timing systems.

## EVIDENCE THE AUDITORY SYSTEM HAS PRIVILEGED ACCESS TO TIMING SYSTEMS

Considering the auditory system's timing advantage along with the prominence of the motor system in timing processing, we suggest that the auditory system's advantage in timing stems from its stronger coupling to the motor system. Auditory timing compared to visual timing tasks often yield more activation in motor structures, such as the SMA and premotor cortex (Jäncke et al., 2000). Even when visual SMS tasks employed the modality-appropriate *moving* visual metronomes, audio-motor synchronization with auditory beeps yielded greater activation in the putamen (Hove et al., 2013b). Likewise, priming a visual rhythm with a similar auditory rhythm resulted in increased putamen activation compared to a visual rhythm

alone, while a visual rhythm yielded no priming effect on an auditory rhythm (Grahn et al., 2011). The finding that the increased visual synchronization ability provided by a bouncing ball does not transfer to purely perceptual rhythm perception provides further evidence of the role of motor coupling in timing tasks (Silva and Castro, 2016). Additionally, the privileged link between auditory and motor systems can be seen in Parkinson's disease, a disorder that impairs movement due to cell loss within the basal ganglia (Davie, 2008). For example, Parkinsonian gait can improve when cued by an external rhythm, and these interventions are more effective when synchronizing with auditory metronomes than with flashing visual metronomes (Rochester et al., 2005; Arias and Cudeiro, 2008).

Visual timing activities recruit timing centers within the visual system that, based on behavioral results, are less precise compared to the auditory timing system. In Jäncke et al. (2000), visual timing tasks resulted in increased activity in the right superior cerebellum, vermis, and right inferior parietal lobe compared to auditory timing tasks. Visual timing tasks also recruit areas MT, V5, and the superior parietal lobe, tying into the dorsal visual stream (Jantzen et al., 2005), and visual rhythm perception induces increased beta activity at event onsets arising from the visual cortex (Comstock and Balasubramaniam, 2017b). It is unclear if these timing activations in the visual system are the result of compensating for a weaker connection to the motor timing system. It may be that the temporal processing in the visual system is additional processing of visual information required to interface with the motor system.

While differences in coupling strength to the motor system are crucial for modality timing differences, other factors are likely. To that end, it is clear that the visual system is able to pick out high speed temporal information, for example, V1 will phase lock its input/output to up to a 100 Hz visual flashing stimuli (Williams et al., 2004). This suggests that entrainment is not easily transferred to the systems involved in time/rhythm perception, especially at the time frame usually involved in rhythm perception, indicating that the issue may be one of translation. A likely place for that translation would be within the dorsal pathway, which has been found to have neurons with high temporal resolution in macaques, with higher temporal resolution in the auditory dorsal stream (Rauschecker, 2017). If there is a higher temporal resolution of the auditory dorsal stream than in the visual dorsal stream, then it may give explanation as to why the visual system cannot synchronize at the higher frequencies achieved by the auditory system. Of course, it cannot be ruled out that the difference in temporal resolution is due to different levels of timing precision available to the dorsal stream. Reduced timing precision in the visual stream may be caused by increased necessary processing due to richer sensory input of the visual system compared to the auditory system. Indeed, greater processing requirements and longer processing time may help to account for the inability of the visual system to allow for synchronization at the higher tempos allowed by the auditory system.

## ROLE OF THE VESTIBULAR-TACTILE-SOMATOSENSORY SYSTEM

Another link between auditory and motor systems is that auditory rhythm perception may be tied to the vestibular-tactile-somatosensory (VTS) system, which is important for movement and dance, and therefore closely tied to the motor system and attuned to timing (Todd and Lee, 2015). In addition to its ties for movement, the VTS system is clearly tied to the auditory system with regards to rhythm perception (Phillips-Silver and Trainor, 2005, 2007, 2008; Trainor et al., 2009), and through common neural activation (Araneda et al., 2017). These ties between the auditory and VTS system may be an additional factor in the dominance of the auditory system in the temporal domain.

Since VTS rhythms are ubiquitous in fetal life through the mother's gait, heart rate, breathing, etc., and since these networks are tied into auditory rhythm systems, it is likely that the VTS system is heavily tied into the timing systems used in auditory rhythm perception and in motor rhythm production (Provasi et al., 2014). This is further strengthened by the fact that movement and rhythms are linked and proprioception (part of the VTS system) plays a large role in perception of rhythms that is tied into auditory rhythm perception and production (Trainor et al., 2009). Interactions between the VTS system with visual rhythm perception remains mostly unexplored at this point however, so it is unclear how much this system plays a supramodal role in the timing involved in rhythm perception/production, or if it is only tied to the auditory and motor rhythm timing systems. Further research in this area is needed to answer these questions.

## EVOLUTIONARY ORIGINS OF SENSORIMOTOR SYNCHRONIZATION

In an evolutionary context, it makes sense that auditory and motor systems would be tightly interconnected. First, rhythms in language are critical for both perception and production and may be a driver of SMS ability (Patel, 2006). Beyond language, matching movement to sound is a necessary result of human evolution that allows for the social and cultural inclination of humanity via music (Hagen and Bryant, 2003; Brown and Jordania, 2013). Dance is also tightly connected with music and culture and can provide a further explanatory account of human SMS capability and the connection between the motor and auditory systems (Fitch, 2016; Iversen, 2016; Laland et al., 2016; Ravignani and Cook, 2016).

Beyond humans, common adaptations appear to increase SMS ability in several non-human species capable of some level of audio-motor entrainment such as parrots (Patel et al., 2009), bonobos (Large and Gray, 2015), and sea-lions (Cook et al., 2013). Although some animals can exhibit rhythmic capabilities, some remarkably well like Ronan the sea-lion (Rouse et al., 2016), they are in some ways limited compared to humans (Patel and Iversen, 2014; Merker et al., 2015). Even though there are animals that can entrain to auditory rhythms, only humans

appear to be naturally inclined to do so (Wilson and Cook, 2016). Finally, there is some evidence that non-human primates are able to synchronize their movements to predictable visual stimuli (Takeya et al., 2017), yet there has been much less research on visual SMS compared to auditory SMS in non-humans.

## GENERAL SYNTHESIS AND FUTURE DIRECTIONS

In looking at how the brain processes timing information, it is clear that many context sensitive mechanisms interact and coordinate to provide optimal timing output. Much of this interaction appears to happen within the motor system and likely involves the subcortical systems to coordinate the various mechanisms. Current research suggests that oscillations play a key role coordinating the interactions among various timing circuits. However, it is not clear if the various timing systems compute measures of time in the same way. When considering that auditory and visual systems take in very different kinds of information and use it in different ways, i.e., auditory has a stronger temporal precision, and visual has a strong spatial bias, it seems likely that the timing mechanisms themselves may greatly differ.

Consider the difference between extracting timing information between a moving visual rhythm and an auditory rhythm. Moving visual stimuli contain more information than auditory stimuli, such that while entraining to auditory stimuli, prediction of the onset of the next event involves encoding the interval between two events and utilizing that information to predict the onset of the next event. With a moving visual rhythmic stimulus, that interval information is present, but so is information on position/velocity/acceleration. This means predictions of the onset of the next event can be made as part of a continuous process. The fact that even with this information, visual SMS is at best equal to auditory SMS except at fast speeds, begs the question as to why visual SMS is less capable. One possible explanation for this is that the visual system has to encode much more information, and further, encoding that information into a form that is usable by the motor network may require extra processing. This may explain the timing activity found within the visual cortex during visual SMS. Even when there is a simple flashing metronome, there is a measure of timing activity originating from the visual cortex. Considering the reduced temporal ability with visual flashing metronomes, it suggests there may be a translation issue in harnessing a system not optimized to temporal processing the way the auditory system has been, resulting in a weaker connection to the motor timing network.

Different timing systems likely employ varying mechanisms and computational principles that are appropriate to the time scale, cellular properties, and general needs of the system. Existing computational models that capture a range of these phenomena across levels include: pacemaker accumulator models, multiple oscillator models, memory trace models,

random process models, ramping activity models, delay line models, and state space trajectory-based models (Addyman et al., 2016; Hass and Durstewitz, 2016). Such models help illustrate the variety of ways to process timing information within a neural network. Evidence also suggests that cells with specific timing mechanisms exist in the basal ganglia and cerebellum (Lusk et al., 2016), yet other areas with multiple functional properties also process timing, such as in the prefrontal cortex (Hyman et al., 2012) and hippocampus (MacDonald et al., 2011). The areas that have multiple functions, as in the hippocampus and prefrontal cortex, will then likely have different computational approach than more specialized timing structures.

Given that there are multiple ways to process timing, and that many forms of cognition require some form of temporal processing, it would be surprising to find that timing mechanisms are not ubiquitous in the brain. This raises an important question. If many different timing mechanisms are available for a given task, and only one output (through action), how do neural systems arrive at the best timing information to use? A strong candidate explanation for this would implicate a mechanism that helps integration through an optimal Bayesian process (Hass and Durstewitz, 2016). Evidence from multimodal sensory integration suggests that when timing information is presented from multiple modalities, the modalities are combined and weighted based on reliability in Bayesian optimal solution (Ernst and Banks, 2002). Since most timing related activity requires motor output, we would expect that the source of timing to be utilized would be determined before, or as that timing information becomes available to the motor system. This seems to make the case that the striatal cells operating as a comparator may be the seat of the Bayesian process to determine the optimal timing source for motor timing.

Since there is some disparity in the amount of work on auditory and visual SMS error correction, there is a need to further study the error correction capabilities within visual SMS. It is currently unknown if visual error correction can be as fast as auditory error correction when dealing modality appropriate stimuli, such as a moving visual sequence or bouncing ball. Another major area of needed work is in understanding the mechanisms by which the Bayesian optimal timing source is chosen in cases where multiple sources are available. If timing mechanisms are as ubiquitous in the brain as evidence suggests, then there may be a variety of ways these mechanisms interface with the motor timing system to produce a single output. Further imaging and computational work is required to understanding this mechanism.

## AUTHOR CONTRIBUTIONS

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

## FUNDING

This work was partially supported by a grant from the National Science Foundation BCS-1460633.



## REFERENCES

- Addyman, C., French, R. M., and Thomas, E. (2016). Computational models of interval timing. *Curr. Opin. Behav. Sci.* 8, 140–146. doi: 10.1016/j.cobeha.2016.01.004
- Araneda, R., Renier, L., Ebner-Karestinos, D., Dricot, L., and De Volder, A. G. (2017). Hearing, feeling or seeing a beat recruits a supramodal network in the auditory dorsal stream. *Eur. J. Neurosci.* 45, 1439–1450. doi: 10.1111/ejn.13349
- Arias, P., and Cudeiro, J. (2008). Effects of rhythmic sensory stimulation (auditory, visual) on gait in Parkinson's disease patients. *Exp. Brain Res.* 186, 589–601. doi: 10.1007/s00221-007-1263-y
- Bijsterbosch, J. D., Lee, K. H., Hunter, M. D., Tsoi, D. T., Lankappa, S., Wilkinson, I. D., et al. (2011). The role of the cerebellum in sub- and supraliminal error correction during sensorimotor synchronization: evidence from fMRI and TMS. *J. Cogn. Neurosci.* 23, 1100–1112. doi: 10.1162/jocn.2010.21506
- Brown, S., and Jordania, J. (2013). Universals in the world's musics. *Psychol. Music* 41, 229–248. doi: 10.1177/0305735611425896
- Chen, Y., Repp, B. H., and Patel, A. D. (2002). Spectral decomposition of variability in synchronization and continuation tapping: comparisons between auditory and visual pacing and feedback conditions. *Hum. Mov. Sci.* 21, 515–532. doi: 10.1016/S0167-9457(02)00138-0
- Comstock, D., and Balasubramaniam, R. (2017a). "Differences in neural correlates of error correction in auditory and visual sensorimotor synchronization," in *Poster presented at the 24th Annual Cognitive Neuroscience Society Meeting* (San Francisco, CA).
- Comstock, D., and Balasubramaniam, R. (2017b). "Beta-band response synchronizes and predicts rhythmic flashing visual stimuli," in *Poster presented at the 47th Annual Society for Neuroscience Meeting* (Washington, DC).
- Comstock, D. C., and Balasubramaniam, R. (2018). Neural responses to perturbations in visual and auditory metronomes during sensorimotor synchronization. *Neuropsychologia* 117, 55–66. doi: 10.1016/j.neuropsychologia.2018.05.013
- Cook, P., Rouse, A., Wilson, M., and Reichmuth, C. (2013). A California sea lion (*Zalophus californianus*) can keep the beat: motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *J. Comp. Psychol.* 127:412. doi: 10.1037/a0032345
- 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., Vidal, F., and Burle, B. (2016). When to act, or not to act: that's the SMA's question. *Curr. Opin. Behav. Sci.* 8, 14–21. doi: 10.1016/j.cobeha.2016.01.003
- Davie, C. A. (2008). A review of Parkinson's disease. *Br. Med. Bull.* 86, 109–127. doi: 10.1093/bmb/ldn013
- Doumas, M., Praamstra, P., and Wing, A. M. (2005). Low frequency rTMS effects on sensorimotor synchronization. *Exp. Brain Res.* 167, 238–245. doi: 10.1007/s00221-005-0029-7
- Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429. doi: 10.1038/415429a
- Fitch, W. (2016). Dance, music, meter and groove: a forgotten partnership. *Front. Hum. Neurosci.* 10:64. doi: 10.3389/fnhum.2016.00064
- 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
- Gan, L., Huang, Y., Zhou, L., Qian, C., and Wu, X. (2015). Synchronization to a bouncing ball with a realistic motion trajectory. *Sci. Rep.* 5:11974. doi: 10.1038/srep11974
- Grahn, J. A., Henry, M. J., and McAuley, J. D. (2011). FMRI investigation of cross-modal interactions in beat perception: audition primes vision, but not vice versa. *Neuroimage* 54, 1231–1243. doi: 10.1016/j.neuroimage.2010.09.033
- Grahn, J. A., and Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *J. Neurosci.* 29, 7540–7548. doi: 10.1523/JNEUROSCI.2018-08.2009
- 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
- Gupta, D. S., and Chen, L. (2016). Brain oscillations in perception, timing and action. *Curr. Opin. Behav. Sci.* 8, 161–166. doi: 10.1016/j.cobeha.2016.02.021
- Hagen, E. H., and Bryant, G. A. (2003). Music and dance as a coalition signaling system. *Hum. Nat.* 14, 21–51. doi: 10.1007/s12110-003-1015-z
- Hass, J., and Dürstewitz, D. (2016). Time at the center, or time at the side? Assessing current models of time perception. *Curr. Opin. Behav. Sci.* 8, 238–244. doi: 10.1016/j.cobeha.2016.02.030
- Hove, M. J., Balasubramaniam, R., and Keller, P. E. (2014). The time course of phase correction: a kinematic investigation of motor adjustment to timing perturbations during sensorimotor synchronization. *J. Exp. Psychol. Hum. Percept. Perform.* 40:2243. doi: 10.1037/a0037826
- Hove, M. J., Fairhurst, M. T., Kotz, S. A., and Keller, P. E. (2013b). Synchronizing with auditory and visual rhythms: an fMRI assessment of modality differences and modality appropriateness. *Neuroimage* 67, 313–321. doi: 10.1016/j.neuroimage.2012.11.032
- Hove, M. J., Iversen, J. R., Zhang, A., and Repp, B. H. (2013a). Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome. *Psychol. Res.* 77, 388–398. doi: 10.1007/s00426-012-0441-0
- Hove, M. J., and Keller, P. E. (2010). Spatiotemporal relations and movement trajectories in visuomotor synchronization. *Music Percept. Interdisc. J.* 28, 15–26. doi: 10.1525/mp.2010.28.1.15
- Hove, M. J., Spivey, M. J., and Krumhansl, C. L. (2010). Compatibility of motion facilitates visuomotor synchronization. *J. Exper. Psychol. Hum. Percept. Perform.* 36:1525. doi: 10.1037/a0019059
- Hyman, J. M., Ma, L., Balaguer-Ballester, E., Dürstewitz, D., and Seamans, J. K. (2012). Contextual encoding by ensembles of medial prefrontal cortex neurons. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5086–5091. doi: 10.1073/pnas.1114415109
- Iversen, J. R. (2016). "21 In the beginning was the beat: evolutionary origins of musical rhythm in humans," in *The Cambridge Companion to Percussion*, eds R. Hartenberger (Cambridge, MA: Cambridge University Press), 281–295.
- Iversen, J. R., and Balasubramaniam, R. (2016). Synchronization and temporal processing. *Curr. Opin. Behav. Sci.* 8, 175–180. doi: 10.1016/j.cobeha.2016.02.027
- Iversen, J. R., Patel, A. D., Nicodemus, B., and Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition* 134, 232–244. doi: 10.1016/j.cognition.2014.10.018
- 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–doi: 10.1111/j.1749-6632.2009.04579.x
- Janata, P., Tomic, S. T., and Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *J. Exp. Psychol. Gen.* 141:54. doi: 10.1037/a0024208
- Jäncke, L., Loose, R., Lutz, K., Specht, K., and Shah, N. J. (2000). Cortical activations during paced finger-tapping applying visual and auditory pacing stimuli. *Cogn. Brain Res.* 10, 51–66. doi: 10.1016/S0926-6410(00)00022-7
- Jang, J., Jones, M., Milne, E., Wilson, D., and Lee, K. H. (2016). Contingent negative variation (CNV) associated with sensorimotor timing error correction. *Neuroimage* 127, 58–66. doi: 10.1016/j.neuroimage.2015.11.071
- Jantzen, K. J., Ratcliff, B. R., and Jantzen, M. G. (2018). Cortical Networks for correcting errors in sensorimotor synchronization depend on the direction of asynchrony. *J. Mot. Behav.* 50, 235–248. doi: 10.1080/00222895.2017.1327414
- Jantzen, K. J., Steinberg, F. L., and Kelso, J. A. (2005). Functional MRI reveals the existence of modality and coordination-dependent timing networks. *Neuroimage* 25, 1031–1042. doi: 10.1016/j.neuroimage.2004.12.029
- Kotz, S. A., Brown, R. M., and Schwartze, M. (2016). Cortico-striatal circuits and the timing of action and perception. *Curr. Opin. Behav. Sci.* 8, 42–45. doi: 10.1016/j.cobeha.2016.01.010
- Laland, K., Wilkins, C., and Clayton, N. (2016). The evolution of dance. *Curr. Biol.* 26, R5–R9. doi: 10.1016/j.cub.2015.11.031
- Large, E. W., and Gray, P. M. (2015). Spontaneous tempo and rhythmic entrainment in a bonobo (*Pan paniscus*). *J. Comp. Psychol.* 129:317. doi: 10.1037/com0000011



- Large, E. W., and Snyder, J. S. (2009). Pulse and meter as neural resonance. *Ann. N.Y. Acad. Sci.* 1169, 46–57. doi: 10.1111/j.1749-6632.2009.04550.x
- Lorås, H., Sigmundsson, H., Talcott, J. B., Öhberg, F., and Stensdotter, A. K. (2012). Timing continuous or discontinuous movements across effectors specified by different pacing modalities and intervals. *Exp. Brain Res.* 220, 335–347. doi: 10.1007/s00221-012-3142-4
- Lusk, N. A., Petter, E. A., MacDonald, C. J., and Meck, W. H. (2016). Cerebellar, hippocampal, and striatal time cells. *Curr. Opin. Behav. Sci.* 8, 186–192. doi: 10.1016/j.cobeha.2016.02.020
- MacDonald, C. J., Lepage, K. Q., Eden, U. T., and Eichenbaum, H. (2011). Hippocampal “time cells” bridge the gap in memory for discontinuous events. *Neuron* 71, 737–749. doi: 10.1016/j.neuron.2011.07.012
- 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
- Merchant, H., and Yarrow, K. (2016). How the motor system both encodes and influences our sense of time. *Curr. Opin. Behav. Sci.* 8, 22–27. doi: 10.1016/j.cobeha.2016.01.006
- Merker, B., Morley, I., and Zuidema, W. (2015). Five fundamental constraints on theories of the origins of music. *Phil. Trans. R. Soc. B*, 370:20140095. doi: 10.1098/rstb.2014.0095
- Morein-Zamir, S., Soto-Faraco, S., and Kingstone, A. (2003). Auditory capture of vision: examining temporal ventriloquism. *Cogn. Brain Res.* 17, 154–163. doi: 10.1016/S0926-6410(03)00089-2
- Nozaradan, S., Peretz, I., Missal, M., and Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *J. Neurosci.* 31, 10234–10240. doi: 10.1523/JNEUROSCI.0411-11.2011
- Okawa, H., Suefusa, K., and Tanaka, T. (2017). Neural Entrainment to Auditory Imagery of Rhythms. *Front. Hum. Neurosci.* 11:493. doi: 10.3389/fnhum.2017.00493
- Parise, C., and Spence, C. (2008). Synesthetic congruency modulates the temporal ventriloquism effect. *Neurosci. Lett.* 442, 257–261. doi: 10.1016/j.neulet.2008.07.010
- Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Percept. Interdisc. J.* 24, 99–104. doi: 10.1525/mp.2006.24.1.99
- Patel, A. D., and Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Front. Sys. Neurosci.* 8:57. doi: 10.3389/fnsys.2014.00057
- Patel, A. D., Iversen, J. R., Bregman, M. R., and Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Curr. Biol.* 19, 827–830. doi: 10.1016/j.cub.2009.03.038
- Phillips-Silver, J., and Trainor, L. J. (2005). Feeling the beat: movement influences infant rhythm perception. *Science* 308, 1430–1430. doi: 10.1126/science.1110922
- Phillips-Silver, J., and Trainor, L. J. (2007). Hearing what the body feels: auditory encoding of rhythmic movement. *Cognition* 105, 533–546. doi: 10.1016/j.cognition.2006.11.006
- Phillips-Silver, J., and Trainor, L. J. (2008). Vestibular influence on auditory metrical interpretation. *Brain Cogn.* 67, 94–102. doi: 10.1016/j.bandc.2007.11.007
- Praamstra, P., Turgeon, M., Hesse, C. W., Wing, A. M., and Perryer, L. (2003). Neurophysiological correlates of error correction in sensorimotor-synchronization. *Neuroimage* 20, 1283–1297. doi: 10.1016/S1053-8119(03)00351-3
- Provasi, J., Anderson, D. I., and Barbu-Roth, M. (2014). Rhythm perception, production, and synchronization during the perinatal period. *Front. Psychol.* 5:1048. doi: 10.3389/fpsyg.2014.01048
- Rauschecker, J. P. (2017). Where, When, and How: are they all Sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex* 98, 262–268. doi: 10.1016/j.cortex.2017.10.020
- Ravignani, A., and Cook, P. F. (2016). The evolutionary biology of dance without frills. *Curr. Biol.* 26, R878–R879. doi: 10.1016/j.cub.2016.07.076
- Repp, B. H. (2000). Compensation for subliminal timing perturbations in perceptual-motor synchronization. *Psychol. Res.* 63, 106–128. doi: 10.1007/PL00008170
- Repp, B. H. (2005). Sensorimotor synchronization: a review of the tapping literature. *Psychon. Bull. Rev.* 12, 969–992. doi: 10.3758/BF03206433
- Repp, B. H. (2001a). Phase correction, phase resetting, and phase shifts after subliminal timing perturbations in sensorimotor synchronization. *J. Exp. Psychol. Hum. Percept. Perform.* 27:600. doi: 10.1037/0096-1523.27.3.600
- Repp, B. H. (2001b). Processes underlying adaptation to tempo changes in sensorimotor synchronization. *Hum. Mov. Sci.* 20, 277–312. doi: 10.1016/S0167-9457(01)00049-5
- Repp, B. H. (2003). Rate limits in sensorimotor synchronization with auditory and visual sequences: the synchronization threshold and the benefits and costs of interval subdivision. *J. Motor Behav.* 35, 355–370. doi: 10.1080/00222890309603156
- Repp, B. H., and Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: effects of intention, attention, and awareness. *Q. J. Exp. Psychol. A*, 57, 499–521. doi: 10.1080/02724980343000369
- Repp, B. H., and Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *J. Exp. Psychol. Hum. Percept. Perform.* 28, 1085–1099.
- Repp, B. H., and Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychol. Res.* 68, 252–270. doi: 10.1007/s00426-003-0143-8
- 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
- Rochester, L., Hetherington, V., Jones, D., Nieuwboer, A., Willems, A. M., Kwakkel, G., et al. (2005). The effect of external rhythmic cues (auditory and visual) on walking during a functional task in homes of people with Parkinson's disease. *Arch. Phys. Med. Rehabil.* 86, 999–1006. doi: 10.1016/j.apmr.2004.10.040
- Ross, J. M., Iversen, J. R., and Balasubramaniam, R. (2016a). Motor simulation theories of musical beat perception. *Neurocase* 22, 558–565. doi: 10.1080/13554794.2016.1242756
- Ross, J. M., Iversen, J. R., and Balasubramaniam, R. (2018). The role of posterior parietal cortex in beat-based timing perception: a continuous theta burst stimulation study. *J. Cogn. Neurosci.* 30, 634–643. doi: 10.1162/jocn\_a\_01237
- Ross, J. M., Warlaumont, A. S., Abney, D. H., Rigoli, L. M., and Balasubramaniam, R. (2016b). Influence of musical groove on postural sway. *J. Exp. Psychol. Hum. Percept. Perform.* 42:308. doi: 10.1037/xhp0000198
- Rouse, A. A., Cook, P. F., Large, E. W., and Reichmuth, C. (2016). Beat keeping in a sea lion as coupled oscillation: implications for comparative understanding of human rhythm. *Front. Neurosci.* 10:257. doi: 10.3389/fnins.2016.00257
- Sarnthein, J., Petsche, H., Rappelsberger, P., Shaw, G. L., and von Stein, A. (1998). Synchronization between prefrontal and posterior association cortex during human working memory. *Proc. Natl. Acad. Sci. U.S.A.* 95, 7092–7096. doi: 10.1073/pnas.95.12.7092
- Schwartz, M., Rothermich, K., and Kotz, S. A. (2012). Functional dissociation of pre-SMA and SMA-proper in temporal processing. *Neuroimage* 60, 290–298. doi: 10.1016/j.neuroimage.2011.11.089
- Shuler, M. G. (2016). Timing in the visual cortex and its investigation. *Curr. Opin. Behav. Sci.* 8, 73–77. doi: 10.1016/j.cobeha.2016.02.006
- Silva, S., and Castro, S. L. (2016). Moving stimuli facilitate synchronization but not temporal perception. *Front. Psychol.* 7:1798. doi: 10.3389/fpsyg.2016.01798
- Takeya, R., Kameda, M., Patel, A. D., and Tanaka, M. (2017). Predictive and tempo-flexible synchronization to a visual metronome in monkeys. *Sci. Reports* 7:6127. doi: 10.1038/s41598-017-06417-3
- Teki, S., Grube, M., Kumar, S., and Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *J. Neurosci.* 31, 3805–3812. doi: 10.1523/JNEUROSCI.5561-10.2011
- Todd, N. P., and Lee, C. S. (2015). The sensory-motor theory of rhythm and beat induction 20 years on: a new synthesis and future perspectives. *Front. Hum. Neurosci.* 9:444. doi: 10.3389/fnhum.2015.00444
- Trainor, L. J., Gao, X., Lei, J. J., Lehtovaara, K., and Harris, L. R. (2009). The primal role of the vestibular system in determining musical rhythm. *Cortex* 45, 35–43. doi: 10.1016/j.cortex.2007.10.014
- von Stein, A., and Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *Int. J. Psychophysiol.* 38, 301–313. doi: 10.1016/S0167-8760(00)00172-0

- Vroomen, J., Bertelson, P., and de Gelder, B. (2001). Directing spatial attention towards the illusory location of a ventriloquized sound. *Acta Psychol.* 108, 21–33. doi: 10.1016/S0001-6918(00)00068-8
- 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
- Wiener, M., Turkeltaub, P., and Coslett, H. B. (2010). The image of time: a voxel-wise meta-analysis. *Neuroimage* 49, 1728–1740. doi: 10.1016/j.neuroimage.2009.09.064
- Williams, P. E., Mechler, F., Gordon, J., Shapley, R., and Hawken, M. J. (2004). Entrainment to video displays in primary visual cortex of macaque and humans. *J. Neurosci.* 24, 8278–8288. doi: 10.1523/JNEUROSCI.2716-04.2004
- Wilson, M., and Cook, P. F. (2016). Rhythmic entrainment: why humans want to, fireflies can't help it, pet birds try, and sea lions have to be bribed. *Psychon. Bull. Rev.* 23, 1647–1659. doi: 10.3758/s13423-016-1013-x
- Wing, A. M., and Kristofferson, A. (1973). The timing of interresponse intervals. *Percept. Psychophys.* 13, 455–460. doi: 10.3758/bf03205802
- Zatorre, R. J., Chen, J. L., and Penhune, V. B. (2007). When the brain plays music: auditory–motor interactions in music perception and production. *Nat. Rev. Neurosci.* 8:547. doi: 10.1038/nrn2152
- Zhou, B., Yang, S., Mao, L., and Han, S. (2014). Visual feature processing in the early visual cortex affects duration perception. *J. Exp. Psychol. Gen.* 143:1893. doi: 10.1037/a0037294

**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 © 2018 Comstock, Hove and Balasubramaniam. 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) and the copyright owner(s) 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.



# Effect of Stimulus Contrast and Visual Attention on Spike-Gamma Phase Relationship in Macaque Primary Visual Cortex

Aritra Das and Supratim Ray\*

Centre for Neuroscience, Indian Institute of Science, Bangalore, India

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

Pedro E. Maldonado,  
Universidad de Chile, Chile  
Ko Sakai,  
University of Tsukuba, Japan

### \*Correspondence:

Supratim Ray  
sray@iisc.ac.in

**Received:** 02 April 2018

**Accepted:** 20 July 2018

**Published:** 14 August 2018

### Citation:

Das A and Ray S (2018) Effect of Stimulus Contrast and Visual Attention on Spike-Gamma Phase Relationship in Macaque Primary Visual Cortex. *Front. Comput. Neurosci.* 12:66. doi: 10.3389/fncom.2018.00066

Brain signals often show rhythmic activity in the so-called gamma range (30–80 Hz), whose magnitude and center frequency are modulated by properties of the visual stimulus such as size and contrast, as well as by cognitive processes such as attention. How gamma rhythm can potentially influence cortical processing remains unclear; previous studies have proposed a scheme called phase coding, in which the intensity of the incoming stimulus is coded in the position of the spike relative to the rhythm. Using chronically implanted microelectrode arrays in the primary visual cortex (area V1) of macaques engaged in an attention task while presenting stimuli of varying contrasts, we tested whether the phase of the gamma rhythm relative to spikes varied as a function of stimulus contrast and attentional state. A previous study had found no evidence of gamma phase coding for either contrast or attention in V1, but in that study spikes and local field potential (LFP) were recorded from the same electrode, due to which spike-gamma phase estimation could have been biased. Further, the filtering operation to obtain LFP could also have biased the gamma phase. By analyzing spikes and LFP from different electrodes, we found a weak but significant effect of attention, but not stimulus contrast, on gamma phase relative to spikes. The results remained consistent even after correcting the filter induced lags, although the absolute magnitude of gamma phase shifted by up to  $\sim 15^\circ$ . Although we found a significant effect of attention, we argue that a small magnitude of phase shift as well as the dependence of phase angles on gamma power and center frequency limits a potential role of gamma in phase coding in V1.

**Keywords:** attention, spike-field coherence, spike-gamma phase, contrast, area V1, stLFP

## INTRODUCTION

Gamma oscillations are rhythmic fluctuations in a frequency range between 30 and 80 Hz in brain signals (Buzsáki, 2006; Buzsáki et al., 2013), which have been consistently linked with high-level cognitive processes such as attention (Fries et al., 2001; Gregoriou et al., 2009), perception (Rodríguez et al., 1999) and feature binding (Singer, 1999). In recordings from the primary visual cortex (area V1), gamma is also known to be highly dependent on the properties of visual stimulus, such as size (Gieselmann and Thiele, 2008; Ray and Maunsell, 2011a; Jia et al., 2013),

orientation (Berens et al., 2008; Jia et al., 2011), and contrast (Ray and Maunsell, 2010; Jia et al., 2013). Although several hypotheses about how gamma rhythm could influence neural processing have been proposed, such as binding by synchrony (Singer, 1999) and communication-through-coherence (Fries, 2015), whether gamma plays a functional role remains unclear (Ray and Maunsell, 2015).

Here we test a specific hypothesis called phase coding (PC), originally proposed in the context of theta rhythm in the hippocampus (O'Keefe and Recce, 1993; Buzsáki and Chrobak, 1995), in which information is coded in the position of the spike relative to the rhythm. In the context of gamma rhythm (Fries et al., 2007), which is thought to be associated with an inhibitory network of interneurons (Bartos et al., 2007; Cardin et al., 2009; Sohal et al., 2009), this hypothesis posits that the rhythmic network inhibition interacts with excitatory input to pyramidal cells such that the more excited cells (which can overcome the inhibition earlier) fire earlier in the gamma cycle. Thus, stimulus intensity can be coded in the gamma phase relative to the spike. However, whether gamma PC occurs is controversial, with evidence both in favor and against the hypothesis. We have earlier shown that in macaque secondary somatosensory cortex, the phase of gamma rhythm does not vary with stimulus intensity (Ray et al., 2008). In V1, one study showed some evidence of PC with different orientations (which they took as a proxy for stimulus intensity), at least for sites that had weak gamma power and weak gamma-spike phase locking (Vinck et al., 2010). Other studies in V1 showed no evidence of PC when the stimulus contrast (a more direct index of stimulus intensity as compared to orientation) was varied (Chalk et al., 2010; Ray and Maunsell, 2010). Importantly, Chalk and colleagues further showed that even attention, which increases the effective contrast of the stimulus (Carrasco et al., 2004), does not cause a shift in spike-gamma phase. They also showed that in V1, attention causes a reduction in gamma power and spike-gamma coupling (Chalk et al., 2010), exactly opposite of what has been shown in higher cortical areas such as V4 (Fries et al., 2001).

While Chalk et al. (2010) failed to provide evidence in favor of PC, one limitation of their study was that spikes and local field potential (LFP) were collected from the same electrode, which can potentially bias the spike-gamma phase relationship because of the presence of spike-related transients (Ray, 2015). Specifically, spikes are associated with a "transient" in the LFP recorded from the same electrode, which could be due to synaptic activity that leads to the spike as well as low-frequency component of the action potential (spike "bleed-through"; see Ray, 2015, for details). The remaining studies either removed this transient using signal processing techniques such as Matching Pursuit (Ray et al., 2008), or used spikes and LFP from different electrodes (Ray and Maunsell, 2010; Vinck et al., 2010) that reduces the bias (see section Discussion for more details on this), but none of these reports studied the effect of attention on spike-gamma phase. To test whether stimulus contrast can be coded in the phase of the gamma rhythm, we here trained monkeys to do a demanding attention task while presenting stimuli that varied in contrast to study the effect of both contrast and attention on spike-gamma phase, while

recording from chronically implanted microelectrode arrays such that spike-gamma phase could be estimated using spikes and LFPs recorded from different electrodes. Note that since attention is thought to increase the effective stimulus contrast (Carrasco et al., 2004), testing whether gamma phase varies with attentional state is also a test for gamma PC for contrast. We further studied the effect of the online causal filter used to obtain the LFP, which introduces a delay in the LFP and has been shown to influence spike-LFP relationships (Okun, 2017), but has not been accounted for in previous studies.

## MATERIALS AND METHODS

Experimental procedures have been described in detail in earlier studies (Ray and Maunsell, 2010, 2011b; Shirhatti et al., 2016); we provide a brief description here.

### Ethics Statement

The animal protocols reported in this study were approved by the Institutional Animal Care and Use Committee of Harvard Medical School.

### Electrophysiological Recordings

Two male rhesus monkeys (*Macaca mulatta*) were implanted with a scleral search coil and a head post and were subsequently trained to perform an attentionally demanding task. Once they learned the behavioral task, a microelectrode array (Blackrock Microsystems, 96 active electrodes) was implanted in the right V1 cortex (about 15 mm anterior from the occipital ridge and 15 mm lateral from the midline). The microelectrodes were 1 mm long and 400  $\mu\text{m}$  apart from each other, with impedance between 0.3 and 1  $\text{M}\Omega$  at 1 kHz. Although histological analysis had not been performed to identify the exact location of the microelectrode tips, they are expected to be in cortical layer 2/3 or 4 based on the approximate thickness of V1 (2 mm; Hubel and Wiesel, 1977). Electrical signals were recorded using commercial hardware and software (Blackrock Microsystems), referenced to a wire placed on the dura near the microelectrode grid. Raw electrical signals were filtered between 0.3 Hz (Butterworth filter, 1st order, analog) and 500 Hz (Butterworth, 4th order, digital) and digitized at 2 kHz (16-bit resolution) to get the LFP. Multi-units were extracted by filtering the raw signal between 250 Hz (Butterworth filter, fourth order, digital) and 7,500 Hz (Butterworth filter, third order, analog) followed by an amplitude threshold (set at  $\sim 6.25$  and  $\sim 4.25$  of the signal SD for the two monkeys). To improve the quality of unit isolation, multi-units were subsequently sorted offline (Offline Sorter, Plexon Inc.). The receptive fields, obtained by flashing small Gabor stimuli on a rectangular grid that encompassed the receptive fields of all the electrodes in the array, were located in the lower left quadrant of the visual space at an eccentricity of about  $3\text{--}5^\circ$ . As in previous studies, only electrodes for which stable estimates of the receptive fields could be obtained (27 and 66 electrodes for the two monkeys), were used for subsequent analysis.



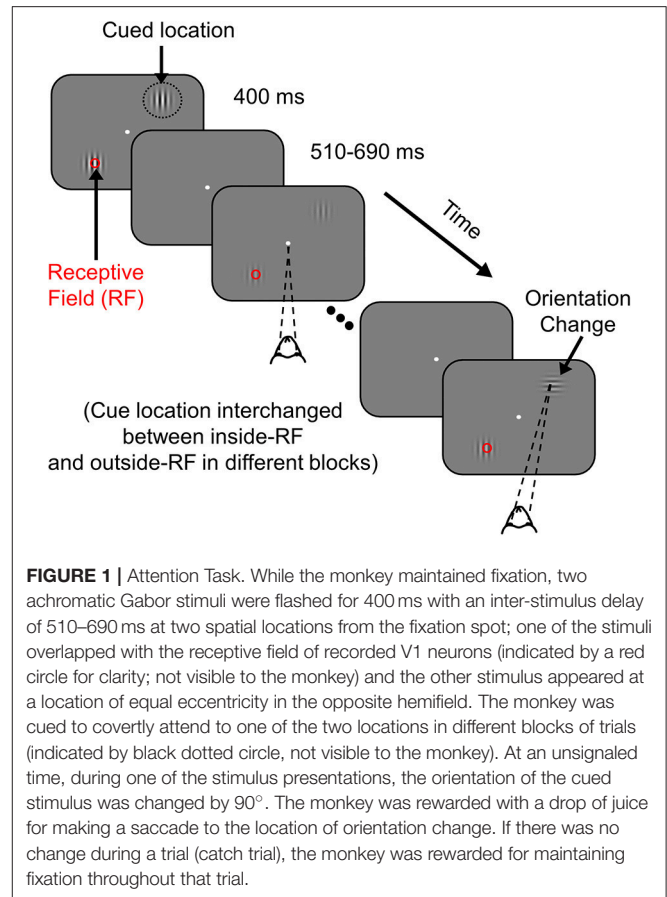
## Behavioral Task Paradigm

The monkeys were required to maintain their gaze within  $1^\circ$  of a small central dot ( $0.05^\circ$ – $0.10^\circ$  diameter) during the task while two achromatic odd-symmetric static Gabor stimuli were synchronously flashed for 400 ms with a mean inter-stimulus period of 600 ms. One of the two Gabor stimuli was centered on the receptive field of one of the recorded sites (new location for every session) while the second stimulus was located at an equal eccentricity on the opposite side of the central fixation point. The monkeys were cued to pay attention to one of the two stimulus locations in different blocks of trials by presenting two instruction trials (not included in the analysis) at the start of the block, in which there was only a single stimulus. The contrasts of the attended and unattended Gabor stimuli were equal on each presentation and could take any of the eight possible values: 0, 1.6, 3.1, 6.2, 12.5, 25, 50, and 100%, chosen pseudo-randomly. At an unstipulated time drawn from an exponential distribution (mean 2,000 ms, range 1,000–7,000 ms for Monkey 1; mean 3,000 ms, range 1,000–7,000 ms for Monkey 2), the orientation of the stimulus at the cued location changed by  $90^\circ$ . An exponential distribution was used to minimize expectation of target appearance and to keep the attentional state uniform during a trial since the hazard function is flat for an exponentially distributed target onset time. The monkeys were rewarded with a drop of juice for making a saccade to the location of the altered stimulus within 500 ms of orientation change. To account for saccade latency and to minimize guessing, monkeys were rewarded only for saccades beginning at least 100 ms after the orientation change. Trials were terminated at 7,000 ms if the target had not appeared, in which case the monkeys were rewarded for maintaining fixation throughout that trial. These catch trials were excluded from analysis (for more details, see Ray and Maunsell, 2010 and **Figure 1**).

The Gabor stimuli used for this task were both static with SD of  $0.5^\circ$ , spatial frequency of 4 cycles per degree, with one of the Gabor stimuli located at the center of the receptive field of one of the recorded sites (new recording site for each session), at its preferred orientation. Data from the two monkeys were collected in 10 and 17 recording sessions, respectively. Only correct trials were used for analysis. For each of the correct trials, only the second stimulus up to the last stimulus before target onset were used for analysis. We only used stimulus contrasts for which salient gamma oscillations were observed (25, 50, and 100% contrasts). For each contrast and attention condition, on average we obtained  $79 \pm 4$  (range 55–101) stimulus repeats for Monkey 1 and  $74 \pm 5$  (range 47–120) for Monkey 2.

## Electrodes and Electrode Pair Selection

Electrodes with receptive field centers within  $0.2^\circ$  of the stimulus center in each of the recording sessions were used for analysis, yielding 63 electrodes (23 unique; many electrodes were selected in several recording sessions) for Monkey 1 and 89 electrodes (53 unique) for Monkey 2. These are referred to as “LFP” electrodes. For spike-field coherence (SFC), spike-triggered LFP (stLFP) and spike-gamma phase histograms, we selected a subset of the LFP electrodes from which at least 20 spikes could be recorded in the analysis interval (150–400 ms after stimulus onset) and the



**FIGURE 1 |** Attention Task. While the monkey maintained fixation, two achromatic Gabor stimuli were flashed for 400 ms with an inter-stimulus delay of 510–690 ms at two spatial locations from the fixation spot; one of the stimuli overlapped with the receptive field of recorded V1 neurons (indicated by a red circle for clarity; not visible to the monkey) and the other stimulus appeared at a location of equal eccentricity in the opposite hemifield. The monkey was cued to covertly attend to one of the two locations in different blocks of trials (indicated by black dotted circle, not visible to the monkey). At an unsignaled time, during one of the stimulus presentations, the orientation of the cued stimulus was changed by  $90^\circ$ . The monkey was rewarded with a drop of juice for making a saccade to the location of orientation change. If there was no change during a trial (catch trial), the monkey was rewarded for maintaining fixation throughout that trial.

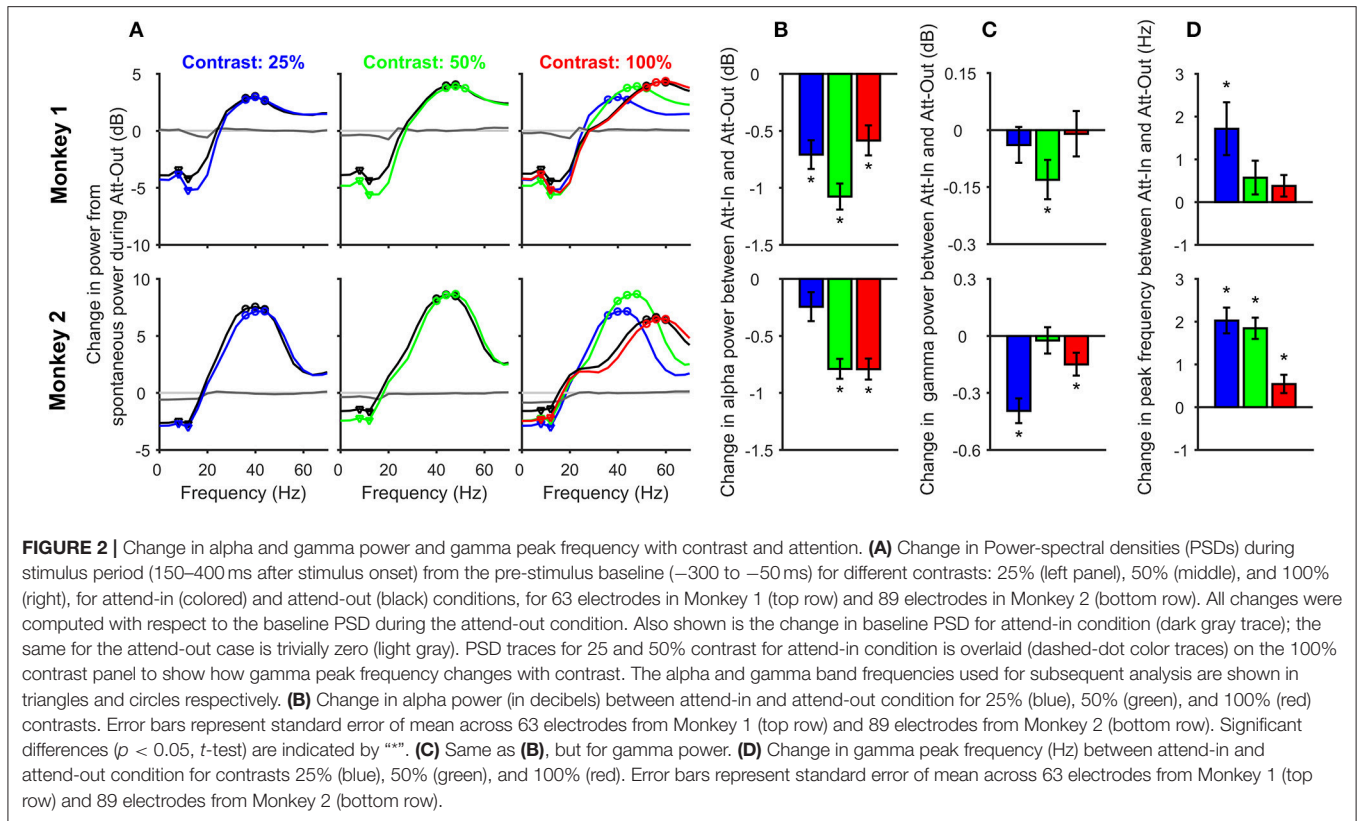
signal to noise ratio of the isolation (Kelly et al., 2007) was greater than 2. This generated 23 (12 unique) and 39 (27 unique) “spike” electrodes for Monkeys 1 and 2, respectively. For each session, we took all combinations of spike and LFP electrodes with receptive fields within  $0.2^\circ$  of the stimulus center, yielding 23 (12 unique) and 39 (27 unique) “same” spike-LFP pairs (**Figure 3**), and 163 (120 unique) and 170 (147 unique) pairs of “different” spike-LFP electrodes for Monkeys 1 and 2, respectively (**Figures 4, 5**).

## DATA ANALYSIS

All data were analyzed using custom codes written in MATLAB (The MathWorks, RRID:SCR\_001622). Individual data analysis methods are briefly summarized below.

### Change in Power Spectral Density (PSD) Plots (**Figure 2**)

Stimulus-induced responses were first obtained by subtracting the mean LFP across all stimulus repeats for each condition (i.e., the event-related potential) from individual single trial time series data. Subsequent analyses were performed on these stimulus-induced responses. Power spectral densities (PSDs) for different stimulus and attention conditions were computed using the multi-taper method with 5 tapers using the Chronux toolbox (Bokil et al., 2010); <http://chronux.org/>, (RRID:SCR\_005547).



The analysis period was selected between 150 and 400 ms after stimulus onset to avoid stimulus onset related transients and compared against a “baseline period” between –300 and –50 ms of stimulus onset. To ensure that the change in power from baseline was not affected due to differences in the baseline power for different attention conditions, change in PSDs were plotted with respect to the baseline response of “attend-out” (attention directed outside the receptive field) condition for each stimulus contrast value:

$$\Delta PSD_i = 10 (\log_{10} (ST)_i - \log_{10} (BL_{Att Out})_i)$$

Here  $i$  represents the contrast condition (25, 50, or 100%),  $\Delta PSD_i$  represents the change in PSD in decibels,  $(ST)_i$  denotes the PSD in the stimulus epoch and  $(BL_{Att Out})_i$  denotes the baseline PSD for attend-out condition.

For the change in alpha power shown in **Figure 2B**, we first averaged the power between 8 and 12 Hz (triangles in **Figure 2A**; note that because we used an analysis interval of 250 ms, we had a frequency resolution of 4 Hz) and subsequently took the difference between the “attend-in” (attention directed inside the receptive field) and attend-out power on a log scale:

$$\Delta Power_i = 10 (\log_{10} (ST_{Att In})_i - \log_{10} (ST_{Att Out})_i)$$

Here  $(ST)_i$  denotes the alpha power in the stimulus epoch for the  $i$ th contrast condition. For gamma power, the same procedure was used with three frequency bins centered around

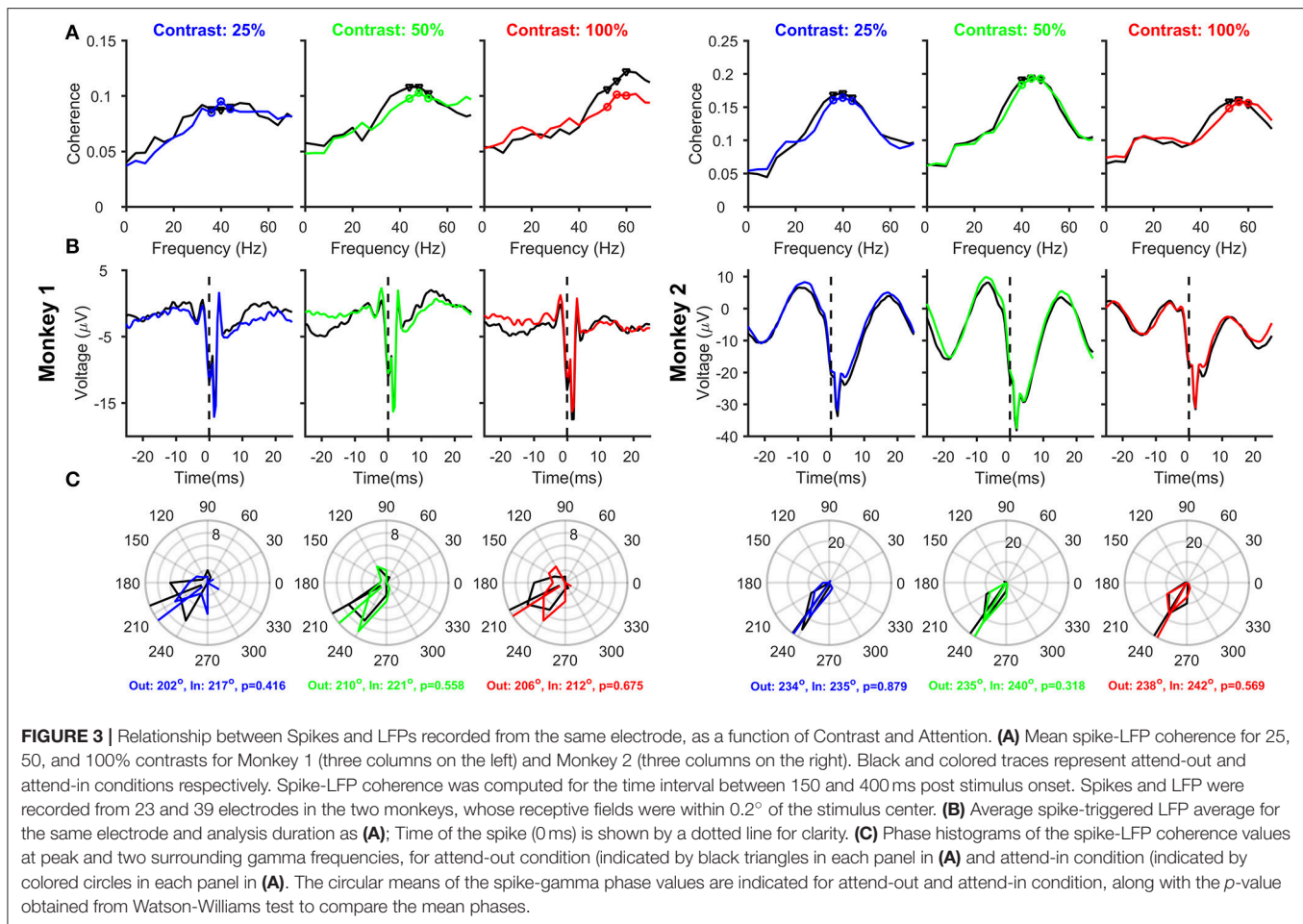
the peak gamma frequency (shown in circles in **Figure 2A**). For computing peak gamma frequency, we choose the frequency bin for which  $\Delta PSD_i$  attained its maximum value between 30 and 60 Hz.

## Coherency Analysis (Figures 3–5)

The coherency between two signals  $x$  and  $y$  is computed using the following equation:

$$Coherency_{xy}(f) = \frac{S_{xy}(f)}{\sqrt{S_{xx}(f)S_{yy}(f)}}$$

Where  $S_{xy}(f)$  denotes the cross-spectrum between the signals  $x$  and  $y$  and  $S_{xx}(f)$  and  $S_{yy}(f)$  denote the auto spectra of each signal. The coherency values were computed using the multi-taper method implemented in Chronux toolbox using five tapers. All the coherence analyses were performed using the sorted multiunit dataset. For spike-field coherence, the spike time series was converted to a binary time series (at 0.5 ms resolution) with a “1” at each time position containing a spike and “0” otherwise (500 data points for the stimulus period). The results were similar for three tapers. All the circular statistical analyses were performed using an open source circular statistics toolbox in MATLAB (CircStat; Berens, 2009). Spike-triggered LFP (stLFP) were computed by taking a  $\pm 25$  ms segment of the LFP around each spike in the stimulus period and subsequently taking the average of those segments.



## Removing Filtering Effect (Figure 5)

Any causal filter necessarily introduces a delay in the signal, which may be dependent on the frequency of the signal. Butterworth filters have a linear relationship between phase delay and frequency, such that the group delay (which roughly translates to how much each frequency component of the signal shifts in time due to the filtering process) is almost constant over a wide frequency range. We removed the low-pass filtering effect by dividing the Fourier Transform of the LFP by the Fourier Transform of the low-pass LFP filter (4th order Butterworth filter with a low-pass cutoff at 500 Hz; constructed in MATLAB using the command “butter”) and subsequently taking the inverse Fourier Transform (Okun, 2017). The correction was only done between 0 and 500 Hz because the power of the LFP (as well as the filter) was very less beyond 500 Hz. The group delay of this Butterworth filter was  $\sim 0.8$  ms over almost the entire frequency range of interest (including the gamma range), such that the stLFP constructed from the corrected LFP signal had a similar shape as the uncorrected stLFP but was shifted leftward by  $\sim 0.8$  ms (Figure 5B vs. Figure 4B).

Note that in addition to this low-pass filter, three other filtering operations also need to be accounted for. The data acquisition system had two analog hardware filters: a high pass

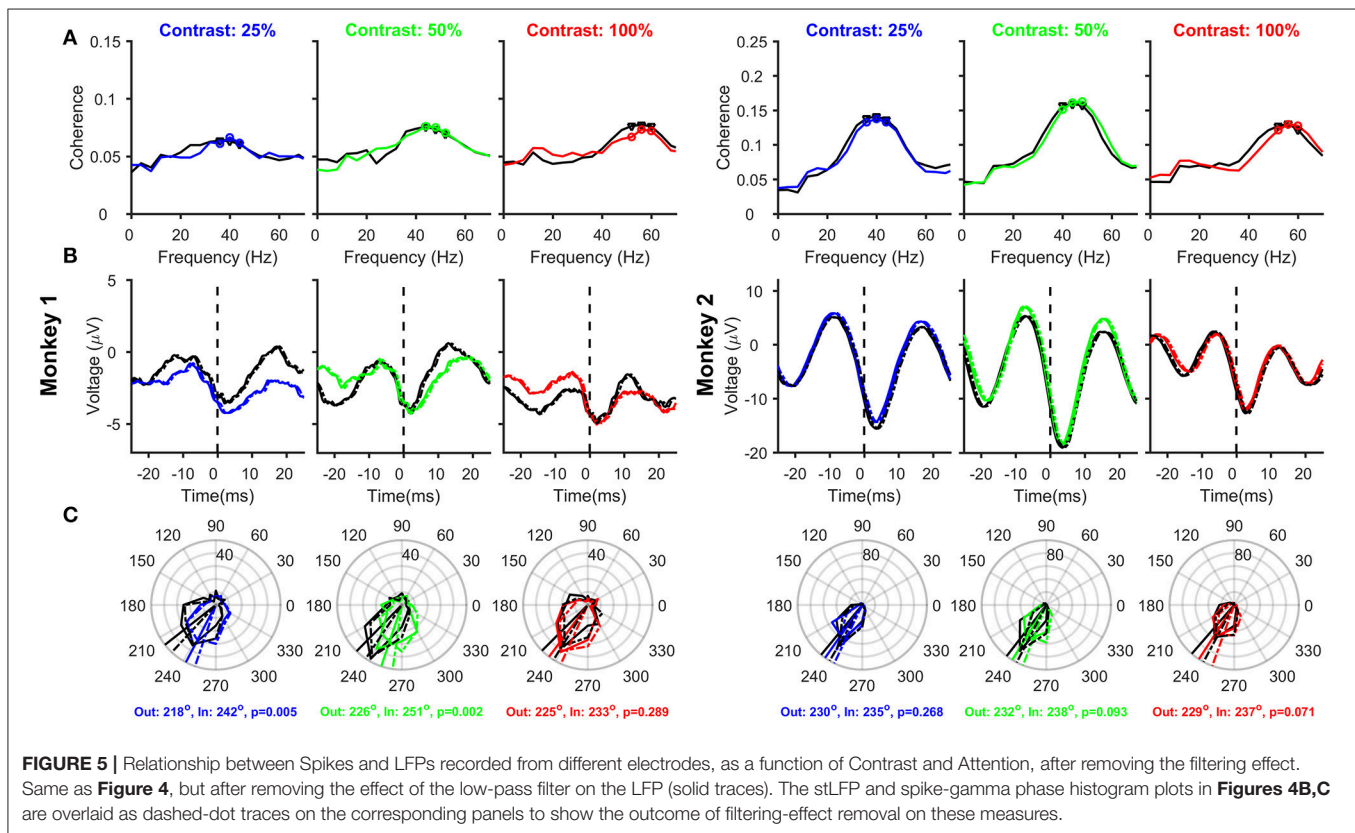
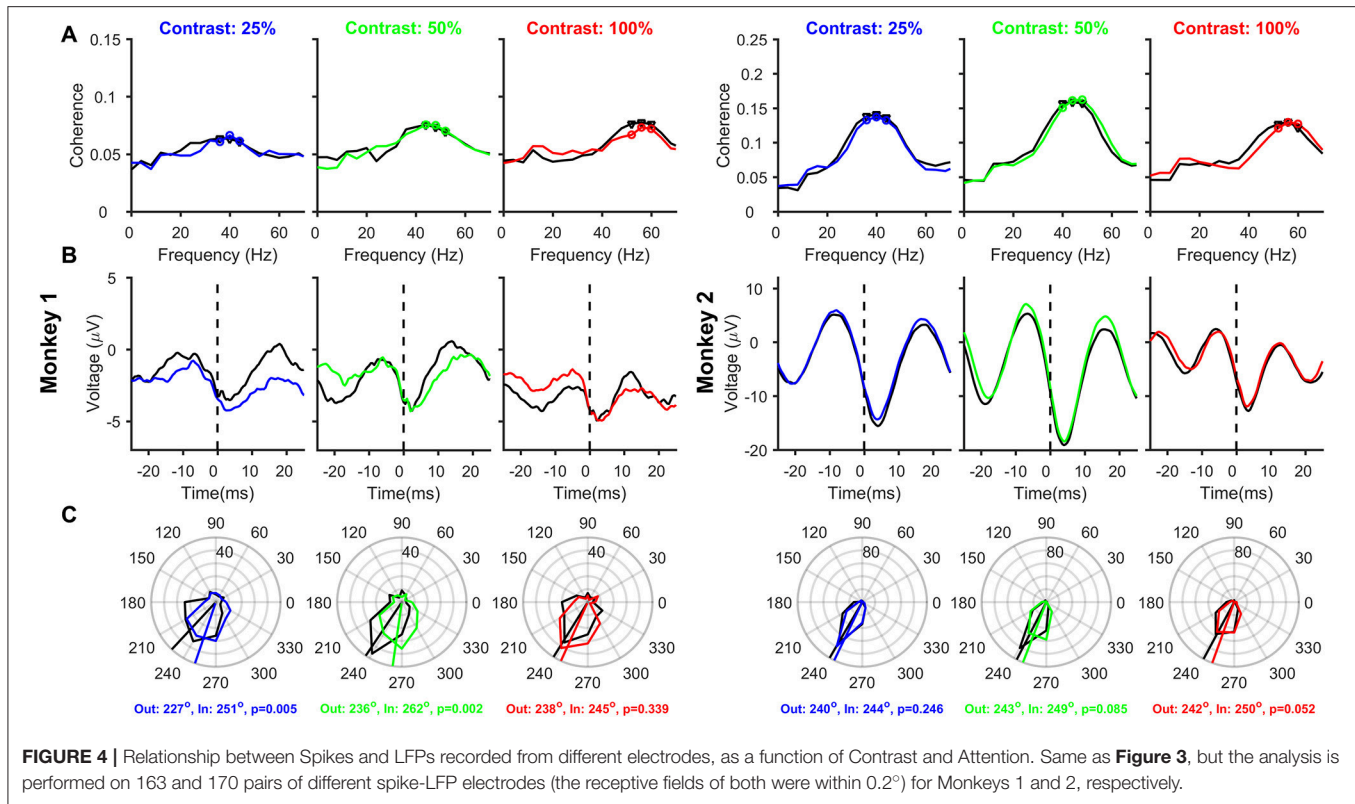
filter at 0.3 Hz (first order, Butterworth) and a low-pass filter at 7,500 Hz (third order, Butterworth). In addition, to obtain spike data, the signal was high-pass filtered at 250 Hz (fourth order, Butterworth, digital). However, all three filters had negligible group delay ( $<0.1$  ms) between 500–5,000 Hz, suggesting that these filtering operations did not change the position of the spike appreciably. Similarly, the high-pass filter had a large group delay at very low frequencies, as shown by Okun (2017), but it was negligible in the gamma range. Therefore, these three filters did not have an appreciable effect on the spike-gamma phase estimation.

## RESULTS

### Spatial Attention Reduces Alpha and Gamma Power and Increases Peak Gamma Frequency in Area V1

We first analyzed changes in alpha and gamma power and gamma peak frequency for attend-in versus attend-out conditions to test whether our results were in agreement with previous attention studies in macaque primary visual cortex (Chalk et al., 2010). Figure 2A shows the average change in PSD during the stimulus







period (150–400 ms after stimulus onset) from the pre-stimulus baseline (−300 to −50 ms), for attend-out (black trace) and attend-in condition (color trace) for 25% (blue), 50% (green), and 100% (red) contrasts. Gamma peak frequency increased with increasing contrast (traces for different contrasts are overlaid in the rightmost plot for comparison), as reported previously (Ray and Maunsell, 2010). To account for potential differences in the baseline activity due to attention, all changes were computed with respect to the baseline activity of the unattended condition (see section Materials and Methods). Consistent with previous studies, we found a strong suppression of alpha power due to attention in both monkeys, which could be observed in the baseline PSD as well (dark gray trace), confirming that the monkeys were indeed attending to the stimuli.

To test these results quantitatively, we first performed a three-way ANOVA test with factors of monkey (2 levels: Monkey 1 and 2), attention (attend-out, attend-in) and contrast (25, 50, and 100%). Alpha power was averaged over 8 and 12 Hz (inverted triangles in **Figure 2A**), while gamma power was averaged in an eight Hz band around the peak frequency for each contrast (40, 48, and 56 Hz for Monkey 1, and 40, 44, and 56 Hz for Monkey 2; **Figure 2A**; see section Materials and Methods for details). The factor monkey was significant for alpha power ( $F_{\text{monkey}} = 243.23$ ,  $p = 9.3 \times 10^{-49}$ ), gamma power ( $F_{\text{monkey}} = 417.45$ ,  $p = 1.3 \times 10^{-76}$ ), and peak gamma frequency ( $F_{\text{monkey}} = 4.52$ ,  $p = 0.004$ ). Thus, we performed a two-way ANOVA with factors attention and contrast separately for the two monkeys for alpha power, gamma power and peak gamma frequency. The effect of contrast on alpha power was not significant for Monkey 1 ( $F_{\text{contrast}} = 0.65$ ,  $p = 0.522$ ) but significant effect for Monkey 2 ( $F_{\text{contrast}} = 7.06$ ,  $p = 9 \times 10^{-4}$ ). The effect of attention on alpha power was significant for both monkeys ( $F_{\text{attention}} = 24.88$ ,  $p = 9.4 \times 10^{-7}$  for Monkey 1 and  $F_{\text{attention}} = 10.17$ ,  $p = 1.5 \times 10^{-3}$  for Monkey 2). However, there was no significant interaction between the two factors on alpha power for either monkey ( $F_{\text{contrast} \times \text{attention}} = 0.87$ ,  $p = 0.42$  for Monkey 1 and  $F_{\text{contrast} \times \text{attention}} = 0.91$ ,  $p = 0.40$  for Monkey 2). For gamma power, the effect of contrast was significant for both monkeys ( $F_{\text{contrast}} = 11.9$ ,  $p = 9.8 \times 10^{-6}$  for Monkey 1 and  $F_{\text{contrast}} = 37.78$ ,  $p = 4.6 \times 10^{-16}$  for Monkey 2) but the effect of attention was not significant ( $F_{\text{attention}} = 0.3$ ,  $p = 0.58$  for Monkey 1 and  $F_{\text{attention}} = 1.72$ ,  $p = 0.19$  for Monkey 2). Again, there was no interaction between the factors ( $F_{\text{contrast} \times \text{attention}} = 0.05$ ,  $p = 0.95$  for Monkey 1 and  $F_{\text{contrast} \times \text{attention}} = 0.17$ ,  $p = 0.85$  for Monkey 2). For peak gamma frequency, there was significant effect of both contrast and attention in both monkeys ( $F_{\text{contrast}} = 415.46$ ,  $p = 1.6 \times 10^{-95}$  for Monkey 1 and  $F_{\text{contrast}} = 973.36$ ,  $p = 7.7 \times 10^{-178}$  for Monkey 2;  $F_{\text{attention}} = 5.07$ ,  $p = 0.025$  for Monkey 1 and  $F_{\text{attention}} = 33.31$ ,  $p = 1.4 \times 10^{-8}$  for Monkey 2). The interaction of the two factors was significant only in Monkey 2 ( $F_{\text{contrast} \times \text{attention}} = 1.11$ ,  $p = 0.33$  for Monkey 1 and  $F_{\text{contrast} \times \text{attention}} = 3.37$ ,  $p = 0.04$  for Monkey 2). Similar results were obtained in a 2-factor ANOVA performed on the data pooled over the two monkeys.

These results were further confirmed using pairwise t-tests. In almost all conditions, alpha power significantly reduced with attention (**Figure 2B**; Monkey 1:  $t$ -test,  $N = 63$ ,  $p = 3.9 \times$

$10^{-7}$ ,  $1.1 \times 10^{-13}$ ,  $3.5 \times 10^{-5}$  for 25, 50, and 100% contrasts respectively;  $p = 8.9 \times 10^{-22}$  for all the contrast conditions combined; Monkey 2:  $t$ -test,  $N = 89$ ,  $p = 0.057$ ,  $2.4 \times 10^{-14}$ ,  $3.1 \times 10^{-13}$  and  $8.6 \times 10^{-20}$  for 25, 50, 100% contrasts and the combined condition, respectively). The reduction in gamma power was significant only for the 50% contrast condition for Monkey 1 ( $t$ -test,  $N = 63$ ,  $p = 0.41$ , 0.013, 0.86, and 0.052 for 25, 50, 100% and combined contrast conditions, respectively), and for 25 and 100% contrasts for Monkey 2 ( $t$ -test,  $N = 89$ ,  $p = 3.4 \times 10^{-8}$ , 0.72, 0.014 and  $1.5 \times 10^{-6}$  for 25, 50, 100% and combined contrast conditions, respectively). Similarly, the increase in gamma peak frequency was significant only for 25% contrast condition for Monkey 1 ( $t$ -test,  $N = 63$ ,  $p = 0.007$ , 0.15, 0.13, and  $7.8 \times 10^{-4}$  for 25, 50, 100% and combined contrast conditions, respectively), and all contrasts for Monkey 2 ( $t$ -test,  $N = 89$ ,  $p = 1.7 \times 10^{-9}$ ,  $6.7 \times 10^{-11}$ , 0.013 and  $6.7 \times 10^{-19}$  for 25, 50, 100% and combined contrast conditions, respectively).

The weak effect of attention on gamma power is not surprising for two reasons. First, because we recorded from a chronically implanted microelectrode array, the stimuli were optimized only for a single site in each session and therefore were non-optimal for most electrodes, unlike the study by Chalk et al. (2010) where stimuli were better optimized. Second, since the stimuli were only presented for 400 ms (to minimize attentional fluctuations within the stimulus duration) and the analysis duration was only 250 ms, the frequency resolution was 4 Hz, which made it difficult to correctly estimate peak frequency shifts that are typically only 2–3 Hz in V1 (Ray and Maunsell, 2010; Bosman et al., 2012). Note that the second limitation can be partially overcome by using Matching Pursuit (Chandran et al., 2016), which allowed us to better characterize the gamma peak frequency shifts in a previous study (see Supplementary Figure 2 of Ray and Maunsell, 2010); we have used multi-taper analysis here because the spike-field coherence (SFC), which was also used to get spike-gamma phase, was obtained using the same technique. In general, the effects of attention on V1 were consistent with the findings of Chalk et al. (2010), and were almost always significant when the results were pooled across contrasts.

### Effect of Contrast and Attention on SFC, stLFP, and Spike-Gamma Phase Computed Using Spike and LFP Recorded From the Same Electrode

Next, we analyzed how attention modulated SFC, stLFP and spike-gamma phase when spikes and LFP were recorded from the same electrode (23 and 39 sites for the two monkeys; see Materials and Methods for details), as was the case in the study by Chalk et al. (2010). The magnitude of the SFC (**Figure 3A**) showed clear peaks in the gamma frequency range, and the peak gamma frequency shifted with an increase in contrast. Consistent with Chalk et al. (2010) and the results obtained using power (**Figure 2**), we found a reduction in SFC magnitude and an increase in peak gamma frequency with attention in almost all conditions. The stLFP plots (**Figure 3B**) showed the presence of a prominent rhythm around the time of the spike, especially for Monkey 2, whose trough was shifted 3–4 ms away from

zero. These results were reflected in the spike-gamma phase histograms (**Figure 3C**), obtained by taking the average angle of the SFC across the three frequency bins around the peak gamma frequency (as highlighted in **Figure 3A**). Following the convention used by Chalk and colleagues, phase angles were defined such that trough of the gamma rhythm was at  $180^\circ$  and rightward shift of the trough increased the phase angle. The mean phase angles were  $\sim 210^\circ$  for Monkey 1 and  $\sim 235^\circ$  for Monkey 2 and were not significantly different across attention conditions (circular mean phases and the associated p-values obtained from Watson-Williams test are shown in the legend). Even when pooled across contrasts, the mean phases between attend-in and attend-out conditions were not significantly different (Watson-Williams multi-sample test,  $p = 0.28$  and  $0.33$  for Monkeys 1 and 2). Similarly, the mean phases at different contrasts were not significantly different from each other in either attend-out or attend-in conditions (Watson-Williams multi-sample test,  $p = 0.89$  (attend-out) and  $p = 0.9$  (attend-in) for Monkey 1;  $p = 0.66$  (attend-out) and  $p = 0.55$  (attend-in) for Monkey 2). These results are consistent with Chalk et al. (2010), who obtained a median phase of  $\sim -0.65\pi$ , which translates to  $\sim 243^\circ$ . An offset of  $\sim 30^\circ$ – $50^\circ$  from the trough ( $180^\circ$ ) is also consistent with the findings of Vinck et al. (2010) and Ray and Maunsell (2010), although in these two studies the convention was chosen such that rightward shift of the trough led to a reduction of phase angle below  $180^\circ$  (such that the phase angles were between  $\sim 130^\circ$  and  $\sim 150^\circ$ ).

Although our results are consistent with previous studies, there are two serious flaws in these results, which can be clearly observed in the stLFP plots (**Figure 3B**). First, there is a large spike-related transient (sharp negative dip near time zero), which biases the estimation of the gamma phase. Specifically, this transient can be decomposed into a series of sinusoids with their troughs aligned to the trough of the transient, effectively “pulling” the phase of any true phase-locked rhythm toward  $180^\circ$  (for a detailed discussion, see Ray, 2015). This can be observed in the two monkeys: the estimated spike-gamma phase is closer to  $180^\circ$  for Monkey 1 compared to Monkey 2 ( $\sim 210^\circ$  vs.  $\sim 235^\circ$ ), simply because the relative magnitude of the transient compared to the gamma rhythm is larger for Monkey 1. The second flaw is that the spike-related transient, which should be around the time of the spike itself, is shifted toward positive values. We address both these concerns below.

### Effect of Using Different Electrodes for Spikes and LFP on SFC, stLFP, and Spike-Gamma Phase

One popular method to reduce the spike-related transient is to take spikes and LFP from different electrodes (Ray and Maunsell, 2010; Vinck et al., 2010; Ray, 2015). We, therefore, repeated the analysis on 163 and 170 “different” spike-LFP pairs for the two monkeys, such that the receptive fields of both were located within  $0.2^\circ$  of stimulus center (see section Materials and Methods for details). Mean SFC showed similar results as before, with clear peaks in the gamma frequency range and an increase in peak gamma frequency with increasing

contrast, and a slight reduction in SFC magnitude and an increase in peak frequency with attention in some cases. Spike-related transient, which was prominent in **Figure 3B**, was now much reduced, better revealing the true gamma rhythm in the stLFP (**Figure 4B**) whose trough was 3–4 ms after the spike in both the monkeys. Mean spike-gamma phases were now  $\sim 235^\circ$  and  $\sim 245^\circ$  for the two monkeys (**Figure 4C**; note that the shift in mean phase between **Figures 3, 4** is much larger for Monkey 1 because the spike transient was relatively much larger for that monkey). Interestingly, for both monkeys and for all contrast conditions, attention appeared to shift the mean gamma-spike phase away from  $180^\circ$ . Although this phase difference did not reach significance for many contrast levels (circular means and p-values obtained using Watson-Williams multi-sample test are shown in the bottom of **Figure 4C**), the phase differences were highly significant when combined across contrasts (Watson-Williams multi-sample test,  $p = 7 \times 10^{-5}$  and  $5 \times 10^{-3}$  for Monkeys 1 and 2), albeit the actual magnitude of the difference was small ( $\sim 19^\circ$  and  $\sim 6^\circ$ ). The mean phases at different contrasts were not significantly different from each other in either the attend-out or the attend-in condition [Watson-Williams multi-sample test,  $p = 0.36$  (attend-out),  $p = 0.11$  (attend-in) for Monkey 1 and  $p = 0.7$  (attend-out),  $p = 0.34$  (attend-in) for Monkey 2].

### Effect of Removing the Filtering Artifact on SFC, stLFP, and Spike-Gamma Phase Relation

The rightward shift of the spike-related transient away from zero (**Figures 3B, 4B**) is simply due to the effect of the filtering operation to obtain the LFP. We, therefore, removed this filtering effect (see Materials and Methods for details) and reanalyzed SFC, stLFP and spike-gamma phase for “different” pair condition (**Figure 5**; for the “same” electrode condition, this operation caused the trough of the spike-transient to shift near zero; data not shown). While this operation did not change any of the results shown in **Figure 4**, the mean phases decreased by  $\sim 10^\circ$  at both 25 and 50% and  $\sim 13^\circ$  at 100% contrast (for the same shift in time, the shift in degrees depends on the frequency of the rhythm; **Figures 4B,C** are overlaid as dashed-dot traces on the corresponding panels in **Figure 5** to show the outcome of filtering-effect removal). Otherwise, like **Figure 4**, the effect of attention on spike-gamma phase remained significant when phases were pooled across contrast conditions (Watson-Williams multi-sample test,  $p = 6 \times 10^{-5}$  for Monkey 1 and  $8 \times 10^{-3}$  for Monkey 2). Similarly, the mean phases at different contrasts were not significantly different from each other in either the attend-out or the attend-in condition [Watson-Williams multi-sample test,  $p = 0.58$  (attend-out),  $p = 0.07$  (attend-in) for Monkey 1 and  $p = 0.69$  (attend-out),  $p = 0.73$  (attend-in) for Monkey 2].

## DISCUSSION

We investigated whether increasing stimulus contrast or allocating more attention to a stimulus (which increases its

effective contrast) shifts the position of the spike relative to the phase of the gamma rhythm, as posited by the PC hypothesis. We highlighted two issues that can bias the phase estimation: the presence of the spike-related transient and the effect of filtering to obtain the LFP. After accounting for these issues, we found no effect of stimulus contrast and a weak but significant effect of attention on spike-gamma phase. Although these results are consistent with the PC hypothesis in the context of attention, we discuss three issues that severely limit the efficacy of gamma PC in V1.

### Issue 1: Magnitude of Gamma PC in V1

For a rhythm occurring at 50 Hz (time period of 20 ms), the interval between the peak and the subsequent trough (the interval over which the inhibition fades away) is 10 ms, which is the maximum range over which PC can operate. It is clear from our results, as well as prior reports, that even if PC occurs, it only uses a small sub-interval within this interval. Since spikes occur away from the trough of the rhythm with increasing stimulus intensity under PC, the delay of the trough from the spike at 100% contrast sets the dynamic range of this coding scheme. In our data, spikes occurred at  $\sim 230^\circ$  at 100% contrast, similar to the value reported by Chalk et al. (2010) ( $\sim 240^\circ$ ) and Vinck et al. (2010) ( $\sim 137^\circ$  for preferred orientation, which translates to  $\sim 223^\circ$  as per our convention). For a rhythm at  $\sim 50$  Hz, a shift of  $\sim 50^\circ$  translates to only  $\sim 3$  ms out of the available  $\sim 10$  ms for coding. Further, even when contrast was reduced to 25%, there was no discernable change in the trough position. The only study that did show any evidence of phase coding (Vinck et al., 2010) showed a shift of  $\sim 20^\circ$  between the best and worst orientation, which translates to only  $\sim 1$  ms shift (in addition, see other issues with their results below). In our data, the shift in phase due to attention is even lesser, especially for Monkey 2 (in addition, see Issue 3 below). It can be argued that the phase could shift down to  $180^\circ$  for very low contrasts (providing a dynamic range of  $\sim 3$  ms), but it is well known that gamma rhythm itself is weak or absent at very low contrasts (Henrie and Shapley, 2005; Jia et al., 2013) and also peaks at a lower frequency (see Issue 3). Thus, if we consider the range of contrasts for which gamma is reliable, the magnitude of PC (i.e., the range over which the spike varies with respect to the rhythm) appears to be very small in V1. In this context, our filtering correction becomes significant, since even though the group delay is only  $\sim 0.8$  ms, it still decreases the dynamics range by a further  $\sim 20$ – $25\%$ .

### Issue 2: Effect of Changing Gamma Amplitude

As shown in Figure 3, spikes are associated with a transient in the LFP recorded from the same electrode, which biases the spike-LFP phase analysis. Because the amplitude of an extracellular action potential generally decreases rapidly as a microelectrode is moved away from the neuron (Gold et al., 2006; Schomburg et al., 2012), the spatial spread of a spike is thought to be very local (for example, Xing and colleagues used a range between 30 and  $100\ \mu\text{m}$  for single units; Xing et al., 2009). Therefore,

one way to reduce the spike transient is to take the LFP from a neighboring electrode that is separated from the spike electrode by at least a few hundred microns (for a representative case, see Vinck et al., 2010). There are, however, two issues with this approach. First, although taking spikes and LFPs from different electrodes drastically reduces the spike-related transient, it does not completely eliminate it (Ray, 2015). For example, as shown in Figures 2A,E of Ray and Maunsell (2011b) where stLFPs were constructed using spikes and LFP electrodes separated by different distances for the same two monkeys as used in this study, the spike-transient could be seen up to electrode pairs separated by  $\sim 400\ \mu\text{m}$  for Monkey 1 and  $\sim 0.4$ – $1.6$  mm for Monkey 2, albeit the magnitude of the spike-transient was much smaller than when stLFP was constructed from the same electrode ( $d = 0$  condition in those plots). This happens because neurons near the LFP electrode are often correlated with the neuron being recorded from the spike electrode, and those neurons produce a transient in the LFP electrode that are locked to the spikes on the spike electrode. The second issue is that this procedure implicitly assumes that gamma oscillations recorded from two nearby electrodes are similar, but the spatial spread of LFP itself is a topic of debate. While some studies have shown that the spatial spread of LFP could be large (up to a few mm; Kajikawa and Schroeder, 2011), others have shown that it could be only a few hundred microns (Katzner et al., 2009; Xing et al., 2009; Dubey and Ray, 2016). A modeling study showed that the spread could depend on the degree of correlation in the neural population (Lindén et al., 2011). Consequently, there might be differences in the gamma recorded from neighboring microelectrodes. For example, we have shown that when a Gabor stimulus is presented, two microelectrodes separated by as little as  $0.2^\circ$  can exhibit significantly different center frequencies (Ray and Maunsell, 2010). Therefore, some studies have used other techniques to remove the spike-transient, such as Matching Pursuit (Ray et al., 2008) or a Bayesian Framework (Zanos et al., 2011). All these methods substantially reduce the spike-transient, although it is unlikely that they completely eliminate it.

A small spike-transient is unlikely to influence the estimation of gamma phase when the rhythm itself is very strong but may shift the phase toward  $180^\circ$  when the rhythm is weak. For example, even when spikes and LFPs were recorded from separate electrodes (Figure 4), the mean phases for Monkey 1 were about  $\sim 10^\circ$  less than Monkey 2, who had a much stronger gamma rhythm than Monkey 1. A visual inspection of the stLFP (Figure 4B) reveals a small spike-transient like structure in Monkey 1, which could have contributed to the reduction in spike-gamma phase as compared to Monkey 2. Importantly, in cases where the magnitude of gamma itself varies across conditions, an apparent shift in spike-gamma phase could just be due to a differential contribution of the spike-transient which “pulls” the phase toward  $180^\circ$ . For example, Vinck et al. (2010) showed that gamma PC was stronger when gamma power and gamma phase locking was very weak (see their Figure 6). Because they did not show the stLFPs, it is unclear whether the apparent phase shift they documented was because of a genuine leftward shift of the gamma trough or the presence of a spike-transient



whose contribution was larger when the gamma rhythm itself was weak.

### Issue 3: Effect of Changing Gamma Peak Frequency

PC hypothesis makes sense when the rhythm has a stable frequency. However, the center frequency of gamma rhythm varies systematically with changes in a variety of stimulus parameters, such as size (Gieselmann and Thiele, 2008; Ray and Maunsell, 2011a; Jia et al., 2013), contrast (Ray and Maunsell, 2010; Bosman et al., 2012; Jia et al., 2013), and drift rates (Gray and Viana Di Prisco, 1997; Friedman-Hill et al., 2000). For example, although we show that the spike-gamma phase angles do not vary with stimulus contrast, note that these angles are computed for different gamma frequencies, making it harder to interpret and compare these phase values. Vinck et al. (2010) used different orientations for comparison, but gamma center frequencies can vary even for different orientations, although the trends are not always consistent (see Figure 2D of Jia et al., 2013 and Figure 1 of Murty et al., 2018). For the same delay between the spike and gamma trough, the effective phase angle is greater when the rhythm is faster. For example, in our data, the stLFP troughs appear to coincide between the attend-in and attend-out cases in almost all conditions (Figure 5B). However, since attention slightly increases the gamma frequency, the effective phase lag in degrees could be larger, which could explain the small but consistent increase in phase angles.

We note, however, that we computed phase over a 250 ms window (similar results were obtained for 200 ms window), which cover more than 10 cycles of the rhythm. During natural vision, we make 3–4 saccades every second (even during fixation, we make several micro-saccades per second), and such eye movements can change or reset the phase of LFP oscillations (Bosman et al., 2009; Ito et al., 2011). It is possible that PC occurs within a single or a few cycles of gamma rhythm, for which gamma need not even have stable frequency over time. It is also possible that PC occurs differently during natural viewing as opposed to a paradigm where animals are trained to fixate for long durations. For example, Ito and colleagues showed that in freely viewing monkeys, fixation-related spike synchronization occurred at an early phase of the rate response after fixation-onset, and the first spikes after the onset of a fixation were locked to a specific epoch of the LFP modulation (Ito et al., 2011). Other studies have also shown that gamma rhythm tends to appear in short bursts over a few cycles (Xing et al., 2012; Lundqvist et al., 2016; Chandran Ks et al., 2017), and therefore PC could theoretically occur over shorter duration than what was considered here. Comparable recordings from monkeys during natural viewing conditions as well as advanced signal processing techniques are required to test this hypothesis.

### Weak Effect of Attention in V1

The effect of attention was weaker in our data than the findings of Chalk et al. (2010), possibly due to the use

of sub-optimal stimuli for many sites, fewer sites, and a shorter analysis window. However, it is unlikely that our results would change drastically if these limitations could be overcome. First, the effect of attention on gamma in V1 is in general weak (Chalk et al., 2010; Buffalo et al., 2011). Second, although the reduction in gamma power and SFC with attention were small, we obtained a pronounced reduction in alpha power in all cases. Similarly, the increase in gamma peak frequency (1–3 Hz in our data) was comparable to a previous study by Bosman and colleagues, who reported an increase in gamma peak frequency of 2–3 Hz (Bosman et al., 2012). Third, although the analysis window was shorter than previous studies, which yielded a poor frequency resolution, the stLFP plots were computed in the time-domain itself and therefore did not suffer from the poor frequency resolution, but even these did not show a substantial rightward shift as is expected from the PC hypothesis. Fourth, while we had fewer recording sites that may have yielded less statistical power for power analysis (Figure 2), we had a substantial number of pairs (163 and 170 for the two monkeys), so the main result regarding the PC hypothesis (Figure 5) did not suffer from the lack of statistical power. Finally, although the effect of attention was weak, contrast had a strong effect on gamma power and frequency, but the PC hypothesis for contrast did not yield a significant result.

In summary, although we did find a weak effect of attention on spike-gamma phase relationship, based on the variety of issues that we have discussed, gamma PC is at best expected to play a minor role in the coding the stimulus contrast in V1.

## AUTHOR CONTRIBUTIONS

AD and SR conceived the idea of research. SR collected data; AD and SR analyzed data. AD and SR wrote the paper.

## FUNDING

This work was supported by Wellcome Trust/DBT India Alliance (Grant 500145/Z/09/Z; intermediate Fellowship to SR), Tata Trusts, Department of Biotechnology-Indian Institute of Science (DBT-IISc) Partnership Programme and a Junior Research Fellowship awarded by IISc from Ministry of Human Resource Development, Government of India (to AD).

## ACKNOWLEDGMENTS

We thank Dr. John Maunsell for his help in experimental design and data collection and Steven Slebocka and Vivian Imamura for technical support.



## REFERENCES

- Bartos, M., Vida, I., and Jonas, P. (2007). Synaptic mechanisms of synchronized gamma oscillations in inhibitory interneuron networks. *Nat. Rev. Neurosci.* 8, 45–56. doi: 10.1038/nrn2044
- Berens, P. (2009). CircStat: a MATLAB Toolbox for circular statistics. *J. Stat. Softw.* 31, 1–21. doi: 10.18637/jss.v031.i10
- Berens, P., Keliris, G. A., Ecker, A. S., Logothetis, N. K., and Tolias, A. S. (2008). Comparing the feature selectivity of the gamma-band of the local field potential and the underlying spiking activity in primate visual cortex. *Front. Syst. Neurosci.* 2:2. doi: 10.3389/neuro.06.002.2008
- Bokil, H., Andrews, P., Kulkarni, J. E., Mehta, S., and Mitra, P. P. (2010). Chronux: a platform for analyzing neural signals. *J. Neurosci. Methods* 192, 146–151. doi: 10.1016/j.jneumeth.2010.06.020
- Bosman, C. A., Schoffelen, J.-M., Brunet, N., Oostenveld, R., Bastos, A. M., Womelsdorf, T., et al. (2012). Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron* 75, 875–888. doi: 10.1016/j.neuron.2012.06.037
- Bosman, C. A., Womelsdorf, T., Desimone, R., and Fries, P. (2009). A microsaccadic rhythm modulates gamma-band synchronization and behavior. *J. Neurosci.* 29, 9471–9480. doi: 10.1523/JNEUROSCI.1193-09.2009
- Buffalo, E. A., Fries, P., Landman, R., Buschman, T. J., and Desimone, R. (2011). Laminar differences in gamma and alpha coherence in the ventral stream. *Proc. Natl. Acad. Sci. U.S.A.* 108, 11262–11267. doi: 10.1073/pnas.1011284108
- Buzsáki, G. (2006). *Rhythms of the Brain*. New York, NY: Oxford University Press.
- Buzsáki, G., and Chrobak, J. J. (1995). Temporal structure in spatially organized neuronal ensembles: a role for interneuronal networks. *Curr. Opin. Neurobiol.* 5, 504–510. doi: 10.1016/0959-4388(95)80012-3
- Buzsáki, G., Logothetis, N., and Singer, W. (2013). Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron* 80, 751–764. doi: 10.1016/j.neuron.2013.10.002
- Cardin, J. A., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., et al. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature* 459, 663–667. doi: 10.1038/nature08002
- Carrasco, M., Ling, S., and Read, S. (2004). Attention alters appearance. *Nat. Neurosci.* 7, 308–313. doi: 10.1038/nn1194
- Chalk, M., Herrero, J. L., Gieselmann, M. A., Delicato, L. S., Gotthardt, S., and Thiele, A. (2010). Attention reduces stimulus-driven gamma frequency oscillations and spike field coherence in V1. *Neuron* 66, 114–125. doi: 10.1016/j.neuron.2010.03.013
- Chandran Ks, S., Seelamantula, C. S., and Ray, S. (2017). Duration analysis using matching pursuit algorithm reveals longer bouts of gamma rhythm. *J. Neurophysiol.* 119, 808–821. doi: 10.1152/jn.00154.2017
- Chandran, K. S. S., Mishra, A., Shirhatti, V., and Ray, S. (2016). Comparison of matching pursuit algorithm with other signal processing techniques for computation of the time-frequency power spectrum of brain signals. *J. Neurosci.* 36, 3399–3408. doi: 10.1523/JNEUROSCI.3633-15.2016
- Dubey, A., and Ray, S. (2016). Spatial spread of local field potential is band-pass in the primary visual cortex. *J. Neurophysiol.* 116, 1986–1999. doi: 10.1152/jn.00443.2016
- Friedman-Hill, S., Maldonado, P. E., and Gray, C. M. (2000). Dynamics of striate cortical activity in the alert macaque: I. Incidence and stimulus-dependence of gamma-band neuronal oscillations. *Cereb. Cortex* 10, 1105–1116. doi: 10.1093/cercor/10.11.1105
- Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034
- Fries, P., Nikolić, D., and Singer, W. (2007). The gamma cycle. *Trends Neurosci.* 30, 309–316. doi: 10.1016/j.tins.2007.05.005
- Fries, P., Reynolds, J. H., Rorie, A. E., and Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291, 1560–1563. doi: 10.1126/science.1055465
- Gieselmann, M. A., and Thiele, A. (2008). Comparison of spatial integration and surround suppression characteristics in spiking activity and the local field potential in macaque V1. *Eur. J. Neurosci.* 28, 447–459. doi: 10.1111/j.1460-9568.2008.06358.x
- Gold, C., Henze, D. A., Koch, C., and Buzsáki, G. (2006). On the origin of the extracellular action potential waveform: a modeling study. *J. Neurophysiol.* 95, 3113–3128. doi: 10.1152/jn.00979.2005
- Gray, C. M., and Viana Di Prisco, G. (1997). Stimulus-dependent neuronal oscillations and local synchronization in striate cortex of the alert cat. *J. Neurosci.* 17, 3239–3253. doi: 10.1523/JNEUROSCI.17-09-03239.1997
- Gregoriou, G. G., Gotts, S. J., Zhou, H., and Desimone, R. (2009). High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* 324, 1207–1210. doi: 10.1126/science.1171402
- Henrie, J. A., and Shapley, R. (2005). LFP power spectra in V1 cortex: the graded effect of stimulus contrast. *J. Neurophysiol.* 94, 479–490. doi: 10.1152/jn.00919.2004
- Hubel, D. H., and Wiesel, T. N. (1977). Ferrier lecture. functional architecture of macaque monkey visual cortex. *Proc. R. Soc. Lond. B Biol. Sci.* 198, 1–59. doi: 10.1098/rspb.1977.0085
- Ito, J., Maldonado, P., Singer, W., and Grün, S. (2011). Saccade-related modulations of neuronal excitability support synchrony of visually elicited spikes. *Cereb. Cortex* 21, 2482–2497. doi: 10.1093/cercor/bhr020
- Jia, X., Smith, M. A., and Kohn, A. (2011). Stimulus selectivity and spatial coherence of gamma components of the local field potential. *J. Neurosci. Off. J. Soc. Neurosci.* 31, 9390–9403. doi: 10.1523/JNEUROSCI.0645-11.2011
- Jia, X., Xing, D., and Kohn, A. (2013). No consistent relationship between gamma power and peak frequency in macaque primary visual cortex. *J. Neurosci.* 33, 17–25. doi: 10.1523/JNEUROSCI.1687-12.2013
- Kajikawa, Y., and Schroeder, C. E. (2011). How local is the local field potential? *Neuron* 72, 847–858. doi: 10.1016/j.neuron.2011.09.029
- Katzner, S., Nauhaus, I., Benucci, A., Bonin, V., Ringach, D. L., and Carandini, M. (2009). Local origin of field potentials in visual cortex. *Neuron* 61, 35–41. doi: 10.1016/j.neuron.2008.11.016
- Kelly, R. C., Smith, M. A., Samonds, J. M., Kohn, A., Bonds, A. B., Movshon, J. A., et al. (2007). Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex. *J. Neurosci.* 27, 261–264. doi: 10.1523/JNEUROSCI.4906-06.2007
- Lindén, H., Tetzlaff, T., Potjans, T. C., Pettersen, K. H., Grün, S., Diesmann, M., et al. (2011). Modeling the spatial reach of the LFP. *Neuron* 72, 859–872. doi: 10.1016/j.neuron.2011.11.006
- Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., and Miller, E. K. (2016). Gamma and beta bursts underlie working memory. *Neuron* 90, 152–164. doi: 10.1016/j.neuron.2016.02.028
- Murty, D. V. P. S., Shirhatti, V., Ravishankar, P., and Ray, S. (2018). Large visual stimuli induce two distinct gamma oscillations in primate visual cortex. *J. Neurosci.* 38, 2270–2277. doi: 10.1523/JNEUROSCI.2270-17.2017
- O’Keefe, J., and Recce, M. L. (1993). Phase relationship between hippocampal place units and the EEG theta rhythm. *Hippocampus* 3, 317–330.
- Okun, M. (2017). Artefactual origin of biphasic cortical spike-LFP correlation. *J. Comput. Neurosci.* 42, 31–35. doi: 10.1007/s10827-016-0625-5
- Ray, S. (2015). Challenges in the quantification and interpretation of spike-LFP relationships. *Curr. Opin. Neurobiol.* 31, 111–118. doi: 10.1016/j.conb.2014.09.004
- Ray, S., Hsiao, S. S., Crone, N. E., Franaszczuk, P. J., and Niebur, E. (2008). Effect of stimulus intensity on the spike-local field potential relationship in the secondary somatosensory cortex. *J. Neurosci. Off. J. Soc. Neurosci.* 28, 7334–7343. doi: 10.1523/JNEUROSCI.1588-08.2008
- Ray, S., and Maunsell, J. H. (2010). Differences in gamma frequencies across visual cortex restrict their possible use in computation. *Neuron* 67, 885–896. doi: 10.1016/j.neuron.2010.08.004
- Ray, S., and Maunsell, J. H. (2011a). Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol.* 9:e1000610. doi: 10.1371/journal.pbio.1000610
- Ray, S., and Maunsell, J. H. (2011b). Network rhythms influence the relationship between spike-triggered local field potential and functional connectivity. *J. Neurosci. Off. J. Soc. Neurosci.* 31, 12674–12682. doi: 10.1523/JNEUROSCI.1856-11.2011
- Ray, S., and Maunsell, J. H. (2015). Do gamma oscillations play a role in cerebral cortex? *Trends Cogn. Sci.* 19, 78–85. doi: 10.1016/j.tics.2014.12.002
- Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B., and Varela, F. J. (1999). Perception’s shadow: long-distance synchronization of human brain activity. *Nature* 397, 430–433.
- Schomburg, E. W., Anastassiou, C. A., Buzsáki, G., and Koch, C. (2012). The spiking component of oscillatory extracellular

- potentials in the rat hippocampus. *J. Neurosci.* 32, 11798–11811. doi: 10.1523/JNEUROSCI.0656-12.2012
- Shirhatti, V., Borthakur, A., and Ray, S. (2016). Effect of reference scheme on power and phase of the local field potential. *Neural Comput.* 28, 882–913. doi: 10.1162/NECO\_a\_00827
- Singer, W. (1999). Neuronal synchrony: a versatile code for the definition of relations? *Neuron* 24, 49–65, 111–125.
- Sohal, V. S., Zhang, F., Yizhar, O., and Deisseroth, K. (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* 459, 698–702. doi: 10.1038/nature07991
- Vinck, M., Lima, B., Womelsdorf, T., Oostenveld, R., Singer, W., Neuenschwander, S., et al. (2010). Gamma-phase shifting in awake monkey visual cortex. *J. Neurosci.* 30, 1250–1257. doi: 10.1523/JNEUROSCI.1623-09.2010
- Xing, D., Shen, Y., Burns, S., Yeh, C.-I., Shapley, R., and Li, W. (2012). Stochastic generation of gamma-band activity in primary visual cortex of awake and anesthetized monkeys. *J. Neurosci.* 32, 13873–13880a. doi: 10.1523/JNEUROSCI.5644-11.2012
- Xing, D., Yeh, C. I., and Shapley, R. M. (2009). Spatial spread of the local field potential and its laminar variation in visual cortex. *J. Neurosci.* 29, 11540–11549. doi: 10.1523/JNEUROSCI.2573-09.2009
- Zanos, T. P., Mineault, P. J., and Pack, C. C. (2011). Removal of spurious correlations between spikes and local field potentials. *J. Neurophysiol.* 105, 474–486. doi: 10.1152/jn.00642.2010

**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 © 2018 Das and Ray. 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) and the copyright owner(s) 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.



# Gender-Dependent Changes in Time Production Following Quadrato Motor Training in Dyslexic and Normal Readers

Tal Dotan Ben-Soussan<sup>1\*</sup> and Joseph Glicksohn<sup>2,3</sup>

<sup>1</sup>Research Institute for Neuroscience, Education and Didactics, Patrizio Paoletti Foundation for Development and Communication, Assisi, Italy, <sup>2</sup>Department of Criminology, Bar-Ilan University, Ramat Gan, Israel, <sup>3</sup>The Leslie and Susan Gonda (Goldschmied) Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat Gan, Israel

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

Marc Wittmann,  
Institut für Grenzgebiete der  
Psychologie und Psychohygiene,  
Germany  
Matthew S. Matell,  
Villanova University, United States

### \*Correspondence:

Tal Dotan Ben-Soussan  
research@  
fondazionepatriziopaoletti.org

**Received:** 26 March 2018

**Accepted:** 08 August 2018

**Published:** 29 August 2018

### Citation:

Ben-Soussan TD and Glicksohn J  
(2018) Gender-Dependent Changes  
in Time Production Following  
Quadrato Motor Training in Dyslexic  
and Normal Readers.  
*Front. Comput. Neurosci.* 12:71.  
doi: 10.3389/fncom.2018.00071

Time estimation is an important component of the ability to organize and plan sequences of actions as well as cognitive functions, both of which are known to be altered in dyslexia. While attention deficits are accompanied by short Time Productions (TPs), expert meditators have been reported to produce longer durations, and this seems to be related to their increased attentional resources. In the current study, we examined the effects of a month of Quadrato Motor Training (QMT), which is a structured sensorimotor training program that involves sequencing of motor responses based on verbal commands, on TP using a pre-post design. QMT has previously been found to enhance attention and EEG oscillatory activity, especially within the alpha range. For the current study, 29 adult Hebrew readers were recruited, of whom 10 dyslexic participants performed the QMT. The normal readers were randomly assigned to QMT ( $n = 9$ ) or Verbal Training (VT, identical cognitive training with no overt motor component, and only verbal response,  $n = 10$ ). Our results demonstrate that in contrast to the controls, longer TP in females was found following 1 month of intensive QMT in the dyslexic group, while the opposite trend occurred in control females. We suggest that this longer TP in the female dyslexics is related to their enhanced attention resulting from QMT. The current findings suggest that the combination of motor and mindful training, embedded in QMT, has a differential effect depending on gender and whether one is dyslexic or not. These results have implications for educational and contemplative neuroscience, emphasizing the connection between specifically-structured motor training, time estimation and attention.

**Keywords:** time production, quadrato motor training, dyslexia, gender difference, time and motion studies

## INTRODUCTION

Timing deficits in dyslexia include those concerned with time estimation (Nicolson et al., 1995; Ramus et al., 2003; Hölzel et al., 2011), rhythm tapping (Wolff et al., 1990; Wolff, 2002), detecting complex timing patterns (Kujala et al., 2007), rapid temporal processing (Tallal et al., 1993), auditory temporal sensitivity (Witton et al., 1998) and visual motion detection

(Talcott et al., 2000). Consequently, dyslexia-related timing deficits have at different times been hypothesized as underlying dyslexics' visual and auditory perception problems, motor coordination problems and fluency and automatization problems (Nicolson et al., 2001), all of which have been proposed as adversely affecting the development of language and literacy skills (Overy et al., 2003). The full extent of the timing deficits is yet to be established, but suggests the need for further investigation.

A critical factor here seems to be the cerebellum, which is implicated in both timing functions (Ivry and Hazeltine, 1992; Ivry, 2000, p. 187; Ivry et al., 2002; Rubia, 2006) and in dyslexia (Reynolds et al., 2003; Kujala et al., 2007; Reynolds and Nicolson, 2007; Ben-Soussan et al., 2014a). Thus, cerebellar deficits are thought to affect articulation and working memory, due to deficits in timing which interfere with automatization of learning (Thomas and Karmiloff-Smith, 2002; Overy et al., 2003; Ram-Tsur et al., 2013).

In fact, cerebellar oscillatory function and its role in motor acquisition and timing have long been acknowledged (Andres et al., 1999; Swinnen, 2002; De Zeeuw et al., 2011). Given that impaired motor skills are often observed in dyslexics, some researchers have attributed dyslexics' cognitive and motor deficiencies to abnormal development and functioning of the cerebellum (Nicolson et al., 1995, 2001). These findings have led to the claim that the role of the cerebellum is not limited to regulating the timing, rate, force, rhythm and accuracy of movements, but also to the speed, capacity, consistency and appropriateness of cognitive processes (Schmahmann, 2004; Hölzel et al., 2011; Buckner, 2013).

Ben-Soussan et al. (2015) have recently presented a general model tying cerebellar function to cognitive improvement, by means of a particular form of motor training, which might be viewed as meditation-in-action, namely Quadrato Motor Training (QMT). QMT has been found to increase creativity, reflectivity and spatial cognition (Ben-Soussan et al., 2013, 2014b, 2015), as well as to increase neuronal synchronization and connectivity, especially within the alpha (8–12 Hz) range (Lasaponara et al., 2017). In addition, a month of daily QMT was found to improve reading and increase cerebellar alpha oscillations in dyslexic adults (Ben-Soussan et al., 2014a). Recently, QMT was further found to increase fractional anisotropy (FA) in tracts related to sensorimotor and cognitive functions and mindfulness, including the corticospinal tracts, anterior thalamic radiations and uncinate fasciculi, as well as in the left inferior fronto-occipital, superior and inferior longitudinal fasciculi (Piervincenzi et al., 2017), reflecting better white matter integrity as a result of greater intravoxel coherence of fiber orientation, axon density and diameter and/or myelination (Beaulieu et al., 1996; Sen and Bassar, 2005; Caminiti et al., 2013).

Let us contrast a hypothesized QMT-based improvement in the functioning of dyslexic adults to that found for another movement-based form of training, which has also reported an improvement in reading fluency for dyslexic participants. Reynolds et al. (2003) study was conducted on dyslexic pupils,

reporting improvements in balance, dexterity and reading. As they subsequently reported in a follow-up study (Reynolds and Nicolson, 2007), both dyslexic and non-dyslexic children benefited from the training, while alternative hypotheses raised by critics of the original study (e.g., based on potential artifacts, such as a Hawthorne effect) could be ruled out. In these studies, training was comprised of a home-based exercise program. In the study we report below, our QMT is also a home-based motor-exercise program. We do, however, also employ a home-based verbal-exercise program, as a suitable control.

Our present question concerns the hypothesized effect of QMT on time perception (specifically, time production, TP) in dyslexic adults. In a recent study (Ben-Soussan et al., 2017), normal reading participants reported a number of changes in their time perception during QMT, including: "an elongation of time, after a while, I had time to move from point to point, I didn't have to be in a hurry. I was faster and the exercise was slower"; "In the last day, when I finish, the precision—I can do many more things in a time, which I didn't think I can do." While QMT has thus been found to affect the subjective experience of time, the effects of QMT on TP have yet to be examined. Given that QMT is viewed as meditation-in-action, one can refer to the literature on meditation and time perception to develop a working hypothesis.

Meditation has been found to lead to a relative overestimation of target durations in passing (Glicksohn, 2001; Berkovich-Ohana et al., 2012; Kramer et al., 2013). Longer produced durations may be explained by a decrease in arousal (due to the decrease in the pacemaker speed of the internal clock), and an increase in size of the subjective time units (Glicksohn, 2001). QMT, viewed as meditation-in-action, should, like other forms of meditation, therefore lead to longer produced durations. We further consider gender, given that males usually make relatively longer TPs (Block et al., 2000, p. 1341; Zakay and Block, 1997, p. 13; Glicksohn and Hadad, 2012), and that male and female dyslexics may differ in the neurocognitive underpinnings of their dyslexia (Lambe, 1999). Hence, we expect to see a lengthening of TP (post—pre QMT), especially so for males. We further expect to see such effects for dyslexic individuals. If there is a gender-dependent change in TP in dyslexics, this would lend further support for searching for gender-dependent patterns of neural activity during this specific task of TP (which usually involves chronometric counting; Glicksohn and Hadad, 2012), as well as among other tasks involving auditory processing (Lambe, 1999, p. 532).

Given that our two reviewers expressed concern that because our participants were probably engaged in chronometric counting (as we, ourselves, have suggested), we might have compromised our study of TP, we shall take this opportunity to engage in debate about this issue. Some researchers argue that chronometric counting should be discouraged (e.g., Mimura et al., 2000; Kladopoulos et al., 2004); others argue that this should be encouraged (e.g., Miró et al., 2003; Myers and Tilley, 2003). Some researchers specifically request their participants to engage in counting so that the same strategy is employed by all participants (e.g., Perbal et al., 2003; Coelho et al., 2004). Counting is a natural strategy to employ in a task



of TP; and, as Fetterman and Killeen (1990, p. 766) argue, “The ubiquity of the practice calls into question experimental psychologists’ attempts to prevent or interfere with subjects’ counting strategies as a means of eliciting ‘uncontaminated’ temporal judgments...” Counting is undeniably a timing task (Brown et al., 2013); both timing without counting, and timing with counting, seem to be correlated to a fair degree (Bartholomew et al., 2015). Furthermore, techniques used to prevent counting may “...be distracting and introduce extraneous variables that can obscure effects specifically related to timing mechanisms” (Gaudreault and Fortin, 2013, p. 598). Consider, for example, the recent study reported by Schreuder et al. (2014). They employed three target durations of 1.33, 1.58 and 2.17 min, each to be produced in their TP task. As they write (p. 3), “we wanted to use intervals that exceeded 1 min, as these seem harder to produce because participants need to concentrate for a longer period of time.” To prevent counting they required their participants to remember, in parallel, an 8-character password (e.g., Z2Hx89bS). There are two problems, to our mind, with this procedure. First, these target durations are beyond an outer bound of 100 s for what would be considered to be time perception; as Wackermann (2007, p. 20) suggests, beyond this upper bound “time is merely cognitively (re)constructed, not actually experienced or “perceived,” a fact that is frequently ignored by contemporary time perception research.” Second, one does not usually try to retain in memory an 8-character password. Hence, what exactly is being investigated in this particular task of TP? In the present study, our participants were most probably employing chronometric counting as a natural strategy, hence were involved in timing *per se*, and not in adopting what might well be for them a suboptimal and unfamiliar strategy (not counting). Given our interest in the performance of our dyslexic participants, in particular, this seems to be ecologically wise.

## MATERIALS AND METHODS

### Participants and Design

For the current study, 29 adult Hebrew readers (19 women (F) and 10 men (M), mean age  $\pm$  SD:  $28 \pm 5$ ) were recruited, of whom 10 were dyslexic. The normal readers were randomly assigned to QMT ( $n = 9$ ; 7 F + 2 M) or *Verbal Training* (VT,  $n = 10$ ; 7 F + 3 M); the dyslexics were assigned to QMT (5 F + 5 M). The study was approved by the ethics committee of Bar-Ilan University. Upon entering the lab, all participants gave written informed consent. The study included three phases: pre-training assessment (Day 1), 28 days of daily training, and post-training assessment of TP. Pre and Post-training assessment took place at the lab. On the other training days, the participants performed the task at home. Compliance was controlled using a diary and daily recording of the training using a webcam. In addition, a semi-structured oral interview regarding QMT-induced experience was conducted, which included three open-ended questions regarding the participant’s physical, emotional and cognitive experiences during and following QMT (Ben-Soussan et al., 2017).

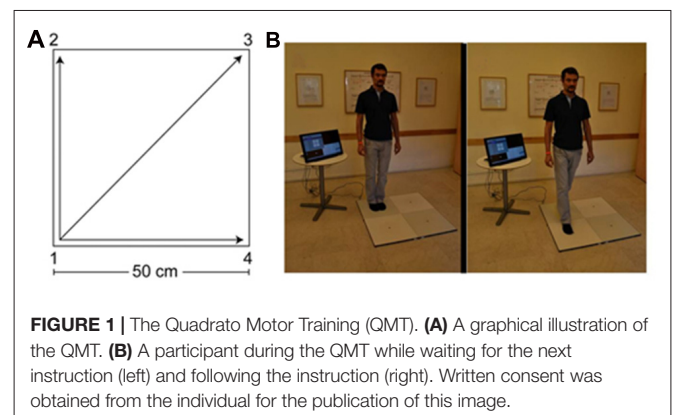
## The Training Groups

### Quadrato Motor Training (QMT)

The QMT group practiced the QMT in full. The QMT requires standing at one corner of a 0.5 m  $\times$  0.5 m square and making movements to different corners of the square in response to verbal instructions given by an audio tape recording indicating the next corner to which the participant should move. There are three optional directions of movement, and the movement is always in one step. We used a specific sequence of movements provided by Patrizio Paoletti, founder of the QMT program, translated from Italian to Hebrew by the first author. Each movement can be forward, backward, left, right, or diagonal. The instructions direct participants to keep the eyes focused straight ahead, hands loose at the side of the body. They are also told to immediately continue with the next instruction and not to stop due to mistakes. At each corner, there are three possible directions to move (for example, from corner 1 the participant can move to corner 2, to corner 3 or to corner 4). The training thus consists of 12 possible movements (3 directions  $\times$  4 corners): 2 forward, 2 backward, 2 left, 2 right and 4 diagonals. The participant is required to move from one corner to another according to the number on the recording. For example, if the sequence required is 1, 2, 1, 2, 1, 2, 3, 2, 4, 3, 1... this means moving to the first corner, then to the second, then back to the first, and so on.

The practice comprised 69 commands (23 sequences of movements that last  $\sim 88$  s; with a  $\sim 25$  s interval between each set of 23 commands for calibration). Thus, in total the whole QMT session lasted  $\sim 6$  min. Each movement has two instructions: the starting current position and the target position (“one four” means move from corner 1 to corner 4). Between noting the starting current position and the target position there was a randomized Inter-Stimulus Interval (ISI) of between 1,100 ms and 1,300 ms. ISI between trials (namely, between the previous target position and the next trial) was a randomized ISI of between 2,300 ms and 2,650 ms; see **Figure 1**.

In the current study we aimed at controlling limb velocity, by using a movement sequence comprising a total of 69 instruction steps, paced at a rate of 0.5 Hz (similar to a slow walking rate), which was the same for all participants. We also controlled for the decision regarding the responding limb by instructing



**FIGURE 1 |** The Quadrato Motor Training (QMT). **(A)** A graphical illustration of the QMT. **(B)** A participant during the QMT while waiting for the next instruction (left) and following the instruction (right). Written consent was obtained from the individual for the publication of this image.

participants to begin all movements with the leg closest to the center of the square.

### Verbal Training (VT)

The VT group stood 1 m in front of the square, but did not move on the corners of it. Instead, their instructions were to respond to the taped commands verbally by stating what direction of movement would be required in order to reach the corner specified by the command. For a movement from corner 1 to corner 2, they were required to say “straight;” for a movement from corner 1 to corner 3, they were required to say “diagonal.” The following is a list of all possible combinations and the appropriate response: 1–2, 4–3, “straight;” 2–1, 3–4, “back;” 1–3, 4–2, 3–1, 2–4, “diagonal;” 1–4, 2–3, “right;” 4–1, 3–2, “left.”

### Time Production (TP) Task

Four target durations of 4, 8, 16 and 32 s served for the TP task. The participant was required to remain with eyes closed while producing each of these target durations by pressing a finger button (Glicksohn, 1996) for the required period of time. Each target interval was produced twice, the target durations being presented in random order to the participant. Produced ( $P$ ) and target ( $T$ ) durations (in seconds) were log-transformed (to base 2), with required durations rendering thereby a linear scale ranging between 2 and 5, with a midpoint value of 3.5; produced duration was then regressed on required duration. We look at three dependent measures: (1) mean  $\log(P)$ ; (2) the *slope* of  $\log(P)$  regressed on  $\log(T)$ ; and (3) the *intercept* of that regression line (Glicksohn and Hadad, 2012).

## RESULTS

**Figures 2–4** presents individual log-log plots of produced duration as a function of target duration, blocked according to Group and Gender. For the controls assigned to VT (CV), one may note the essential linearity of the data in the log-log plot. We have fitted the linear regression lines for one CV male, and for two CV females, to exemplify this. The diagonal in each plot indicates what would be veridical TP (i.e., produced time = target duration). Note that for two of the CV male participants, produced duration post-training is longer than that of pre-training, and for the third male, the opposite is the case. Given this small group size, these opposite trends will easily cancel out, leaving no clear post-pre difference in TP, as we will subsequently show. For the CV females, one notes an increase in produced duration post-training for two participants, a decrease in produced duration post-training for three participants, and no noticeable change post vs. pre for the remaining two participants in this group. Again, this will result in a canceling out of effects, as with the male participants.

Turning to the participants assigned to QMT, we note that for the controls (CQ), one of the two males produces longer durations post-training (as can be seen on comparing the regression lines), while four of the females produce shorter durations post-training, and for the remaining three there is no

noticeable change post vs. pre. This is a clear effect for Gender, for these controls, as we will subsequently show. In contrast, when looking at the dyslectic participants, one notes that three of the five males produce somewhat longer durations post-training, and three of the five females produce markedly longer durations post-training. We turn now to a formal analysis of these trends, using analysis of variance (ANOVA).

For each of our three dependent measures (mean, slope, intercept), we ran a Group (dyslectics assigned to QMT, controls assigned to QMT, controls assigned to VT)  $\times$  Gender (male, female)  $\times$  Time (pre, post) ANOVA, adopting the Greenhouse-Geisser  $p$ -value for each effect. **Figure 5** presents mean ( $\pm$ SE) values for mean  $\log(P)$ . The three-way interaction for this measure was significant ( $F_{(2,23)} = 3.85$ ,  $MSE = 0.068$ ,  $p < 0.05$ ). There was no main effect for Gender ( $F_{(1,23)} = 1.71$ , ns), Group ( $F_{(2,23)} < 1$ ) or Time ( $F_{(1,23)} = 2.95$ ,  $p = 0.10$ ), and no two-way interactions.

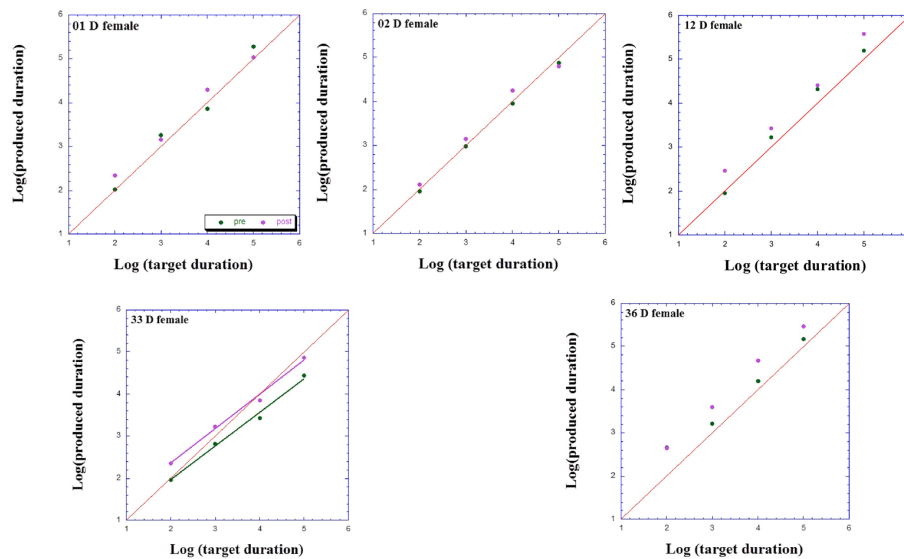
We found a significant lengthening of produced time for the female dyslectics following QMT ( $t_{(4)} = -3.80$ ,  $p < 0.05$ ;  $n = 5$ ), in contrast there was a decrease in produced time for the control females following QMT ( $t_{(6)} = 2.56$ ,  $p < 0.05$ ;  $n = 7$ ). No such difference was found in the VT group, either for females ( $t_{(6)} = 0.28$ , ns;  $n = 7$ ), or for males ( $t_{(2)} = 0.05$ , ns;  $n = 3$ ).

We ran a comparable ANOVA, this time with  $\log(P)$  comprising a profile of four mean values for each of the four target durations (Target Duration). In this analysis, we uncovered a Target Duration  $\times$  Gender interaction ( $F_{(3,63)} = 3.86$ ,  $MSE = 0.036$ ,  $p < 0.05$ ), as well as the expected main effect for Target Duration ( $F_{(3,63)} = 1,481.26$ ,  $MSE = 0.036$ ,  $p < 0.001$ ). In short, increasing target duration results in increasing produced duration, for both male and female participants, as one would expect; furthermore, for the longer target durations, male participants produce longer durations than do female participants, while for the specific target duration of 2 s, females produce longer target durations than males. Given that these effects are not dependent on Group, our focus on mean  $\log(P)$  and on the three-way Group  $\times$  Gender  $\times$  Time interaction for this is supported.

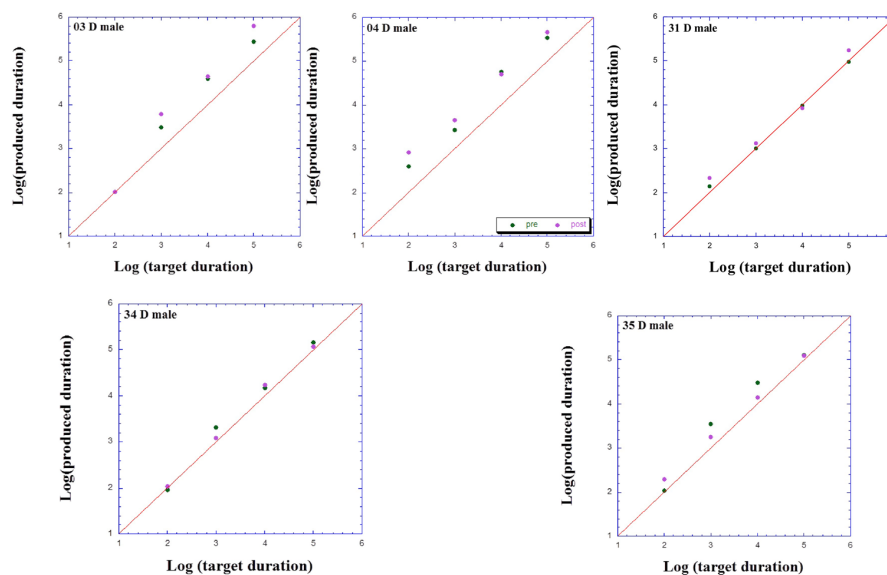
Turning to the other two measures, we found no three-way interaction for the slope ( $F_{(2,23)} < 1$ ) nor for the intercept ( $F_{(2,23)} < 1$ ). Mean slope values range between 0.92 and 1.06, hence do not deviate from an expected slope of 1.00 (Glicksohn and Hadad, 2012), while mean intercept values range between  $-0.11$  and  $0.59$ . For the slope, there was no main effect for Gender ( $F_{(1,23)} = 3.17$ ,  $p = 0.09$ ) or Group ( $F_{(2,23)} < 1$ ), and no two-way interaction. For the intercept, there was also no main effect for Gender ( $F_{(1,23)} < 1$ ), but there was a main effect for Group ( $F_{(2,23)} = 3.58$ ,  $MSE = 0.185$ ,  $p < 0.05$ ), whereby the controls assigned to VT had a higher mean intercept (0.461) than the dyslectics assigned to QMT (0.321) and the controls assigned to QMT (0.016).

There were no baseline differences in TP between dyslexic and normal readers for any of the three measures. Specifically, for mean  $\log(P)$ , slope and intercept, there was no main effect for Group ( $F_{(2,23)} = 1.56, 0.11, 1.73$ , respectively, all ns), for Gender ( $F_{(1,23)} = 0.55, 2.18, 0.53$ , respectively, all ns) nor their interaction ( $F_{(2,23)} < 1$  for each).

## D: Females



## D: Males



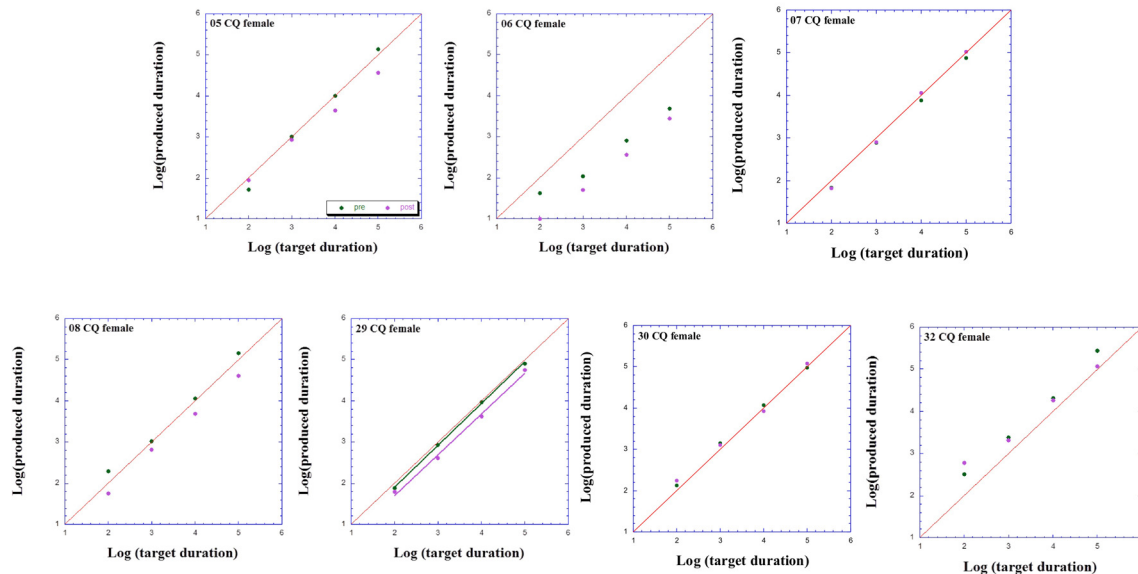
**FIGURE 2 |** Individual log-log plots of produced duration of the dyslexics assigned to QMT group (D), as a function of target duration, blocked according to Gender.

## First-Person Reports

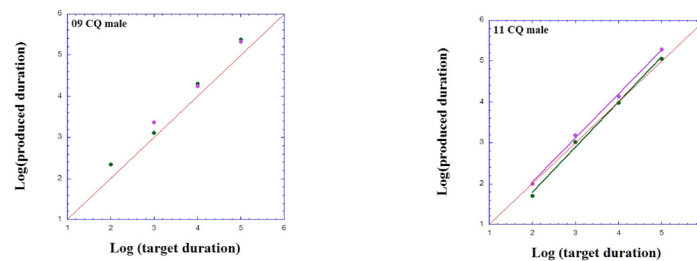
The semi-structured interview revealed that eight participants from the QMT group reported having increased attention and relaxation. More specifically, participant 03, a dyslexic male, reported: “I had to focus that the inner leg will be first but with practice it became less hard. In the beginning, I really had to concentrate for that. Maybe more balance and equilibrium.” Participant 34, a dyslexic male, reported having “more sharpness. Things are retrieved faster. More focused. Yes, it contributes to focus, you have focus.” Participant 33, a dyslexic female, reported having more “Attention, and listening more what people say. Not

just hearing the voice but the listening.” In addition, two dyslexic and two normal readers reported a sense of relaxation and calmness following the training. Participant 1, a dyslexic female, reported: “I am calmer. I don’t know if it is because of it, but in some place my stress decreased from all things and their meaning, let’s say if I don’t find an apartment, I will go abroad, it’s not critical. Acceptance.” Participant 31, a dyslexic male, reported “relaxation as a result of the training. As a result of the relaxation, looking at the decisions in a more reasonable or concentrated way. The practice felt a bit meditative”. Participants 08 and 06, normal female readers, reported “feeling calmer;” and

## CQ: Females



## CQ: Males



**FIGURE 3 |** Individual log-log plots of produced duration of the controls assigned to QMT (CQ) as a function of target duration, blocked according to Gender.

“feelings of relaxation and calmness. It has a bit of an effect like meditation. It enters into a state of mind that I should do the experiment. And also when my thoughts wandered, you need to be focused and to the thing to keep me in the frame,” respectively.

The only participant who reported being more aroused following the training was actually from the verbal control group: “I felt two things. One, is that when I am tired, I am more concentrated during the training and there were times I really awoke after.”

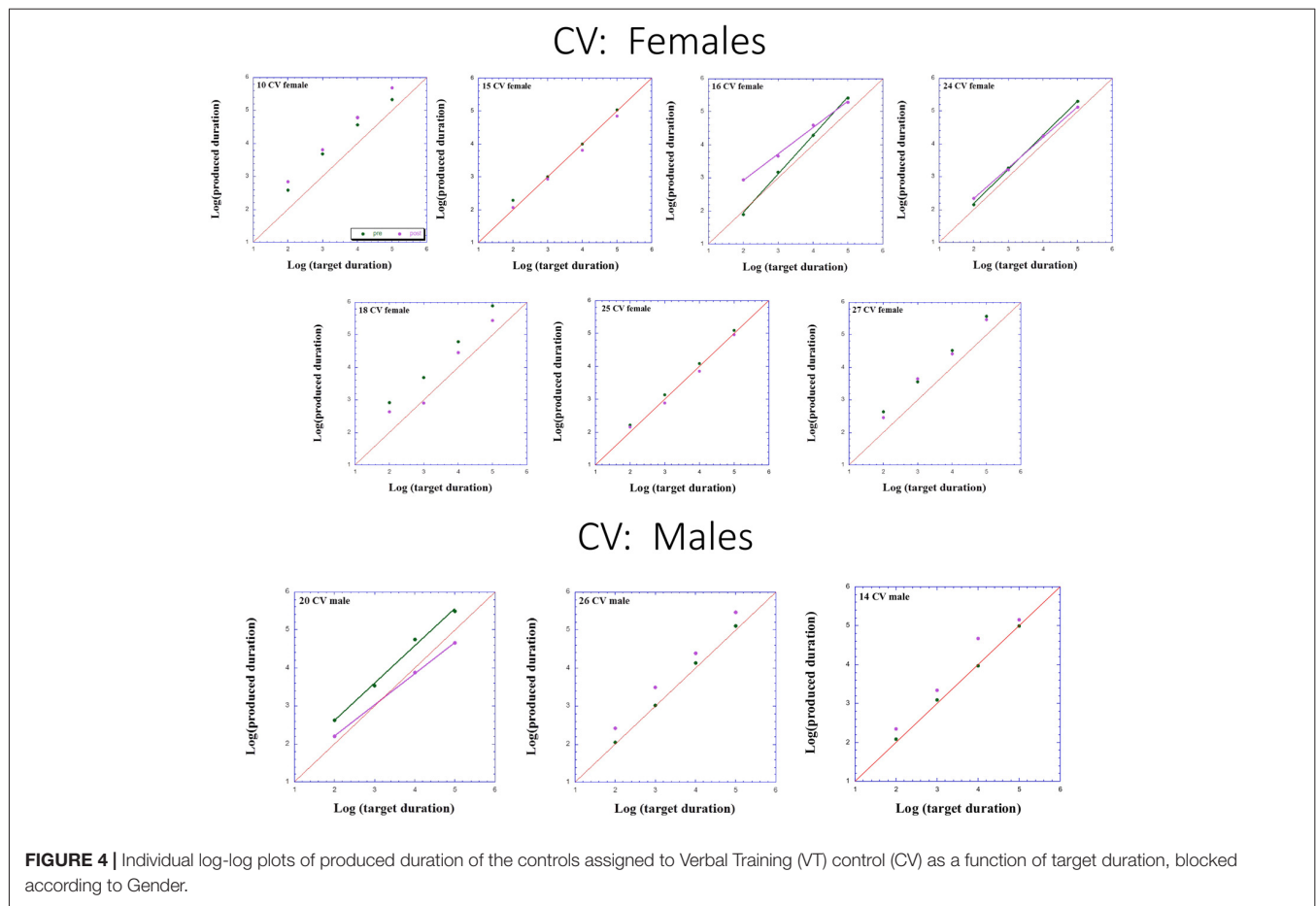
## DISCUSSION

Time estimation is an important component of the ability to organize and plan sequences of actions as well as cognitive functions. While attention deficits are accompanied by short TPs, expert meditators have been reported to produce longer durations, related to their increased attentional resources. In

the current study, we examined the effects of a month of QMT, a structured sensorimotor training program that involves sequencing of motor responses based on verbal commands, on TP using a pre-post design. Our results demonstrate that in contrast to the controls, longer TP was found following 1 month of intensive QMT in the female dyslexic group, while shorter TP was found for the control females following QMT. We suggest that this may be related to three mediating inter-related mechanisms, including enhanced attention resulting from QMT, better working memory and better cerebellar functioning. The semi-structured interview confirmed this hypothesis and revealed that participants from the QMT group reported having increased attention and relaxation. This is in line with our previous report (Ben-Soussan et al., 2017).

The involvement of the cerebellum in cognition has been overshadowed by years of focus on its motor role. Yet, the cerebellum, possibly but not exclusively through its connections with frontal and prefrontal areas, contributes to cognition,





learning and language (Beaton and Mariën, 2010; Pesce and Ben-Soussan, 2016), leading also to the notion of the linguistic cerebellum (Jansen et al., 2005; Stoodley and Schmahmann, 2009). More specifically, it has been suggested that cerebellar dysfunction may be involved in dyslexia due to the cerebellum's role as an oscillator, producing synchronized activity within neuronal networks, including sensorimotor networks critical for reading, timing and attention (Buhusi and Meck, 2005; Ben-Soussan et al., 2014a).

Within an internal-clock framework, a change in attentional resources can result in longer perceived duration (Kramer et al., 2013). Such a practice-enhanced attention results in better working memory (Davis and Hayes, 2011), which is a main deficit in dyslexics (Jeffries and Everatt, 2004; Smith-Spark and Fisk, 2007). In turn, working memory is closely related to the cerebellum (Justus and Ivry, 2001; Ravizza et al., 2006). In fact, it has been suggested that the cerebellar impairments in dyslexia, which are linked to reduced articulation speed, may lead to impaired working memory, and in turn to the language impairments (Nicolson et al., 2001).

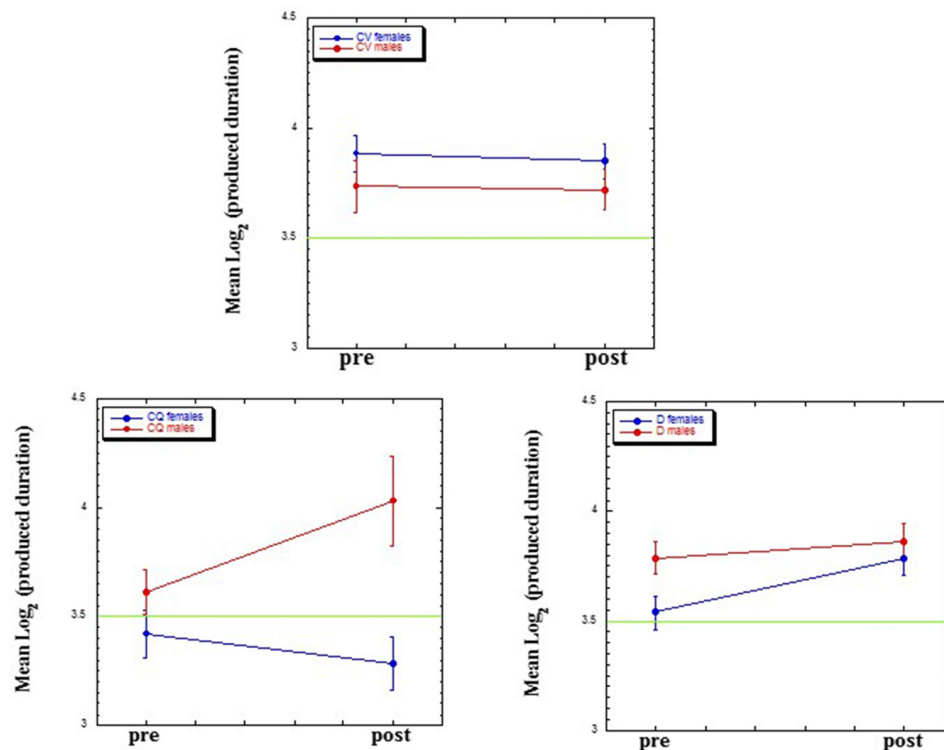
In addition to the centrality of phonological mechanisms in dyslexia, recent evidence also supports an important role for attentional mechanisms (Shaywitz and Shaywitz, 2008; Shaywitz et al., 2017). The lengthening of TP following QMT in the female dyslexic group could be related to increased attention and

activation of the cerebellum. In fact, QMT was previously found to increase mindfulness and attentional effort (Ben-Soussan et al., 2017), as well as to improve white matter integrity of neuronal pathways related to attention and learning (Piervincenzi et al., 2017) in normal readers. Furthermore, given that participants predominantly employ chronometric counting when engaging with our TP task (Glicksohn and Hadad, 2012), the (right) cerebellum was surely activated (Tracy et al., 2000; O'Leary et al., 2003; Hinton et al., 2004). QMT, similar to other mindfulness training, involves deliberately staying in the present moment (Kramer et al., 2013; Ben-Soussan et al., 2014b). We note that mindfulness meditation trains attentional skills and produces increased attentional resources (Lutz et al., 2008).

The shortening of TP for the control females following QMT could be related to their induced arousal as a consequence of the motor training, speeding up the internal clock rate, hence leading to a shortening of TP (Ozel et al., 2004). In fact, a similar trend has been previously observed on using MEG, wherein both dyslexic and control groups improved reading performance; cerebellar alpha oscillations increased in the dyslexic group, while the opposite trend occurred in the normal reader group (Ben-Soussan et al., 2014a).

A common element of many theories related to the cause of dyslexia is the conviction that timing skills, and particularly rapid timing skills and motor timing skills, are a fundamental problem

# Time Production Data



**FIGURE 5 |** Mean ( $\pm$ SE) values for mean  $\log_2(P)$  as a function of Group (dyslectics assigned to QMT, D; controls assigned to QMT, CQ and controls assigned to VT, CV), Gender (male, female) and Time (pre, post).

area. In fact, it has been previously suggested that a deficit in rapid temporal processing can cause specific auditory perception problems, leading to specific phonological perception problems (Tallal et al., 1993). Yet, in the current study we did not find baseline differences in TP between dyslexic and normal readers in any of the three measures.

We expected to see a lengthening of TP (post—pre QMT), especially so for males. In contrast to our hypothesis, we observed a lengthening of TP, for the female participants. A trend of a lengthening of TP occurred for both QMT groups, in contrast to the VT control group, yet this lengthening of TP was not statistically significant for the males, probably due to the small number of these participants. We have detected an interaction involving gender, whereby the hypothesized lengthening of TP for both QMT groups is found only for males, while for females this lengthening is found only for the dyslectics, in contrast to a shortening of TP observed for the controls, and this is intriguing. Ingallhalikar et al. (2014) have recently shown that “In all supratentorial regions, males had greater within-hemispheric connectivity, as well as enhanced modularity and transitivity, whereas between-hemispheric connectivity and cross-module participation predominated in females. However, this effect was reversed in the cerebellar connections.” Can our results be related

to these gender differences in cerebellar connectivity? There is much to explore here in future studies.

The current study is a preliminary attempt to examine the connection between sensorimotor training, TP and dyslexia. The main limitations of the current study are the small sample size and the use of only one training paradigm. In the future, a study on a larger sample that includes dyslexic no-training and verbal control groups may extend the current results.

## CONCLUSION/SIGNIFICANCE

The current findings suggest that the combination of motor and mindful training, embedded in QMT, has a differential effect depending on one's gender and whether one is dyslexic or not. This may have valuable implications for educational and contemplative neuroscience, in emphasizing the connection between specifically-structured motor training, time estimation and attention.

## AUTHOR CONTRIBUTIONS

TDB-S and JG conceived and designed the experiments, analyzed the data and wrote the article.

## REFERENCES

- Andres, F. G., Mima, T., Schulman, A. E., Dichgans, J., Hallett, M., and Gerloff, C. (1999). Functional coupling of human cortical sensorimotor areas during bimanual skill acquisition. *Brain* 122, 855–870. doi: 10.1093/brain/122.5.855
- Bartholomew, A. J., Meck, W. H., and Cirulli, E. T. (2015). Analysis of genetic and non-genetic factors influencing timing and time perception. *PLoS One* 10:e0143873. doi: 10.1371/journal.pone.0143873
- Beaton, A., and Mariën, P. (2010). Language, cognition and the cerebellum: grappling with an enigma. *Cortex* 46, 811–820. doi: 10.1016/j.cortex.2010.02.005
- Beaulieu, C., Does, M. D., Snyder, R. E., and Allen, P. S. (1996). Changes in water diffusion due to Wallerian degeneration in peripheral nerve. *Magn. Reson. Med.* 36, 627–631. doi: 10.1002/mrm.1910360419
- Ben-Soussan, T. D., Avirame, K., Glicksohn, J., Goldstein, A., Harpaz, Y., and Ben-Schachar, M. (2014a). Changes in cerebellar activity and inter-hemispheric coherence accompany improved reading capacity following Quadrato Motor Training. *Front. Syst. Neurosci.* 8:81. doi: 10.3389/fnsys.2014.00081
- Ben-Soussan, T. D., Berkovich-Ohana, A., Glicksohn, J., and Goldstein, A. (2014b). A suspended act: increased reflectivity and gender-dependent electrophysiological change following Quadrato Motor Training. *Front. Psychol.* 5:55. doi: 10.3389/fpsyg.2014.00055
- Ben-Soussan, T. D., Glicksohn, J., and Berkovich-Ohana, A. (2015). From cerebellar activation and connectivity to cognition: a review of the Quadrato Motor Training. *Biomed Res. Int.* 2015:954901. doi: 10.1155/2015/954901
- Ben-Soussan, T. D., Glicksohn, J., and Berkovich-Ohana, A. (2017). Attentional effort, mindfulness, and altered states of consciousness experiences following Quadrato Motor Training. *Mindfulness* 8, 59–67. doi: 10.1007/s12671-015-0469-5
- Ben-Soussan, T. D., Glicksohn, J., Goldstein, A., Berkovich-Ohana, A., and Donchin, O. (2013). Into the square and out of the box: the effects of Quadrato Motor Training on creativity and  $\alpha$  coherence. *PLoS One* 8:e55023. doi: 10.1371/journal.pone.0055023
- Berkovich-Ohana, A., Glicksohn, J., and Goldstein, A. (2012). Mindfulness-induced changes in gamma band activity-implications for the default mode network, self-reference and attention. *Clin. Neurophysiol.* 123, 700–710. doi: 10.1016/j.clinph.2011.07.048
- Block, R. A., Hancock, P. A., and Zakay, D. (2000). Sex differences in duration judgments: a meta-analytic review. *Mem. Cognit.* 28, 1333–1346. doi: 10.3758/bf03211834
- 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
- Buckner, R. L. (2013). The cerebellum and cognitive function: 25 years of insight from anatomy and neuroimaging. *Neuron* 80, 807–815. doi: 10.1016/j.neuron.2013.10.044
- 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
- Caminiti, R., Carducci, F., Piervincenzi, C., Battaglia-Mayer, A., Confalone, G., Visco-Comandini, F., et al. (2013). Diameter, length, speed and conduction delay of callosal axons in macaque monkeys and humans: comparing data from histology and magnetic resonance imaging diffusion tractography. *J. Neurosci.* 33, 14501–14511. doi: 10.1523/jneurosci.0761-13.2013
- Coelho, M., Ferreira, J. J., Dias, B., Sampaio, C., Martins, I. P., and Castro-Caldas, A. (2004). Assessment of time perception: the effect of aging. *J. Int. Neuropsychol. Soc.* 10, 332–341. doi: 10.1017/S1355617704103019
- Davis, D. M., and Hayes, J. A. (2011). What are the benefits of mindfulness? A practice review of psychotherapy-related research. *Psychotherapy* 48, 198–208. doi: 10.1037/a0022062
- De Zeeuw, C. I., Hoebeek, F. E., Bosman, L. W., Schonewille, M., Witter, L., and Koekkoek, S. K. (2011). Spatiotemporal firing patterns in the cerebellum. *Nat. Rev. Neurosci.* 12, 327–344. doi: 10.1038/nrn3011
- Fetterman, J. G., and Killeen, P. R. (1990). A componential analysis of pacemaker-counter timing systems. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 766–780. doi: 10.1037//0096-1523.16.4.766
- Gaudreault, R., and Fortin, C. (2013). To count or not to count: the effect of instructions on expecting a break in timing. *Atten. Percept. Psychophys.* 75, 588–602. doi: 10.3758/s13414-012-0411-z
- Glicksohn, J. (1996). Entering trait and context into a cognitive-timer model for time estimation. *J. Environ. Psychol.* 16, 361–370. doi: 10.1006/jevp.1996.0030
- Glicksohn, J. (2001). Temporal cognition and the phenomenology of time: a multiplicative function for apparent duration. *Conscious. Cogn.* 10, 1–25. doi: 10.1006/ccog.2000.0468
- Glicksohn, J., and Hadad, Y. (2012). Sex differences in time production revisited. *J. Individ. Dif.* 33, 35–42. doi: 10.1027/1614-0001/a000059
- 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
- Hölzel, B. K., Carmody, J., Vangel, M., Congleton, C., Yerramsetti, S. M., Gard, T., et al. (2011). Mindfulness practice leads to increases in regional brain gray matter density. *Psychiatry Res.* 191, 36–43. doi: 10.1016/j.psychresns.2010.08.006
- Ingalhalikar, M., Smith, A., Parker, D., Satterthwaite, T. D., Elliott, M. A., Ruparel, K., et al. (2014). Sex differences in the structural connectome of the human brain. *Proc. Natl. Acad. Sci. U S A* 111, 823–828. doi: 10.1073/pnas.1316909110
- Ivry, R. (2000). Exploring the role of the cerebellum in sensory anticipation and timing: commentary on Tesche and Karhu. *Hum. Brain Mapp.* 9, 115–118. doi: 10.1002/(sici)1097-0193(200003)9:3<115::aid-hbm1>3.3.co;2-x
- Ivry, R. B., and Hazeltine, R. E. (1992). “Models of timing-with-a-timer,” in *Time, Action and Cognition: Towards Bridging the Gap*, eds F. Macar, V. Pouthas and W. J. Friedman (Dordrecht: Kluwer Academic Publishers), 183–189.
- Ivry, R. B., Spencer, R. M., Zelaznik, H. N., and Diedrichsen, J. (2002). “The cerebellum and event timing,” in *The Cerebellum: Recent Developments in Cerebellar Research. Annals of the New York Academy of Science*, (Vol. 978) eds S. M. Highstein and W. T. Thach (New York, NY: The New York Academy of Sciences), 302–317.
- Jansen, A., Flöel, A., Van Randenborgh, J., Konrad, C., Rotte, M., Förster, A. F., et al. (2005). Crossed cerebro-cerebellar language dominance. *Hum. Brain Mapp.* 24, 165–172. doi: 10.1002/hbm.20077
- Jeffries, S., and Everatt, J. (2004). Working memory: its role in dyslexia and other specific learning difficulties. *Dyslexia* 10, 196–214. doi: 10.1002/dys.278
- Justus, T. C., and Ivry, R. B. (2001). The cognitive neuropsychology of the cerebellum. *Int. Rev. Psychiatry* 13, 276–282. doi: 10.1080/09540260127528
- Kladopoulos, C. N., Hemmes, N. S., and Brown, B. L. (2004). Prospective timing under dual-task paradigms: attentional and contextual-change mechanisms. *Behav. Processes* 67, 221–233. doi: 10.1016/j.beproc.2003.12.004
- Kramer, R. S., Weger, U. W., and Sharma, D. (2013). The effect of mindfulness meditation on time perception. *Conscious. Cogn.* 22, 846–852. doi: 10.1016/j.concog.2013.05.008
- Kujala, J., Pammer, K., Cornelissen, P., Roebroek, A., Formisano, E., and Salmelin, R. (2007). Phase coupling in a cerebro-cerebellar network at 8–13 Hz during reading. *Cereb. Cortex* 17, 1476–1485. doi: 10.1093/cercor/bhl059
- Lambe, E. K. (1999). Dyslexia, gender, and brain imaging. *Neuropsychologia* 37, 521–536. doi: 10.1016/s0028-3932(98)00146-8
- Lasaponara, S., Mauro, F., Carducci, F., Paoletti, P., Tombini, M., Quattrocchi, C. C., et al. (2017). Increased  $\alpha$  band functional connectivity following the Quadrato Motor Training: a longitudinal study. *Front. Hum. Neurosci.* 11:282. doi: 10.3389/fnhum.2017.00282
- Lutz, A., Slagter, H. A., Dunne, J. D., and Davidson, R. J. (2008). Attention regulation and monitoring in meditation. *Trends Cogn. Sci.* 12, 163–169. doi: 10.1016/j.tics.2008.01.005
- Mimura, M., Kinsbourne, M., and O'Connor, M. (2000). Time estimation by patients with frontal lesions and by Korsakoff amnesics. *J. Int. Neuropsychol. Soc.* 6, 517–528. doi: 10.1017/s1355617700655017
- Miró, E., Cano, M. C., Espinosa-Fernández, L., and Buela-Casal, G. (2003). Time estimation during prolonged sleep deprivation and its relation to activation measures. *Hum. Factors* 45, 148–159. doi: 10.1518/hfes.45.1.148.27227
- Myers, P. M., and Tilley, A. (2003). The relationship between diurnal type and time duration estimation at morning and evening times of day. *Pers. Individ. Dif.* 35, 1141–1150. doi: 10.1016/s0191-8869(02)00324-0

- Nicolson, R. I., Fawcett, A. J., and Dean, P. (1995). Time estimation deficits in developmental dyslexia: evidence of cerebellar involvement. *Proc. Biol. Sci.* 259, 43–47. doi: 10.1098/rspb.1995.0007
- Nicolson, R. I., Fawcett, A. J., and Dean, P. (2001). Developmental dyslexia: the cerebellar deficit hypothesis. *Trends Neurosci.* 24, 508–511. doi: 10.1016/s0166-2236(00)01896-8
- O'Leary, D. S., Block, R. I., Turner, B. M., Koeppel, J., Magnotta, V. A., Ponto, L. B., et al. (2003). Marijuana alters the human cerebellar clock. *Neuroreport* 14, 1145–1151. doi: 10.1097/00001756-200306110-00009
- Overly, K., Nicolson, R. I., Fawcett, A. J., and Clarke, E. F. (2003). Dyslexia and music: measuring musical timing skills. *Dyslexia* 9, 18–36. doi: 10.1002/dys.233
- Ozel, S., Larue, J., and Dosseville, F. (2004). Effect of arousal on internal clock speed in real action and mental imagery. *Can. J. Exp. Psychol.* 58, 196–205. doi: 10.1037/h0087444
- Perbal, S., Couillet, J., Azouvi, P., and Pouthas, V. (2003). Relationships between time estimation, memory, attention and processing speed in patients with severe traumatic brain injury. *Neuropsychologia* 41, 1599–1610. doi: 10.1016/s0028-3932(03)00110-6
- Pesce, C., and Ben-Soussan, T. D. (2016). “Cogito ergo sum” or “ambulo ergo sum”? new perspectives in developmental exercise and cognition research,” in *Exercise-Cognition Interaction: Neuroscience Perspectives*, ed. T. McMorris (London: Elsevier), 251–282.
- Piervincenzi, C., Ben-Soussan, T. D., Mauro, F., Mallio, C. A., Errante, Y., Quattrocchi, C. C., et al. (2017). White matter microstructural changes following Quadrato Motor Training: a longitudinal study. *Front. Hum. Neurosci.* 11:590. doi: 10.3389/fnhum.2017.00590
- Ram-Tsur, R., Nissim, M., Zion, M., Ben-Soussan, T. D., and Mevarech, Z. (2013). Language development: the effect of aquatic and on-land motor interventions. *Creat. Educ.* 4, 41–50. doi: 10.4236/ce.2013.49b009
- Ramus, F., Pidgeon, E., and Frith, U. (2003). The relationship between motor control and phonology in dyslexic children. *J. Child Psychol. Psychiatry* 44, 712–722. doi: 10.1111/1469-7610.00157
- Ravizza, S. M., McCormick, C. A., Schlerf, J. E., Justus, T., Ivry, R. B., and Fiez, J. A. (2006). Cerebellar damage produces selective deficits in verbal working memory. *Brain* 129, 306–320. doi: 10.1093/brain/awh685
- Reynolds, D., and Nicolson, R. I. (2007). Follow-up of an exercise-based treatment for children with reading difficulties. *Dyslexia* 13, 78–96. doi: 10.1002/dys.331
- Reynolds, D., Nicolson, R. I., and Hambly, H. (2003). Evaluation of an exercise-based treatment for children with reading difficulties. *Dyslexia* 9, 48–71. doi: 10.1002/dys.235
- Rubia, K. (2006). “The neural correlates of timing functions,” in *Timing the Future: The Case for a Time-Based Prospective Memory*, eds J. Glicksohn and M. S. Myslobodsky (London: World Scientific), 213–238.
- Schmahmann, J. D. (2004). Disorders of the cerebellum: ataxia, dysmetria of thought and the cerebellar cognitive affective syndrome. *J. Neuropsychiatry Clin. Neurosci.* 16, 367–378. doi: 10.1176/jnp.16.3.367
- Schreuder, E., Hoeksma, M. R., Smeets, M. A., and Semin, G. R. (2014). The effects of odor and body posture on perceived duration. *Front. Neurobot.* 8:6. doi: 10.3389/fnbot.2014.00006
- Sen, P. N., and Basser, P. J. (2005). A model for diffusion in white matter in the brain. *Biophys. J.* 89, 2927–2938. doi: 10.1529/biophysj.105.063016
- Shaywitz, S. E., and Shaywitz, B. A. (2008). Paying attention to reading: the neurobiology of reading and dyslexia. *Dev. Psychopathol.* 20, 1329–1349. doi: 10.1017/s0954579408000631
- Shaywitz, S., Shaywitz, B., Wietecha, L., Wigal, S., McBurnett, K., Williams, D., et al. (2017). Effect of atomoxetine treatment on reading and phonological skills in children with dyslexia or attention-deficit/hyperactivity disorder and comorbid dyslexia in a randomized, placebo-controlled trial. *J. Child Adolesc. Psychopharmacol.* 27, 19–28. doi: 10.1089/cap.2015.0189
- Smith-Spark, J. H., and Fisk, J. E. (2007). Working memory functioning in developmental dyslexia. *Memory* 15, 34–56. doi: 10.1080/09658210601043384
- Stoodley, C. J., and Schmahmann, J. D. (2009). The cerebellum and language: evidence from patients with cerebellar degeneration. *Brain Lang.* 110, 149–153. doi: 10.1016/j.bandl.2009.07.006
- Swinnen, S. P. (2002). Intermanual coordination: from behavioural principles to neural-network interactions. *Nat. Rev. Neurosci.* 3, 348–359. doi: 10.1038/nrn807
- Talcott, J. B., Hansen, P. C., Assoku, E. L., and Stein, J. F. (2000). Visual motion sensitivity in dyslexia: evidence for temporal and energy integration deficits. *Neuropsychologia* 38, 935–943. doi: 10.1016/s0028-3932(00)00020-8
- Tallal, P., Miller, S., and Fitch, R. H. (1993). Neurobiological basis of speech: a case for the preeminence of temporal processing. *Ann. N Y Acad. Sci.* 682, 27–47. doi: 10.1111/j.1749-6632.1993.tb22957.x
- Thomas, M., and Karmiloff-Smith, A. (2002). Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *Behav. Brain Sci.* 25, 727–750. doi: 10.1017/s0140525x02000134
- Tracy, J. I., Faro, S. H., Mohamed, F. B., Pinsk, M., and Pinus, A. (2000). Functional localization of a “Time Keeper” function separate from attentional resources and task strategy. *Neuroimage* 11, 228–242. doi: 10.1006/nimg.2000.0535
- Wackermann, J. (2007). Inner and outer horizons of time experience. *Span. J. Psychol.* 10, 20–32. doi: 10.1017/s1138741600006284
- Witton, C., Talcott, J. B., Hansen, P. C., Richardson, A. J., Griffiths, T. D., Rees, A., et al. (1998). Sensitivity to dynamic auditory and visual stimuli predicts nonword reading ability in both dyslexic and normal readers. *Curr. Biol.* 8, 791–797. doi: 10.1016/s0960-9822(98)70320-3
- Wolff, P. (2002). Timing precision and rhythm in developmental dyslexia. *Read. Writ.* 15, 179–206. doi: 10.1023/A:1013880723925
- Wolff, P., Michel, G. F., Ovrut, M., and Drake, C. (1990). Rate and timing precision of motor coordination in developmental dyslexia. *Dev. Psychol.* 26, 349–359. doi: 10.1037//0012-1649.26.3.349
- Zakay, D., and Block, R. A. (1997). Temporal cognition. *Curr. Dir. Psychol. Sci.* 6, 12–16. doi: 10.1111/1467-8721.ep11512604

**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 © 2018 Ben-Soussan and Glicksohn. 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) and the copyright owner(s) 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.





# Metastable States of Multiscale Brain Networks Are Keys to Crack the Timing Problem

Tommaso Gili<sup>1,2\*</sup>, Valentina Ciullo<sup>2,3</sup> and Gianfranco Spalletta<sup>2,4\*</sup>

<sup>1</sup> IMT School for Advanced Studies Lucca, Lucca, Italy, <sup>2</sup> Laboratory of Neuropsychiatry, IRCCS Santa Lucia Foundation, Rome, Italy, <sup>3</sup> Department of Neurosciences, Psychology, Drug Research and Child Health, University of Florence, Florence, Italy, <sup>4</sup> Division of Neuropsychiatry, Menninger Department of Psychiatry and Behavioral Sciences, Baylor College of Medicine, Houston, TX, United States

The dynamics of the environment where we live in and the interaction with it, predicting events, provided strong evolutionary pressures for the brain functioning to process temporal information and generate timed responses. As a result, the human brain is able to process temporal information and generate temporal patterns. Despite the clear importance of temporal processing to cognition, learning, communication and sensory, motor and emotional processing, the basal mechanisms of how animals differentiate simple intervals or provide timed responses are still under debate. The lesson we learned from the last decade of research in neuroscience is that functional and structural brain connectivity matter. Specifically, it has been accepted that the organization of the brain in interacting segregated networks enables its function. In this paper we delineate the route to a promising approach for investigating timing mechanisms. We illustrate how novel insight into timing mechanisms can come by investigating brain functioning as a multi-layer dynamical network whose clustered dynamics is bound to report the presence of metastable states. We anticipate that metastable dynamics underlie the real-time coordination necessary for the brain's dynamic functioning associated to time perception. This new point of view will help further clarifying mechanisms of neuropsychiatric disorders.

**Keywords:** brain networks, multiscale modeling, metastable state brain dynamics, timing and time perception, functional MRI, electrophysiology

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

Andrew A. Fingelkurts,  
BM-Science, Finland  
Martin Kröger,  
ETH Zürich, Switzerland

### \*Correspondence:

Tommaso Gili  
tommaso.gili@imtlucca.it  
Gianfranco Spalletta  
g.spalletta@hsantalucia.it

**Received:** 27 April 2018

**Accepted:** 17 August 2018

**Published:** 11 September 2018

### Citation:

Gili T, Ciullo V and Spalletta G (2018)  
Metastable States of Multiscale Brain  
Networks Are Keys to Crack the  
Timing Problem.  
*Front. Comput. Neurosci.* 12:75.  
doi: 10.3389/fncom.2018.00075

## THE VIEW

Timing is an umbrella term that encompasses a variety of processes based on the prediction and estimation of temporal intervals across a wide range of scales, from hundreds of milliseconds to seconds. Theoretical models, mainly based on the existence of an internal clock (Gibbon, 1977), have been challenged by compelling behavioral findings that enhance suspects about its biological plausibility (Karmarkar and Buonomano, 2007). Alternate models have been proposed, describing timing as an ensemble of neural processes emerging from the activity of neural circuits inherently capable of temporal processing as a result of the complexity of cortical networks coupled with the presence of time-dependent neuronal properties (Buonomano and Maass, 2009). In this view, neural systems can benefit from the temporal evolution of their states, caused by the variation in neural and synaptic properties. The overall effect results in an adaptation of cerebral networks that

could be tuned to discriminate temporal intervals (Bueno et al., 2017). State-dependent models can be extended to be consistent with the majority of timing models (Hass and Durstewitz, 2016), with the different models indicating specific constraints on what would collapse the state space. Although a route is traced toward a comprehensive description of timing, it is still unclear whether brain networks states are part of a coding scheme used to track time or a by-product of other processes that could generate a time-decodable signal. A possible theoretical framework could be the multi-scale description of brain networks both in space and time. On one hand it would be able to capture the local-to-global properties of neural processes that give rise to timing, on the other hand it would allow to grasp the integration processes among brain regions responsible for timing by means of metastability of network states (Friston, 1997; Fingelkurts and Fingelkurts, 2004, 2017; Deco and Kringelbach, 2016). Accordingly, our perspective view about the best strategy able to provide a coherent and complete description of timing can be divided in three steps: (1) the choice of tasks involving different aspects of timing (Coull and Nobre, 1998, 2008; Coull, 2004; Coull et al., 2013; Ciullo et al., 2018a) to be administered on a steady-state fashion (Gonzalez-Castillo and Bandettini, 2018; Tommasin et al., 2018) in order to saturate the activity of the areas interacting during the specific task; (2) the brain activity should be monitored by means of different techniques able to highlight different temporal and spatial scales (e.g., fMRI, hd-EEG, MEG). Specifically the different scales can be cast in a common framework according to the multilayer representation (De Domenico et al., 2013) (different spatial scales for the same time scale or different temporal scales for the same spatial one); (3) the temporal dynamics from each task will be finally analyzed and fitted to theoretical models of neuronal synchronization (Deco et al., 2017; Cavanna et al., 2018) in order to cluster the dynamics of brain's activity during time processing.

In the following paragraphs the core of each step is clarified and a review of the state-of-the-art is proposed.

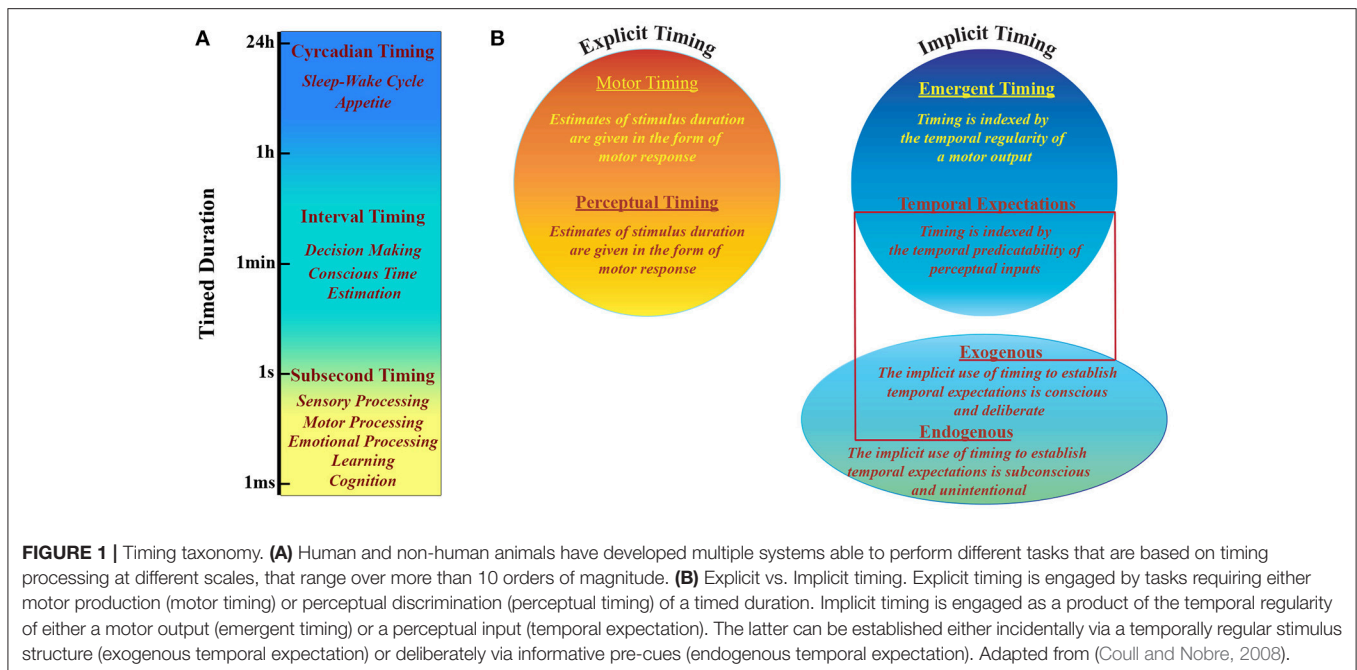
## TIMING IN HUMAN AND NON-HUMAN ANIMALS

The perception of what happens around us and the capacity to respond to it are crucially based on our ability of keeping track of time. Since both perception and action change over time, timing is necessary to estimate environmental dynamics, evaluate interplay between events and predict the consequences of our actions. Throughout normal development we acquire a sense of duration and rhythm that is basic to many behavioral aspects (Allman et al., 2012). Even if there is no specific system that senses time, human and non-human animals can estimate temporal intervals across a wide range of scales (Mauk and Buonomano, 2004; Buhusi and Meck, 2005). Intervals ranging from hundreds of milliseconds to seconds are typically associated with sensory and motor processing, learning, cognition and emotional processing (Figure 1), while larger intervals include processes that range from decision making to sleep-wake cycles (Buhusi and Meck, 2005). There is experimental evidence that

timing is an intrinsic computational ability of every circuit in the cortex and that it can be performed locally. This notion implies that during perception tasks cortical networks can tell time as a result of time-dependent changes in synaptic properties, which influence any population response to sensory events in a history dependent fashion (Karmarkar and Buonomano, 2007). Furthermore, with the above mentioned sensory timing, motor timing is supposed to depend on the activity of highly connected cortical recurrent networks able to self-sustain activity (Mauk and Buonomano, 2004).

Psychophysical experiments suggest that sensory timing can be local (Johnston et al., 2006; Burr et al., 2007; van Wassenhove and Nagarajan, 2007), even if other results suggest that temporal performance variability in different contexts may be better described by a hybrid model (Merchant et al., 2008). Neuroimaging research suggests that a partially distributed timing mechanism sustains contextual flexibility. It is supposed to be integrated by core structures such as the cortico-thalamic-basal ganglia (CTBG) circuit and regions that are selectively engaged by different behavioral contexts (Buhusi and Meck, 2005; Coull et al., 2011). Cell activity changes, associated with temporal processing in behaving monkeys, have been described in areas composing different circuits responsible for sensorimotor processing via the skeletomotor or oculomotor effector systems (Perrett, 1998; Lebedev et al., 2007; Tanaka, 2007; Genovesio et al., 2009; Mita et al., 2009). Most of these studies reported climbing activity during different timing conditions: discrimination of time, time estimation, single interval reproduction and delay related response. Specifically, Merchant et al. (2013) showed a variable discharge rate of cells of Medial Premotor Cortex (MPC) as a function of interval durations with a synchronization-continuation tapping task. This suggested the MPC might contain a representation of interval duration, in the hundred of milliseconds, where diverse populations of interval-tuned cells are typically activated according to the duration of the produced interval. Ramping activity of MPC cells encodes either the elapsed or the remaining time for a temporalized movement such that the dynamic organization of motor intentions and action is sustained by ramping cells. Accordingly, interval tuning on the overall discharge rate affects more cognitive facets of temporal processing.

By moving to larger temporal and spatial scales, functional magnetic resonance imaging (fMRI) studies in humans showed that interval timing is regulated by distributed brain networks whose involvement is flexibly adapted according to task demands: timing emerges from the interaction among diverse brain regions rather than from processing in a specific one (Livesey et al., 2007; Coull et al., 2008; Harrington et al., 2010; Fingelkurts, 2014). For example pattern of timing-related activation in bilateral caudate and putamen was found to be distinguished from that found for most other brain regions in time-perception tasks. Only the anterior insula was found to exhibit the same activation pattern. This region crucially integrates processing from disparate domains (e.g., interoception, emotion, and cognition), including time (Kosillo and Smith, 2010; Wittmann et al., 2010), via its dense pattern



of connections with most association areas in the basal ganglia and the occipital, temporal and prefrontal cortex. The connectedness of anterior insula with frontal cognitive control areas suggests that it supports the perceptual integration of sensory information (Eckert et al., 2009). By stimulating the supramarginal gyrus of the right hemisphere with transcranial magnetic stimulation (TMS) a dilation of perceived duration was induced because of its effect on interval encoding (Wiener et al., 2012). This result indicates that the neural circuitry that encodes time crucially includes the right supramarginal gyrus, confirming the detrimental effect of right parietal damage on time perception (Harrington et al., 1998). These findings support also the hypothesis of a network of multiple central clocks and distributed processes of timing mechanisms (Merchant et al., 2008). The ability to organize behaviors within periods in the range of seconds to minutes, depends on a cognitive system that requires multiple neuropsychological functions (Buhusi and Meck, 2005; Coull and Nobre, 2008), consequently pathophysiological distortions in time might reflect neuropsychological deficits typical of definite neuropsychiatric disorders as schizophrenia (Ciullo et al., 2016, 2018a), acquired brain injury (Piras et al., 2014), Parkinson's disease (Wearden et al., 2008), Huntington's disease (Beste et al., 2007) and attention-deficit hyperactivity disorder (Zelaznik et al., 2012). Thus, the understanding of timing mechanisms and of the related cognitive processes may also allow the realization of a model system aiming to characterize cognitive dysfunctions in order to define novel tools for early diagnosis and to develop novel targeted cognitive therapies. However, despite intensive investigations and substantial progress, the absence of a definitive framework encompassing the multifaceted nature of timing processes indicate that our understanding of the principles and mechanisms underlying brain functioning during

time perception remains still incomplete. Nonetheless, all the results described above emphasize the role of interactions among distributed neuronal populations at different spatiotemporal scales in enabling flexible cognitive operations that give rise to sense of time (Fingelkurts and Fingelkurts, 2006). Given the functional specialization and integration that sustain the sense of time, a promising framework able to provide a modeling of time perception in the brain from an explicitly integrative perspective is represented by complex network theory. Recent developments in the quantitative analysis of complex networks, based largely on graph theory, have been rapidly translated to studies of brain network organization. Accordingly, the brain is described as a network of nodes and edges, while analytic advancements in network science and statistics allow us to represent and quantify functional interactions among brain regions of interest in order to make inferences about its organizational properties both at rest and as a function of cognitive demands. To our knowledge, a network based description of brain regions integration in timing is still largely incomplete and actually available only in Ciullo et al. (2018b) and Ghaderi et al. (2018).

This kind of cerebral systems modeling (Bassett and Sporns, 2017), will be crucially beneficial in the close future to an organic description of brain functioning during the estimation of temporal intervals and eventually to a better description of disorders characterized by impaired time perception.

## MULTISCALE BRAIN NETWORKS

A tentative modeling of time perception processes necessarily points to a description of brain functioning based on the interplay of multi-scale brain networks (Fingelkurts et al., 2010; Bassett and Siebenhühner, 2013). The meaning of "scale" can vary according

to the context: (i) a network's spatial scale, which refers to the resolution at which its connected regions of interest (nodes) and connections (edges) are defined, and can range from individual cells and synapses size (Jarrell et al., 2012; Shimono and Beggs, 2015; Lee et al., 2016), to brain regions and fiber tracts (Bullmore and Bassett, 2011) and (ii) temporal scales with precision ranging from sub-millisecond (Burns et al., 2014), to lifetime (Betz et al., 2014; Gu et al., 2015). Although it is important to understand the functioning of individual elements, at each scale it is crucial to understand the sets of pair-wise relations that arrange the elements into the larger description of a totally interconnected system, namely the local and global topology of the network (Fingelkurts et al., 2010; Barabasi, 2016). Together these scales define a three-dimensional space in which the evolution of the brain network complexity is reported, being each point identified by three coordinates: space, time, and topology (Betz et al., 2017). Most descriptions of time perception mechanisms exist as single points in this space being based on analyses focused on networks defined singularly at one spatial, temporal, and topological scale. We anticipate that, while such studies have proven illuminating, in order to better understand the brain's true multi-scale, multi-level nature, it is essential that analyses begin to form bridges that link different scales to one another in order to offer a comprehensive description of the mechanisms that govern timing.

One promising approach to study a network that changes over multiple timescales is to make use of multi-layer network models of temporal networks (De Domenico et al., 2013; Kivelä et al., 2014). The multi-layer network model can treat estimates of the network's topology at different points of the time-scale as "layers." This implies the necessity to integrate different modalities of investigation spanning different time-scales. It could be done by creating a multi-layer from different non-invasive neuroimaging techniques: from high-density electroencephalography (hd-EEG) (Liu et al., 2017), to magnetoencephalography (MEG) (de Pasquale et al., 2010), fast fMRI (Lewis et al., 2016), classical fMRI (Telesford et al., 2016) and combined EEG-fMRI (Mullinger and Bowtell, 2010; Yu et al., 2016). On the other hand, invasive approaches are able to detect multiple single neuron signals in non-human animals (Logothetis, 2012) and in human patients that need deep brain stimulation (Okun, 2014). Traditional analysis would characterize each layer independently of one another, while multi-layer network analysis treats the ensemble of layers as a single unit, characterizing its structure as a whole to explicitly bridge multiple temporal scales. Since the multi-layer network model doesn't depend on the timescales represented by each layer, it can include any timescale made accessible using neuroimaging technologies.

As well as for time, the space dimension can be also investigated at multiple scales (**Figure 2A**). MEG and fMRI analyses of human brain networks are limited by the accuracy of the inverse source localization of signal generators (MEG), and the spatial granularity of the individual voxel (fMRI). Nonetheless, it is possible to probe multiple spatial scales by appropriately aggregating the minimal units of interest into parcels or regions of interest. Several parcellation approaches have been proposed, distinguishing to one another according to

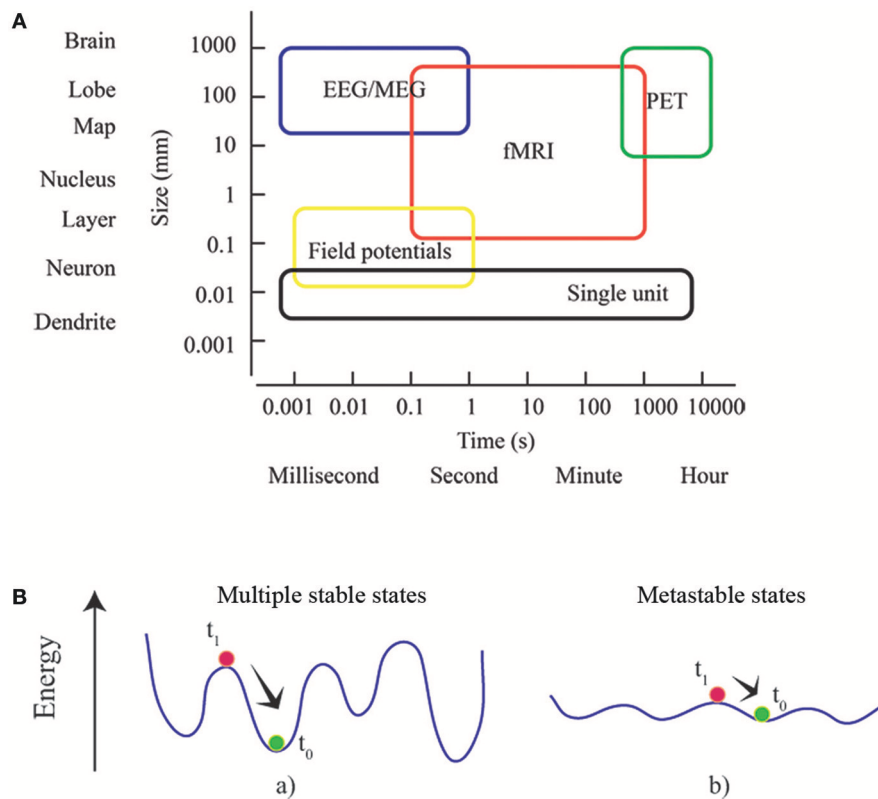
different criteria as spatial variation in functional connectivity, myelination, cytoarchitectonics, etc. (Tzourio-Mazoyer et al., 2002; Craddock et al., 2012; Wang et al., 2015; Glasser et al., 2016; Gordon et al., 2016). Since the choice of parcellation conditions the network's topology (Wang et al., 2009; Zalesky et al., 2010), it must be checked if any result is not driven by the specific choice of parcellation, and is reproducible (at least qualitatively) using a different set of parcels at the same resolution (Bassett et al., 2011). A route for future research is to apply multi-scale topological analysis to voxel-level networks during the execution of tasks. It will allow identifying different parcels differentially involved in different brain states in order to sub-divide specific brain areas responsible for sustaining different cognitive engagements.

## METASTABILITY: A RESOURCE OF BRAIN NETWORKS FOR SUSTAINING TIME PERCEPTION MECHANISMS

Large-scale brain networks have been showed to be organized according to multiple segregated sub-networks of interacting areas. It has been suggested that a dynamic, adaptable brain network arrangement in response to environmental stimulations underlies successful cognition (Bressler and Kelso, 2001; Fries, 2005). Dynamic combination of responses to sensory inputs, and spontaneous processing is at the core of brain activity, where task evoked responses should not be interpreted only in terms of localized processing, but should also take into account distributed processing occurring as activity flow across intrinsic networks (Smith et al., 2009; Zalesky et al., 2014; Sadaghiani et al., 2015; Cole et al., 2016; Shine et al., 2016). This allows a description of brain functioning in terms of a continuous recruitment of neuronal populations in a temporally coordinated fashion both during tasks execution, and at rest (Fingelkurts and Fingelkurts, 2005). Recently, it has been found that the neuronal engagement follows a precise hierarchy, according to two distinct sets of networks, or metastates, that the brain tends to cycle within (Vidaurre et al., 2017).

Metastates or metastable cerebral states are the core of a prominent conceptual framework known as Metastability (Scott Kelso, 1995; Fingelkurts and Fingelkurts, 2004, 2017; Freeman and Holmes, 2005; Werner, 2007). It offers a description of the reciprocal influence among interconnected parts and processes when pure synchronization does not occur. In coordination dynamics, such synchronization corresponds to stable fixed points of collective states (Friston, 1997). Metastability can be better understood by defining an energy landscape for the ensemble of possible states experienced by the brain: the phase space of the brain system (Fingelkurts and Fingelkurts, 2004, 2017). Generally, a system dynamically evolves attracted toward states of minimum energy, which can be either local or global. After being temporarily attracted toward a local state of minimum energy, an externally driven system can flee the basin of attraction and experience other equilibrium states. The dynamics of a metastable system is characterized by states that only transiently attract the dynamics. Since during its dynamic evolution the system





**FIGURE 2 |** Multiscale and multistable nature of the brain. **(A)** Rather than considering the brain as a list of parts defined at a particular scale, brain network theory take advantage of the complexity of the interactions between the parts, and identifies the dependence of phenomena across scales. Box dimensions give outer bounds of the spatial and temporal scales at which relational data are measured and interactions unfold. Adapted from an image of neuroscience recording methods (Sejnowski et al., 2014). **(B)** The concepts of energy landscape and metastability. Points of these landscapes correspond to particular states of the system. The system at equilibrium (green) is perturbed (at  $t_0$ ) toward a state (red) that subsequently (at  $t_1$ ) relaxes. In (a) the system is stable and local minima (equilibrium points) are deep: dynamics are rapidly restored and the effects of perturbation are short-lasting. In (b) the energy landscape is almost flat and the stability of local minima decreases: the system can easily explore different (metastable) states without an external driving or endogenous fluctuation.

tends to linger around these metastable states, the idea of a repertoire of conditions or configurations can be introduced (**Figure 2B**). Consequently, components are able to influence each other's destiny without being caught in a sustained state of synchronization, unable to create collectively new information (Scott Kelso, 1995; Tognoli and Kelso, 2014). The emergence of metastable dynamics has been theoretically showed to be contingent upon the coupling between modules of a dynamical system (Friston, 1997; Strogatz, 2001; Shanahan, 2010; Cabral et al., 2011). Specifically, dynamic patterns of functional brain networks, consistent with metastable dynamics, come out when coupling is topologically characterized by short average path lengths and high clustering (Wildie and Shanahan, 2012) of modules. The efficiency of task-related brain activity has been showed to depend on metastability of spontaneous brain activity, which allows for optimal experience of the dynamical repertoire (Cabral et al., 2014). Recently metastability in brain networks has been investigated in aging, consciousness and neuronal communication in healthy subjects (Deco and Kringelbach, 2016; Deco et al., 2017; Naik et al.,

2017; Cavanna et al., 2018) and in Schizophrenia and Alzheimer's disease patients (Córdova-Palomera et al., 2017; Koutsoukos and Angelopoulos, 2018). A variety of methods are described in order to capture synchronization and metastability in brain functioning.

Since metastability is a fundamental concept to grasp the behavior of complex systems theoretically and empirically, we anticipate that a form of metastability exists in time processing systems that parallels the metastability observed in many other aspects of brain functioning. The need for metastability in time perception modeling follows right from the definition. Metastability is the simultaneous occurrence of two competing tendencies: the inclination of individual components to exist as interacting entities and the propensity for the components to be characterized just by their independent behavior (Kelso, 2012). As a consequence metastability may be thought as a dynamical condition that allows the coordination of heterogeneous elements as it happens during time perception (brain areas having disparate intrinsic dynamics or brain areas whose activity is associated with different sensory, motor

and cognitive processes) (Fingelkurts and Fingelkurts, 2006; Fingelkurts, 2014). Metastable brain theory may ameliorate timing modeling as it does not favor extremes, e.g., integrated vs. segregated processes, but it tends to reconcile them. Since metastability is a characteristic of the full complexity of the brain, it reaches a maximum when the balance between segregative and integrative forces is found. Furthermore, metastability doesn't need active induction since no disengagement mechanisms are required, as it happens in timing processing (Kononowicz et al., 2016). Finally, the crucial importance of time to perception and action necessitates metastability, in order to explain the ease with which timing can be performed by a range of different neural architectures. Clustering the dynamics of brain's activity during time processing may unearth the presence of metastable states associated with this specific aspect of cognition.

## CONCLUSION

Here we propose that the route along which future research will find novel insight into timing mechanisms is drawn in the direction of brain investigation as a multi-layer dynamical network whose clustered dynamics unavoidably reports the presence of metastable states. This perspective paves the way

for future investigations into both the role of timing in other cognitive domains, from learning to agency, and the role that temporal dependency of brain network states has in cognition, elucidating the general characteristics of human cognitive activity that exists at a wide range of spatiotemporal scales. At the same time, our better understanding of dysfunctional timing processes will crucially allow us to develop novel diagnostics of neuropsychiatric diseases, and to design personalized therapeutics for rehabilitation and treatment of brain disorders characterized by distorted time perception.

## AUTHOR CONTRIBUTIONS

TG and GS conceived the paper. TG organized the paper. TG and VC wrote the first draft of the paper and collected and filtered the references. GS supervised the paper.

## ACKNOWLEDGMENTS

This study was conducted within the project Multidimensional study of timing abilities and sense of agency in schizophrenia and bipolar patients funded through 5Xmille 2016 from the Italian Ministry of Health.

## REFERENCES

- Allman, M. J., Pelphrey, K. A., and Meck, W. H. (2012). Developmental neuroscience of time and number: implications for autism and other neurodevelopmental disabilities. *Front. Integr. Neurosci.* 6:7. doi: 10.3389/fnint.2012.00007
- Barabasi, A. L. (2016). *Network Science*. Cambridge, UK: Cambridge University Press
- Bassett, D. S., Brown, J. A., Deshpande, V., Carlson, J. M., and Grafton, S. T. (2011). Conserved and variable architecture of human white matter connectivity. *Neuroimage* 54, 1262–1279. doi: 10.1016/j.neuroimage.2010.09.006
- Bassett, D. S., and Siebenhühner, F. (2013). "Multiscale network organization in the human brain," in *Multiscale Analysis and Nonlinear Dynamics: From Genes to the Brain*, ed M. M. Pesenson (Weinheim: Wiley-VCH), 179–204.
- Bassett, D. S., and Sporns, O. (2017). Network neuroscience. *Nat. Neurosci.* 20, 353–364. doi: 10.1038/nn.4502
- Beste, C., Saft, C., Andrich, J., Müller, T., Gold, R., and Falkenstein, M. (2007). Time processing in Huntington's disease: A group-control study. *PLoS ONE* 2:e1263. doi: 10.1371/journal.pone.0001263
- Betz, R. F., and Bassett, D. S. (2017). Multi-scale brain networks. *Neuroimage* 160, 73–83. doi: 10.1016/j.neuroimage.2016.11.006
- Betz, R. F., Byrge, L., He, Y., Goñi, J., Zuo, X. N., and Sporns, O. (2014). Changes in structural and functional connectivity among resting-state networks across the human lifespan. *Neuroimage* 102, 345–357. doi: 10.1016/j.neuroimage.2014.07.067
- Bressler, S. L., and Kelso, J. A. (2001). Cortical coordination dynamics and cognition. *Trends Cogn. Sci.* 5, 26–36. doi: 10.1016/S1364-6613(00)01564-3
- Bueno, F. D., Morita, V. C., De Camargo, R. Y., Reyes, M. B., Caetano, M. S., and Cravo, A. M. (2017). Dynamic representation of time in brain states. *Sci. Rep.* 7:46053. doi: 10.1038/srep46053
- 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
- Bullmore, E. T., and Bassett, D. S. (2011). Brain graphs: graphical models of the human brain connectome. *Annu. Rev. Clin. Psychol.* 7, 113–140. doi: 10.1146/annurev-clinpsy-040510-143934
- Buonomano, D. V., and Maass, W. (2009). State-dependent computations: spatiotemporal processing in cortical networks. *Nat. Rev. Neurosci.* 10, 113–125. doi: 10.1038/nrn2558
- Burns, S. P., Santaniello, S., Yaffe, R. B., Jouny, C. C., and Crone, N. E. (2014). Network dynamics of the brain and influence of the epileptic seizure onset zone. *Proc. Natl. Acad. Sci. U.S.A.* 111, E5321–E5330. doi: 10.1073/pnas.1401752111
- Burr, D., Tozzi, A., and Morrone, M. C. (2007). Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nat. Neurosci.* 10, 423–425. doi: 10.1038/nn1874
- Cabral, J., Hugues, E., Sporns, O., and Deco, G. (2011). Role of local network oscillations in resting-state functional connectivity. *Neuroimage* 57, 130–139. doi: 10.1016/j.neuroimage.2011.04.010
- Cabral, J., Kringelbach, M. L., and Deco, G. (2014). Exploring the network dynamics underlying brain activity during rest. *Prog. Neurobiol.* 114, 102–131. doi: 10.1016/j.pneurobio.2013.12.005
- Cavanna, F., Vilas, M. G., Palmucci, M., and Tagliazucchi, E. (2018). Dynamic functional connectivity and brain metastability during altered states of consciousness. *Neuroimage* 180, 383–395. doi: 10.1016/j.neuroimage.2017.09.065
- Ciullo, V., Piras, F., Vecchio, D., Banaj, N., Coull, J. T., and Spalletta, G. (2018a). Predictive timing disturbance is a precise marker of schizophrenia. *Schizophr. Res. Cogn.* 12, 42–49. doi: 10.1016/j.scog.2018.04.001
- 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. *Schizophr. Bull.* 42, 505–518. doi: 10.1093/schbul/sbv104
- Ciullo, V., Vecchio, D., Gili, T., Spalletta, G., and Piras, F. (2018b). Segregation of Brain Structural Networks Supports Spatio-Temporal Predictive Processing. *Front. Hum. Neurosci.* 12:212. doi: 10.3389/fnhum.2018.00212
- Cole, M. W., Ito, T., Bassett, D. S., and Schultz, D. H. (2016). Activity flow over resting-state networks shapes cognitive task activations. *Nat. Neurosci.* 19, 1718–1726. doi: 10.1038/nn.4406
- Córdova-Palomera, A., Kaufmann, T., Persson, K., Alnæs, D., Doan, N. T., Moberget, T., et al. (2017). Disrupted global metastability and static and dynamic brain connectivity across individuals in the Alzheimer's disease continuum. *Sci. Rep.* 7:40268. doi: 10.1038/srep40268

- Coull, J., and Nobre, A. (2008). Dissociating explicit timing from temporal expectation with fMRI. *Curr. Opin. Neurobiol.* 18, 137–144. doi: 10.1016/j.conb.2008.07.011
- Coull, J. T. (2004). Functional anatomy of the attentional modulation of time estimation. *Science* 303, 1506–1508. doi: 10.1126/science.1091573
- 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., Davranche, K., Nazarian, B., and Vidal, F. (2013). Functional anatomy of timing differs for production versus prediction of time intervals. *Neuropsychologia* 51, 309–319. doi: 10.1016/j.neuropsychologia.2012.08.017
- 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
- Coull, J. T., and Nobre, A. C. (1998). Where and when to pay attention: the neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *J. Neurosci.* 18, 7426–7435.
- Craddock, R. C., James, G. A., Holtzheimer, P. E., Hu, X. P., and Mayberg, H. S. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Mapp.* 33, 1914–1928. doi: 10.1002/hbm.21333
- De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivela, M., Moreno, Y., Porter, M. A., et al. (2013). Mathematical formulation of multilayer networks. *Phys. Rev. X* 3:041022. doi: 10.1103/PhysRevX.3.041022
- de Pasquale, F., Della Penna, S., Snyder, A. Z., Lewis, C., Mantini, D., Marzetti, L., et al. (2010). Temporal dynamics of spontaneous MEG activity in brain networks. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6040–6045. doi: 10.1073/pnas.0913863107
- Deco, G., and Kringelbach, M. L. (2016). Metastability and coherence: extending the communication through coherence hypothesis using a whole-brain computational perspective. *Trends Neurosci.* 39, 125–135. doi: 10.1016/j.tins.2016.01.001
- Deco, G., Kringelbach, M. L., Jirsa, V. K., and Ritter, P. (2017). The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core. *Sci. Rep.* 7:3095. doi: 10.1038/s41598-017-03073-5
- Eckert, M. A., Menon, V., Walczak, A., Ahlstrom, J., Denslow, S., Horwitz, A., et al. (2009). At the heart of the ventral attention system: the right anterior insula. *Hum. Brain Mapp.* 30, 2530–2541. doi: 10.1002/hbm.20688
- Fingelkurts, A., and Fingelkurts, A. (2017). Information flow in the brain: ordered sequences of metastable states. *Information* 8:22. doi: 10.3390/info8010022
- Fingelkurts, A. A. (2014). Present moment, past, and future: mental kaleidoscope. *Front. Psychol.* 5:395. doi: 10.3389/fpsyg.2014.00395
- Fingelkurts, A. A., and Fingelkurts, A. A. (2004). Making complexity simpler: multivariability and metastability in the brain. *Int. J. Neurosci.* 114, 843–862. doi: 10.1080/00207450490450046
- Fingelkurts, A. A., and Fingelkurts, A. A. (2005). “Mapping of brain operational architectonics,” in *Focus on Brain Mapping Research*, ed F. J. Chen (New York, NY: Nova Science Publishers, Inc).
- Fingelkurts, A. A., and Fingelkurts, A. A. (2006). Timing in cognition and EEG brain dynamics: discreteness versus continuity. *Cogn. Process.* 7, 135–162. doi: 10.1007/s10339-006-0035-0
- Fingelkurts, A. A., Fingelkurts, A. A., and Neves, C. F. H. (2010). Natural world physical, brain operational, and mind phenomenal space-time. *Phys. Life Rev.* 7, 195–249. doi: 10.1016/j.plrev.2010.04.001
- Freeman, W. J., and Holmes, M. D. (2005). Metastability, instability, and state transition in neocortex. *Neural Networks* 18, 497–504. doi: 10.1016/j.neunet.2005.06.014
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474–480. doi: 10.1016/j.tics.2005.08.011
- Friston, K. J. (1997). Transients, metastability, and neuronal dynamics. *Neuroimage* 5, 164–171. doi: 10.1006/nimg.1997.0259
- Genovesio, A., Tsujimoto, S., and Wise, S. P. (2009). Feature- and order-based timing representations in the frontal cortex. *Neuron* 63, 254–266. doi: 10.1016/j.neuron.2009.06.018
- Ghaderi, A. H., Moradkhani, S., Haghighatfard, A., Akrami, F., Khayer, Z., and Balci, F. (2018). Time estimation and beta segregation: an EEG study and graph theoretical approach. *PLoS ONE* 13:e0195380. doi: 10.1371/journal.pone.0195380
- 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
- Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., et al. (2016). A multi-modal parcellation of human cerebral cortex. *Nature* 536, 171–178. doi: 10.1038/nature18933
- Gonzalez-Castillo, J., and Bandettini, P. A. (2018). Task-based dynamic functional connectivity: recent findings and open questions. *Neuroimage* 180, 526–533. doi: 10.1016/j.neuroimage.2017.08.006
- Gordon, E. M., Laumann, T. O., Adeyemo, B., Huckins, J. F., Kelley, W. M., and Petersen, S. E. (2016). Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb. Cortex* 26, 288–303. doi: 10.1093/cercor/bhu239
- Gu, S., Satterthwaite, T. D., Medaglia, J. D., Yang, M., Gur, R. E., Gur, R. C., et al. (2015). Emergence of system roles in normative neurodevelopment. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13681–13686. doi: 10.1073/pnas.1502829112
- Harrington, D. L., Haaland, K. Y., and Knight, R. T. (1998). Cortical networks underlying mechanisms of time perception. *J. Neurosci.* 18, 1085–1095
- 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
- Hass, J., and Durstewitz, D. (2016). Time at the center, or time at the side? Assessing current models of time perception. *Curr. Opin. Behav. Sci.* 8, 238–244. doi: 10.1016/j.cobeha.2016.02.030
- Jarrell, T. A., Wang, Y., Bloniarz, A. E., Brittin, C. A., Xu, M., Thomson, J. N., et al. (2012). The connectome of a decision-making neural network. *Science* 337, 437–444. doi: 10.1126/science.1221762
- 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
- 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
- Kelso, J. A. (2012). Multistability and metastability: understanding dynamic coordination in the brain. *Philos. Trans. R. Soc. B Biol. Sci.* 367, 906–918. doi: 10.1098/rstb.2011.0351
- Kivela, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., and Porter, M. A. (2014). Multilayer networks. *J. Complex Networks* 2, 203–271. doi: 10.1093/comnet/cnu016
- Kononowicz, T. W., van Rijn, H., and Meck, W. H. (2016). Timing and time perception: a critical review of neural timing signatures before, during, and after the To-Be-Timed Interval, in *Stevens’ Handbook of Experimental Psychology and Cognitive Neuroscience, 4th Edn.* ed S. Ghetti (New York, NY: John Wiley & Sons), 1–48.
- Kosillo, P., and Smith, A. T. (2010). The role of the human anterior insular cortex in time processing. *Brain Struct. Funct.* 214, 623–628. doi: 10.1007/s00429-010-0267-8
- Koutsoukos, E., and Angelopoulos, E. (2018). Investigating the role of metastability in the disturbed brain functional connectivity observed under the experience of thought blocks in schizophrenia. *Schizophr. Res.* 193, 98–99. doi: 10.1016/j.schres.2017.07.055
- Lebedev, M. A., O’Doherty, J. E., and Nicolelis, M. A. (2007). Decoding of temporal intervals from cortical ensemble activity. *J. Neurophysiol.* 99, 166–186. doi: 10.1152/jn.00734.2007
- Lee, W. C., Bonin, V., Reed, M., Graham, B. J., Hood, G., Glatfelter, K., et al. (2016). Anatomy and function of an excitatory network in the visual cortex. *Nature* 532, 370–374. doi: 10.1038/nature17192
- Lewis, L. D., Setsompop, K., Rosen, B. R., and Polimeni, J. R. (2016). Fast fMRI can detect oscillatory neural activity in humans. *Proc. Natl. Acad. Sci. U.S.A.* 113, E6679–E6685. doi: 10.1073/pnas.1608117113
- Liu, Q., Farahibozorg, S., Porcaro, C., Wenderoth, N., and Mantini, D. (2017). Detecting large-scale networks in the human brain using high-density electroencephalography. *Hum. Brain Mapp.* 38, 4631–4643. doi: 10.1002/hbm.23688
- Livesey, A. C., Wall, M. B., and Smith, A. T. (2007). Time perception: manipulation of task difficulty dissociates clock functions from other cognitive demands. *Neuropsychologia* 45, 321–331. doi: 10.1016/j.neuropsychologia.2006.06.033

- Logothetis, N. K. (2012). Intracortical recordings and fMRI: an attempt to study operational modules and networks simultaneously. *Neuroimage* 62, 962–969. doi: 10.1016/j.neuroimage.2012.01.033
- Mauk, M. D., and Buonomano, D. V. (2004). The neural basis of temporal processing. *Annu. Rev. Neurosci.* 27, 307–340. doi: 10.1146/annurev.neuro.27.070203.144247
- Merchant, H., Pérez, O., Zarco, W., and Gámez, 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., Bartolo, R., and Prado, L. (2008). The context of temporal processing is represented in the multidimensional relationships between timing tasks. *PLoS ONE* 3:e3169. doi: 10.1371/journal.pone.0003169
- 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
- Mullinger, K., and Bowtell, R. (2010). Combining EEG and fMRI. *Methods Mol. Biol.* 711, 303–326. doi: 10.1007/978-1-61737-992-5\_15
- Naik, S., Banerjee, A., Bapi, R. S., Deco, G., and Roy, D. (2017). Metastability in senescence. *Trends Cogn. Sci.* 21, 509–521. doi: 10.1016/j.tics.2017.04.007
- Okun, M. S. (2014). Deep-brain stimulation — entering the era of human neural-network modulation. *N. Engl. J. Med.* 371, 1369–1373. doi: 10.1056/NEJMp1408779
- Perrett, S. P. (1998). Temporal discrimination in the cerebellar cortex during conditioned eyelid responses. *Exp. Brain Res.* 121, 115–124. doi: 10.1007/s002210050443
- 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
- Sadaghiani, S., Poline, J.-B., Kleinschmidt, A., and D'Esposito, M. (2015). Ongoing dynamics in large-scale functional connectivity predict perception. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8463–8468. doi: 10.1073/pnas.1420687112
- Scott Kelso, J. A. (1995). “The self-organization of brain and behavior,” in *Dynamic Patterns* (Cambridge, MA: MIT Press), 1–28.
- Sejnowski, T. J., Churchland, P. S., and Movshon, J. A. (2014). Putting big data to good use in neuroscience. *Nat. Neurosci.* 17, 1440–1441. doi: 10.1038/nn.3839
- Shanahan, M. (2010). Metastable chimera states in community-structured oscillator networks. *Chaos* 20:013108. doi: 10.1063/1.3305451
- Shimono, M., and Beggs, J. M. (2015). Functional clusters, hubs, and communities in the cortical microconnectome. *Cereb. Cortex* 25, 3743–3757. doi: 10.1093/cercor/bhu252
- Shine, J. M., Bissett, P. G., Bell, P. T., Koyejo, O., Balsters, J. H., Gorgolewski, K. J., et al. (2016). The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron* 92, 544–554. doi: 10.1016/j.neuron.2016.09.018
- Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., et al. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13040–13045. doi: 10.1073/pnas.0905267106
- Strogatz, S. H. (2001). Exploring complex networks. *Nature* 410, 268–276. doi: 10.1038/35065725
- Tanaka, M. (2007). Cognitive signals in the primate motor thalamus predict saccade timing. *J. Neurosci.* 27, 12109–12118. doi: 10.1523/JNEUROSCI.1873-07.2007
- Telesford, Q. K., Lynall, M.-E., Vettel, J., Miller, M. B., Grafton, S. T., and Bassett, D. S. (2016). Detection of functional brain network reconfiguration during task-driven cognitive states. *Neuroimage* 142, 198–210. doi: 10.1016/j.neuroimage.2016.05.078
- Tognoli, E., and Kelso, J. A. (2014). The metastable brain. *Neuron* 81, 35–48. doi: 10.1016/j.neuron.2013.12.022
- Tommasin, S., Mascali, D., Moraschi, M., Gili, T., Hassan, I. E., Fratini, M., et al. (2018). Scale-invariant rearrangement of resting state networks in the human brain under sustained stimulation. *Neuroimage* 179, 570–581. doi: 10.1016/j.neuroimage.2018.06.006
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- van Wassenhove, V., and Nagarajan, S. S. (2007). Auditory cortical plasticity in learning to discriminate modulation rate. *J. Neurosci.* 27, 2663–2672. doi: 10.1523/JNEUROSCI.4844-06.2007
- Vidaurre, D., Smith, S. M., and Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proc. Natl. Acad. Sci. U.S.A.* 114, 12827–12832. doi: 10.1073/pnas.1705120114
- Wang, D., Buckner, R. L., Fox, M. D., Holt, D. J., Holmes, A. J., Stoecklein, S., et al. (2015). Parcellating cortical functional networks in individuals. *Nat. Neurosci.* 18, 1853–1860. doi: 10.1038/nn.4164
- Wang, J., Wang, L., Zang, Y., Yang, H., Tang, H., Gong, Q., et al. (2009). Parcellation-dependent small-world brain functional networks: a resting-state fmri study. *Hum. Brain Mapp.* 30, 1511–1523. doi: 10.1002/hbm.20623
- Wearden, J. H., Smith-Spark, J. H., Cousins, R., Edelstyn, N. M., Cody, F. W., and O'Boyle, D. J. (2008). Stimulus timing by people with Parkinson's disease. *Brain Cogn.* 67, 264–279. doi: 10.1016/j.bandc.2008.01.010
- Werner, G. (2007). Metastability, criticality and phase transitions in brain and its models. *BioSystems* 90, 496–508. doi: 10.1016/j.biosystems.2006.12.001
- Wiener, M., Klotz, D., Turkeltaub, P. E., Hamilton, R. H., Wolk, D. A., and Coslett, H. B. (2012). Parietal influence on temporal encoding indexed by simultaneous transcranial magnetic stimulation and electroencephalography. *J. Neurosci.* 32, 12258–12267. doi: 10.1523/JNEUROSCI.2511-12.2012
- Wildie, M., and Shanahan, M. (2012). Metastability and chimera states in modular delay and pulse-coupled oscillator networks. *Chaos* 22:043131. doi: 10.1063/1.4766592
- Wittmann, M., Simmons, A. N., Aron, J. L., and Paulus, M. P. (2010). Accumulation of neural activity in the posterior insula encodes the passage of time. *Neuropsychologia* 48, 3110–3120. doi: 10.1016/j.neuropsychologia.2010.06.023
- Yu, Q., Wu, L., Bridwell, D. A., Erhardt, E. B., Du, Y., He, H., et al. (2016). Building an EEG-fMRI multi-modal brain graph: a concurrent EEG-fMRI study. *Front. Hum. Neurosci.* 10:476. doi: 10.3389/fnhum.2016.00476
- Zalesky, A., Fornito, A., Cocchi, L., Gollo, L. L., and Breakspear, M. (2014). Time-resolved resting-state brain networks. *Proc. Natl. Acad. Sci. U.S.A.* 111, 10341–10346. doi: 10.1073/pnas.1400181111
- Zalesky, A., Fornito, A., Harding, I. H., Cocchi, L., Yücel, M., Pantelis, C., et al. (2010). Whole-brain anatomical networks: does the choice of nodes matter? *Neuroimage* 50, 970–983. doi: 10.1016/j.neuroimage.2009.12.027
- Zelaznik, H. N., Vaughn, A. J., Green, J. T., Smith, A. L., Hoza, B., and Linnea, K. (2012). Motor timing deficits in children with Attention-Deficit/Hyperactivity disorder. *Hum. Mov. Sci.* 31, 255–265. doi: 10.1016/j.humov.2011.05.003

**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 © 2018 Gili, Ciullo and Spalletta. 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) and the copyright owner(s) 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.





# Slower Is Higher: Threshold Modulation of Cortical Activity in Voluntary Control of Breathing Initiation

Pierre Pouget<sup>1\*</sup>, Etienne Allard<sup>2†</sup>, Tymothée Poitou<sup>2</sup>, Mathieu Raux<sup>1,3</sup>, Nicolas Wattiez<sup>1</sup> and Thomas Similowski<sup>1,3\*</sup>

<sup>1</sup> UMRS 975, INSERM, CNRS 7225, Institute of Brain and Spinal Cord, UPMC - University Pierre and Marie Curie, Paris, France, <sup>2</sup> UMRS1158, INSERM, Neurophysiologie Respiratoire Expérimentale et Clinique, Sorbonne Universités, UPMC - University Pierre and Marie Curie, Paris, France, <sup>3</sup> Service de Pneumologie et Réanimation Médicale (Département "R3S"), AP-HP, Groupe Hospitalier Pitié-Salpêtrière Charles Foix, Paris, France

## OPEN ACCESS

### Edited by:

Arpan Banerjee,  
National Brain Research Centre  
(NBRC), India

### Reviewed by:

Leszek Kubin,  
University of Pennsylvania,  
United States  
Thiago S. Moreira,  
Universidade de São Paulo, Brazil

### \*Correspondence:

Pierre Pouget  
pierre.pouget@upmc.fr  
Thomas Similowski  
thomas.similowski@psl.aphp.fr

<sup>†</sup> These authors are co-first authors

### Specialty section:

This article was submitted to  
Decision Neuroscience,  
a section of the journal  
Frontiers in Neuroscience

**Received:** 12 March 2018

**Accepted:** 04 September 2018

**Published:** 11 October 2018

### Citation:

Pouget P, Allard E, Poitou T, Raux M,  
Wattiez N and Similowski T (2018)  
Slower Is Higher: Threshold  
Modulation of Cortical Activity  
in Voluntary Control of Breathing  
Initiation. *Front. Neurosci.* 12:663.  
doi: 10.3389/fnins.2018.00663

Speech or programmed sentences must often be interrupted in order to listen to and interact with interlocutors. Among many processes that produce such complex acts, the brain must precisely adjust breathing to produce adequate phonation. The mechanism of these adjustments is multifactorial and still poorly understood. In order to selectively examine the adjustment in breath control, we recorded respiratory-related premotor cortical potentials from the scalp of human subjects while they performed a single breathing initiation or inhibition task. We found that voluntary breathing is initiated if, and only if, the cortical premotor potential activity reaches a threshold activation level. The stochastic variability in the threshold correlates to the distribution of initiation times of breathing. The data also fitted a computerized interactive race model. Modeling results confirm that this model is also as effective in respiratory modality, as it has been found to be for eye and hand movements. No modifications were required to account for respiratory cycle inhibition processes. In this overly simplified task, we showed a link between voluntary initiation and control of breathing and activity in a fronto-median region of the cerebral cortex. These results shed light on some of the physiological constraints involved in the complex mechanisms of respiration, phonation, and language.

**Keywords:** motor control, countermanding task, breathing, decision making, inhibition (psychology)

## INTRODUCTION

In vertebrate animals, the central nervous system generates a rhythmic command that drives the contraction of respiratory muscles in order to move air in and out of the lungs. The mechanisms of this automatic control have been deeply investigated. This control relies primarily on groups of brainstem neurons in dynamic interaction (Feldman et al., 2013) that generate the respiratory rhythm and adjust it to the metabolic activity of the body. Voluntary breathing commands can also arise from higher brain structures,

and the respiratory muscles are represented within the primary motor cortex (Smith, 1938; Gandevia and Rothwell, 1987; Similowski et al., 1996). Patients with locked-in syndrome retain emotional influences on breathing but have no voluntary control of respiratory movements (Heywood et al., 1996). Over the last decade, functional studies in humans have shown that neuronal activity in the premotor cortex and the supplementary motor areas are also involved when subjects are exposed to inspiratory resistance and breathe against it without being instructed to do so (Raux et al., 2007, 2013). Cortico-subcortical cooperation in generating the neural drive to breathe has been demonstrated in patients with deficient respiratory automatism (Tremoureaux et al., 2014a), in normal subjects during hypocapnia-related inhibition of the respiratory automatism (Dubois et al., 2016), in patients with inspiratory muscle weakness (Georges et al., 2016), and in patients with abnormally high inspiratory resistances (Launois et al., 2015). In all cases their electroencephalographic activity suggests involvement of the premotor cortex. In this study we pushed the argument further by testing the hypothesis of a causal involvement of the cerebral cortex in the voluntary initiation and inhibition of a single breath command, and tested it by recording respiratory-related premotor cortical potentials in the scalp of awake subjects during such maneuvers.

Although, research on voluntary respiratory control is most often based on neurophysiological studies, computational modeling also has a role to play. Among the multitude of existing computational models, the race model offers an interesting method to investigate breathing modality in the context of an inhibitory task. This task consists of two types of trials: Go trials, where subjects were instructed to perform an action as quickly as possible, and Stop trials, where subjects had to inhibit this action. This paradigm is useful when studying the ability of a subject to inhibit an action, and allows to the time needed to stop an action to be assessed (Stop Signal Reaction Time or SSRT), which is not observable directly. Logan et al. (1984) developed a race model to estimate the SSRT in an oculomotor countermanding task. Beyond the access to the time needed to cancel an action, this model provides a way to explore functional mechanisms responsible for inhibition performance. This model involves two independent units, a go and a stop process, performing a race from a baseline until one of these processes crossed an arbitrary threshold; the winner of this race is the first process that crosses the threshold. More recently, Boucher et al. (2007a) added interactions between go and stop units to account for electrophysiological recordings of single units in macaques. Race models have been tested in eye and hand modality (Boucher et al., 2007b) but never, to our knowledge, on respiratory modality. In this study, we tested the modified Boucher et al.'s race model to assess whether it would be qualitatively adequate to account for respiratory data in the particular context of a countermanding task.

Our study focused on the magnitude and timing of the fronto-median cortical premotor potential activity, examining whether its stochastic variability could account for breathing initiation time and inhibition. This study is an important contribution to an improved understanding of how humans initiate vocalizations

and the disentanglement of the distinct and shared processes within the complex mechanisms of respiration, phonation, and language.

## MATERIALS AND METHODS

### Subjects and Session Design

This study was part of a wider respiratory-related cortical activity research program that has been approved by the local ethical committee. The subjects gave informed consent to participate. Data were collected from six human subjects (five males, mean age 31 years  $\pm$  8). Each subject participated in two sessions with 256 NoStop trials and two sessions with 128 NoStop and 128 Stop trials. Each session lasted for approximately 60 min. The trial order and the session order were both randomized across subjects. All subjects reported having normal or corrected-to-normal vision.

### Ventilatory Movement Recording

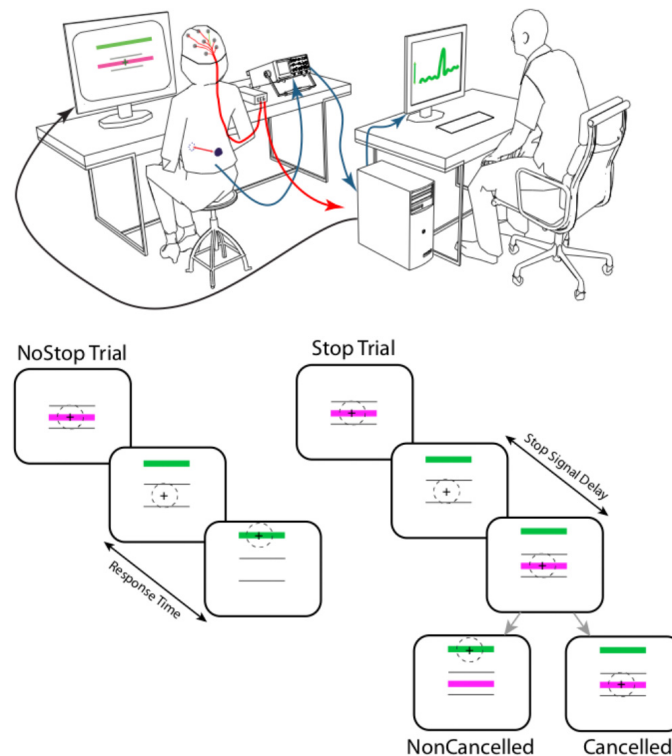
The subjects' ventilatory movements were measured using custom-built magnetometers (Mead et al., 1967). Two magnets were positioned, one ventrally and one dorsally, at the level of the umbilicus using an elastic belt. This allowed us to measure abdominal expansion, a direct indicator of diaphragmatic contraction insofar as the diaphragm is the only muscle whose contraction increases abdominal circumference. A PC with a NI-DAQ analog acquisition card (National Instruments Corp., Austin, TX, United States) running the *Xenomai* operating system for parallel real-time acquisition (sampling frequency 1,000 Hz) recorded ventilatory movements and the various stimuli presented on-screen during the sessions.

### Initiation Breathing Task – NoStop Trials

Subjects sat 57 cm away from a TV display monitor (Dell 21"). Each trial started with the presentation of two purple horizontal bars centered on a black background. The two bars were vertically separated by 10° of visual angle (**Figure 1**). After a random delay ranging from 500 to 1,000 ms, a green bar (Go signal) appeared 6° above the top fixation bar, instructing the subject to initiate inspiration. To provide feedback of the amplitude of the corresponding evoked respiratory response, the subject's abdominal movements were displayed on the screen as a cross that moved up and down with abdominal expansion (see below). A calibration procedure was performed at the beginning of each block to adjust gain and offset so the respiration amplitude modulation remained within the range of the breathing initiation bars.

### Countermanding Breathing Task – Stop Trials

As in the initiation-breathing task, each trial started with the presentation of two purple horizontal bars centered on a black background (**Figure 1**). The two bars were vertically separated by 10° of visual angle. After a random delay ranging from 500 to 1,000 ms, a green bar (Go signal) appeared 6° above the top



**FIGURE 1 |** Experimental setup: the respiratory movements of the subjects were assessed using a magnetometer that measured the distance between two magnets. These were placed on the abdomen and on the back at the height of the umbilicus, and maintained in position with an elastic belt. This allowed us to measure abdominal expansion, an indirect indicator of diaphragmatic contraction. During respiratory tasks, electroencephalographic potential (ERPs) signal were recorded. Each trial started with the presentation of two purple horizontal bars ( $10^\circ$  length of visual angle) centered on a black background. The two bars were vertically separated by  $10^\circ$  of visual angle. After a random delay ranging from 500 to 1,000 ms, a green bar (Go signal) appeared  $6^\circ$  above the top fixation bar, instructing the subject to initiate an inspiration. Abdominal movements were displayed on the subject's screen as a cross that moved up and down in relation to abdominal expansion and contraction to provide feedback to the subject of the amplitude of the evoked respiratory response. On 50% of trials (at random), after a delay (Stop Signal Delay- SSD) ranging from 48 to 640 ms, a green bar was presented in the center of the screen, instructing the subject to stop the inspiration. A failure to inhibit inspiration was classed as a non-cancelled trial, while a successfully inhibited inspiration was classed as a canceled trial.

fixation bar, instructing the subject to initiate an inspiration. Abdominal movements were displayed on the subject's screen as a cross that moved up and down according to abdominal expansion and contraction to provide feedback to the subject of the amplitude of the evoked respiratory response. A calibration procedure was performed at the beginning of each block to adjust gain and offset so the respiration amplitude modulation remained within the range of the breathing initiation bars. In a second type of trial – the task started with the presentation of two purple horizontal bars centered on a black background. On 50% of trials (at random), after a delay (Stop Signal Delay- SSD) ranging from 48 to 640 ms, a green bar was presented in the center of the screen, instructing the subject to stop his/her inspiration. A failure to inhibit inspiration was classed as a non-cancelled trial, while a successfully inhibited inspiration was classed as a canceled trial.

## Reaction Time Measurement

The ventilatory movement signal was processed offline using Matlab (MATLAB Release 2012b, The MathWorks, Inc., Natick, MA, United States). The onset of breathing movement was

determined using a derivative of the abdominal expansion signal based on a threshold limit compared with the resting breathing signal. The reaction time was the difference between a Go signal presentation and the onset of inspiration.

## EEG Signal Recording

During respiratory tasks, electroencephalographic (EEG) signal was recorded using nine active electrodes positioned according to the international 10–20 system, recorded with a V-Amp system, (Brain Product, Munich, Germany). The reference was calculated from the electrodes A1 and A2. The impedance of each electrode was estimated between 5 and 10 k $\Omega$  and was always lower less than 25 k $\Omega$ . Abdominal ventilatory movements and the EEG signal were later synchronized using markers.

## EEG Signal Processing

Ensemble averaging was first performed to improve the signal-to-noise ratio and reveal the potentials, in a manner typical to the study of evoked potentials. The continuously recorded electroencephalographic signal was split into three epochs, each of one second, extending from 0.5 s before to 2.5 s

after the Go signal presentation (green bar). A thresholding method was used to detect artifact and periods exhibiting activity  $\pm 3$  standard deviation of the mean were discarded. The rejection rate was approximately 30% in the various sessions. Trials were sorted into five (NoStop sessions) or three (NoStop trials of Stop sessions) groups according to reaction time, and the EEG signal was averaged point by point.

## Race Model

The interactive race model is composed of a go unit and a stop unit. Their activity is governed by two stochastic differential equations (Usher and McClelland, 2001) with a null leakage factor:

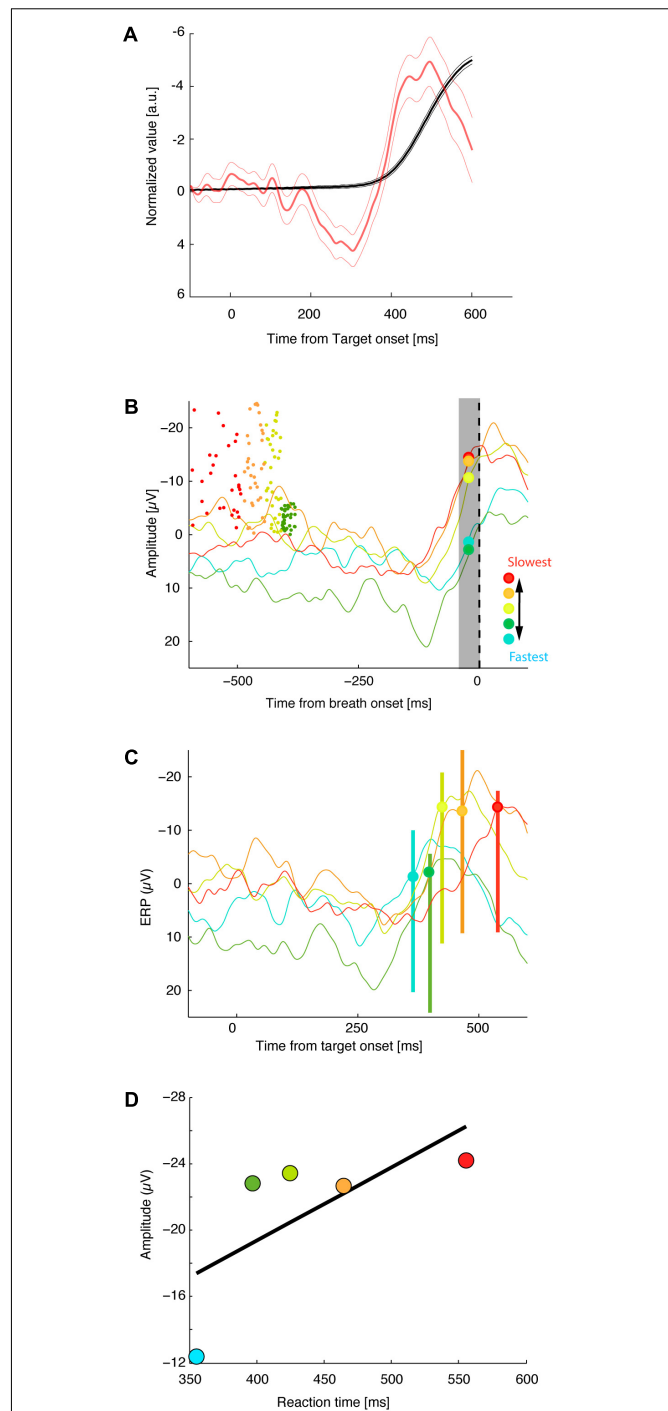
$$da_{go}(t) = \mu_{go} - \beta_{stop} \cdot a_{stop}(t) + \zeta_{go} \quad (1)$$

$$da_{stop}(t) = \mu_{stop} - \beta_{go} \cdot a_{go}(t) + \zeta_{stop} \quad (2)$$

Each unit is defined by three parameters: the mean growth rate ( $\mu$ ); the inhibition parameter on the other unit ( $\beta$ ); and a Gaussian noise term ( $\zeta$ ) with a mean of zero variance of  $\sigma_{go}^2$  or  $\sigma_{stop}^2$ , where  $a$  represents the activity of the unit. The race finished when a unit crossed the threshold, fixed at 1,000 (arbitrary units), within a limit of 800 ms. If unit activity is negative during the race, the activity was reset to zero at this point (non-physiologic value). An additional parameter,  $D$ , was added to the model to take into account the stimulus encoding that occurred in the go and stop units (respectively  $D_{go}$  and  $D_{stop}$ ).  $D_{go}$  was determined and set constant for each subject using values from the EEG recordings.  $D_{go}$  was calculated when the activity signal equaled the baseline mean plus five standard deviations in the EEG signal.  $D_{stop}$ ,  $\mu$ ,  $\beta$ , and  $\sigma$  are the unconstrained parameters of the model.

We tested these interactive race models to find the parameters that best fitted the data from six subjects who performed two sessions of a breathing countermanding task. Two sessions were removed from the analysis because of bimodality observed in reaction times.

We fitted inhibition function, reaction time distributions of correct NoStop trials and failed stop-signal trials. For each fit, we computed a chi-square test between data from the model and the subjects. Inhibition function chi-squares were calculated by summing chi-squares computed at each stop-signal-delay (SSD) between error rates from model and subject data. The chi-squares of the reaction times of correct and failed NoStop stop-signal trials were calculated as follows: each distribution was sorted into five quintiles; a “local” chi-square was computed at each quintile between the proportion of trials from model regarding subjects’ data, then the five “local” chi-squares were added. A general chi-square was then calculated by adding the local chi-squares together. To find the best parameters for each subject, we minimized the general chi-square using a minimization function (*patternsearch* from the global optimization toolbox of Matlab). *Patternsearch* looks for a minimum based on an adaptive *mesh* that is aligned with the coordinate directions. Because minimization functions are sensitive to start point, we



**FIGURE 2 |** Threshold of premotor activity response as a function of breath initiation times. **(A)** Normalized average of the diaphragmatic signal (black lines) and normalized activity at Fz aligned on target onset (red lines). Thick and thin vertical lines represent, respectively, the mean of response times  $\pm$  standard error of mean. **(B)** Average activity at Fz aligned on breath onset. Dot indicates the time of target presentation of each correct trial (the Y values of each dot represents, its trial position during the session). Trials have been sorted by mean reaction time into five groups. **(C)** Activity at Fz aligned on target onset (same convention as in B). **(D)** Average activation level 10–20 ms before breath initiation is plotted against mean reaction time for the five reaction time groups for this recording session.

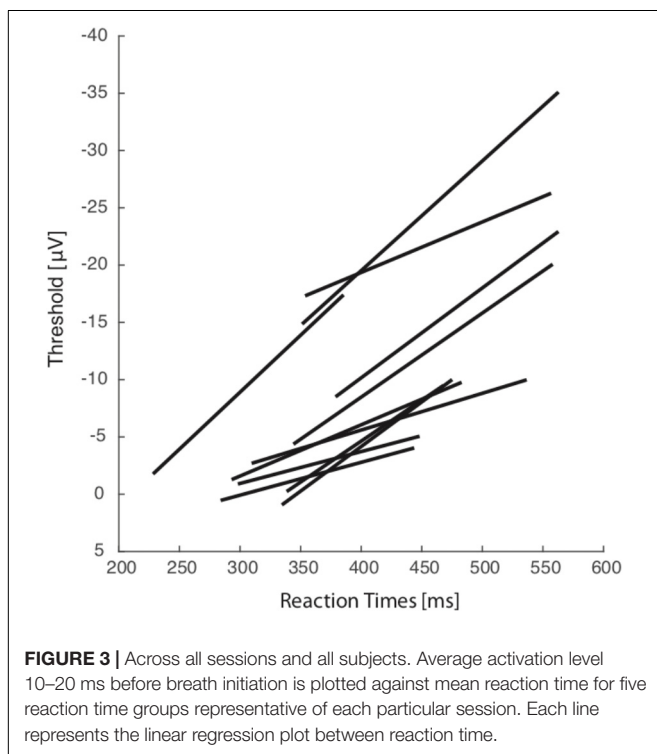


ran *patternsearch* from 50 randomly chosen starting points. Finally, to avoid a local minimum, we again started *patternsearch* from 200 new start points. These starting points were determined by the best parameters from first run and were defined as follows: best parameters  $\pm 0.01$  units and best parameters  $\pm 0.02$  units. All these procedures were performed on a supercomputer cluster (NEC, 40 nodes, 28 cpu Intel Xeon E5-2680 V4 2.4 GHz/node, 128 Go RAM/node).

## RESULTS

**Figure 2** shows the responses from a representative session of 256 NoStop trials. The activity recorded at Fz began to increase  $\sim 100$  ms before breathing initiation, peaking shortly before breathing initiation.

Specific measures of movement-related neural activity were required to evaluate the prediction that the trigger threshold of breathing preparation varied with reaction time. We tested this prediction by measuring the level of neural activation as a function of the time at which the presumed threshold triggering the movement was crossed. On the basis of electrophysiological studies of cortical control in arm, leg, and eye movements, we estimate that measurements of neural activity 10–20 ms before breathing initiation are an accurate index of the level of trigger threshold activation. We defined the threshold activation as the average level of the activation function in the period between 20 and 10 ms before breathing initiation. We compared the activation threshold across groups of NoStop-trials with different reaction times. **Figure 2** presents the activity at Fz for sorted response times aligned on breath onset (top panel)



and target presentation (middle panel). The activation threshold increases for longer reaction times (bottom panel). We divided the distribution of trials into five groups according to their reaction times (Neshige et al., 1988). A linear regression analysis indicated a significant relation between the activation threshold and reaction time. As shown in **Figure 3**, significant changes in activation threshold with reaction time were observed for 11 of the 12 sessions ( $R^2 = 0.78 \pm 0.21$ , mean and standard error, all  $p$ -values were lower than 0.001). The  $p$ -value was computed by transforming the correlation to create a  $t$ -statistic having 3° of freedom. The confidence bounds were based on an asymptotic normal distribution of  $0.5 \cdot \log((1 + r)/(1 - r))$ , with an approximate variance equal to  $1/(N - 3)$ .

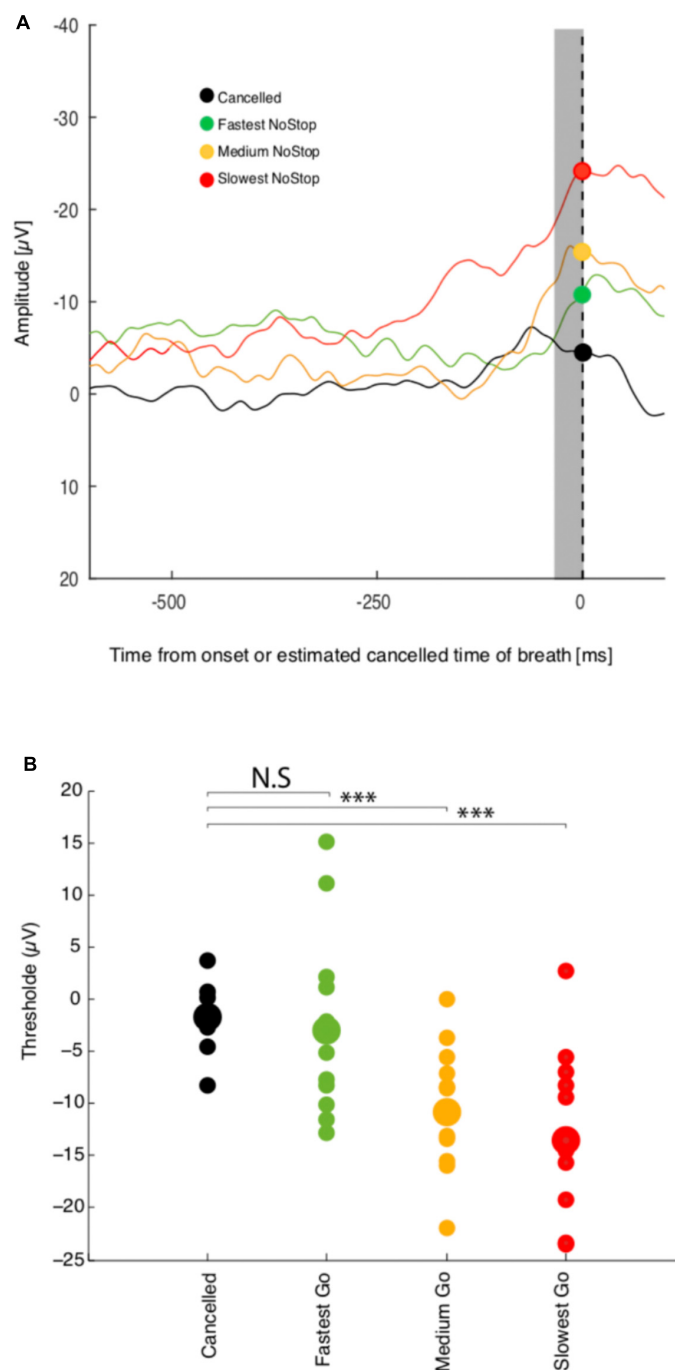
In one session, the impedance value was too high and signal quality was insufficient to be included in the analysis. These results are consistent with variable threshold activity in adjustment of reaction time in breathing.

To determine not only if activation threshold co-varies with reaction time, but if this activity is sufficient to predict whether breathing is going to be initiated or not, we examined the modulation of activity recorded at Fz while the subject was performing a countermanding breathing task (**Figure 1**). **Figure 4A** shows the response of a representative session of 256 non-canceled trials and 256 canceled trials. A linear regression analysis indicates a relation between the threshold activation and reaction time ( $R^2 = 0.48$ ,  $p < 0.001$ ) for this session.

As in the NoStop task, Fz activity in the non-canceled trials recorded began to increase approximately 100 ms before breathing initiation, peaking shortly before it, which we characterize as a failure of inhibition. We compared the activation threshold across groups of non-canceled trials with the activity during canceled trials. The activity in canceled trials remained low and never reached the threshold activity of trials with shorter or longer reaction times as shown in **Figure 4** (top panel). The threshold activation for the group of canceled trials was essentially unchanged compared to baseline activity. Changes in threshold activation with reaction time were observed between subjects: see **Figure 4** (bottom panel). These results are consistent with variable threshold activity and causal linkage between premotor threshold activity and breathing initiation.

To compare response time according to the phase of the spontaneous cycle we divided the response trials into two groups. In the first group, defined as the Inspiratory group, the response times were made in the half cycle that included the initiation of inspiration movement. In the second group, defined as the Expiratory group, the response times were made in the half cycle that included the initiation of expiratory movement. Across sessions and subjects (**Figure 5**), the average mean response times initiated in the half cycle including the initiation of inspiration movement was not significantly different from the response times initiated in the half cycle that included the initiation of expiratory movement (Mann–Whitney  $U$ -test:  $p = 0.79$ ).

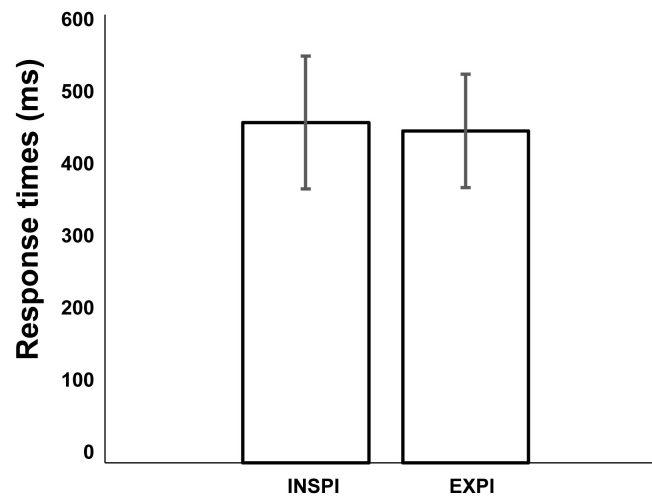
The outputs of data modeling are the best parameters  $\mu$  (mean growth rate),  $\hat{\sigma}$  (noise in simulated signals),  $\beta$  (weight of inhibition of the other unit), and  $D_{\text{stop}}$  (time to encode stop signal) corresponding to the smallest find by minimization procedure.



**FIGURE 4 |** Threshold premotor activity response and breathing inhibition or breathing initiation times. **(A)** Activity at Fz aligned on breath onset. NoStop trials have been sorted by mean reaction time into three groups. Activity recorded during successful Canceled trials are presented by the black curve **(B)** Average activation level 10–20 ms before breath initiation or estimated response time are plotted against mean reaction time for the three reaction time groups for this recording session. \*\*\*Represents significant  $p$  value < 0.001.

Results from behavioral data modeling by interactive race model are summarized in **Table 1** for each subject and session.  $SSRT_s$  represents the Stop Signal Reaction Time (ms) calculated by an integration method (Hanes and Schall, 1995) based on simulated data.

An example of simulated data compared to observed data is shown in **Figure 6** to illustrate the model fitting. This figure represents real data from session 3a and the best parameters from the interactive race model. Panel a) shows the cumulative latencies of NoStop trials; here we can see that these parameters



**FIGURE 5 |** Comparison of mean response times initiated in the half cycle including the initiation of inspiration movement or half cycle including the initiation of expiratory movement (Mann–Whitney *U*-test:  $p = 0.79$ ). Error bars represent the standard deviation of the mean (across 6 subjects and 11 sessions).

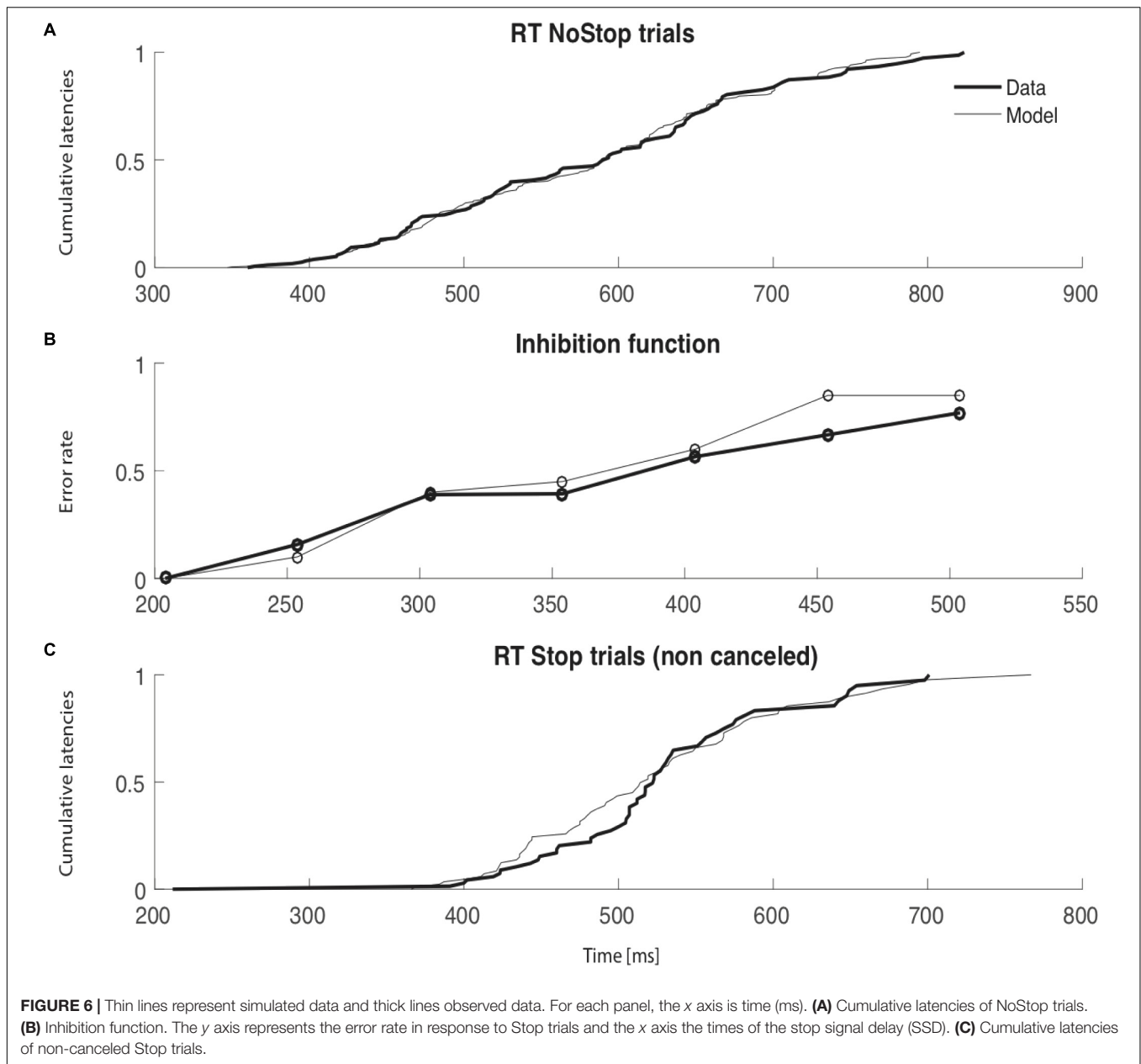
of the model qualitatively fit the real distribution of reaction times. Panel b) shows the inhibition function; the simulated data closely resemble the observed experimental data though when stop signal delay was 450 ms, there was some divergence. Panel c) shows the cumulative distribution of non-canceled Stop trials; again, modeled data closely resemble experimental data, apart from between 400 and 500 ms. In all but one session (session 1a) the initiation of an inhibition curve was observed for SSDs shorter than 200ms (see **Supplementary Figure 1**, for all sessions). In one session (session 1a) the estimation of SSRTs was problematic and therefore require cautious interpretation. Similar outputs and models' mimics exist in the context of countermanding task (modulation of e.g.,  $\mu_{go}$  or DSTOP etc...). Therefor possible combinations of parameters might produce similar estimates of behavioral parameters (Pouget et al., 2011). Simultaneous physiological and behavioral measurement would be required to disentangle such possible discrepancies. In our present context, our models were unsufficiently constraint to be able to conclude. Our modest goal was simply to expose the fact that a already proposed simple model of eye movement control holds with a unique control of breathing initiation.

## DISCUSSION

Our study focused on specific premotor activity potentials in the fronto-medial cortex that increase in relation to voluntary breathing (Macefield and Gandevia, 1991), loaded breathing (Raux et al., 2007), or speech breathing (Tremoureaux et al., 2014b). We first tested whether variability in a single breathing initiation time might be accounted for by modulation of this activity. The results show that voluntary breathing is initiated if, and only if, the cortical premotor potential activity reached a threshold activation level. In the context of countermanding a breathing task, our results show that voluntary breathing is initiated if, and only if, the cortical premotor potential activity reaches a specific threshold activation level. The stochastic variability of this threshold correlates with the distribution of breathing initiation times. Similar premotor potentials have been recorded during voluntary limb, leg and eye movements, with some cerebral potentials being distinguishable in the latter. The major component is a slow negativity, termed the Bereitschaftspotential (readiness potential), which develops before the movement

**TABLE 1 |** Outputs of data modeling for each behavioral session.

Subject/Session	1a	1b	2a	2b	3a	3b	4a	5a	5b	6a
$\mu_{go}$	0.91	1.64	2.71	2.34	1.96	1.4	1.97	2.37	2.31	2.34
$\sigma_{go}$	36.82	46.05	16.51	10.99	11.63	15.53	6.29	23.79	11.96	15.34
$\mu_{stop}$	45.95	17.5	81.72	8.58	66.35	79.69	77.29	81.76	64.68	83.13
$\sigma_{stop}$	50.81	32.22	99.28	81.81	83.16	85.24	54.01	74.41	22.8	97.19
$\beta_{go}$	6.37	9.07	0.51	0.33	0.36	0.35	0.05	0.37	0.13	0.41
$\beta_{stop}$	3.25	3.49	6.39	8.5	5.56	4.63	5.61	5.89	6.48	3.18
$D_{stop}$	93	88	164	152	30	55	161	116	145	126
$\chi^2$	0.06	0.11	0.07	0.27	0.06	0.26	0.12	0.08	0.06	0.25
SSRT <sub>s</sub>	58	206	333	329	196	132	179	236	234	236



(Deecke et al., 1969; Papakostopoulos et al., 1975; Kristeva and Kornhuber, 1980). Based on subdural recordings from the exposed cerebral cortex (Lee et al., 1986; Neshige et al., 1988), and on topographical analysis of scalp recordings (Barrett et al., 1986; Tarkka and Hallett, 1990), this premotor potential is considered to originate in the supplementary motor area and primary motor cortex. This evoked potential does not accompany pathological limb movements generated subcortically (Obeso et al., 1981). Additionally, this potential is present during an array of cortically controlled respiratory movements (see above), but does not accompany involuntary respiratory activity (Macefield and Gandevia, 1991) or abnormal respiratory activity such as hiccups (Raux et al., 2007). To our knowledge this is the first study that demonstrates that

voluntary breathing initiation times might be accounted for by the modulation of cortical premotor potential activity.

Regarding breathing control, several mechanisms are required for adequate speech production. Firstly, as speech is produced during expiration, the control of the duration of vocalized sentences implies inhibition on the automatic breathing pattern generators (for automatic inspiration not to provoke unwanted speech interruptions). In many animals, breath holding during submersion involves strong reflex inhibition of respiratory activity, for example in reaction to snout submersion in the dog (review in Butler, 1982; see also (Lin, 1982). However, natural-diving mammals can perform voluntary apneas spontaneously or in response to training (Ridgway et al., 1969). In humans,



a cortical network capable of substantiating such speech-related inhibitory inputs has been described during voluntary apneas (McKay et al., 2008). Secondly, producing sentences of variable length at a variable loudness implies the ability to prepare these sentences through tailored pre-phonatory breaths. We have previously shown that the corresponding respiratory EEG activities were similar to those involved in the production of voluntary respiratory maneuvers such as sniffing (Tremoureux et al., 2014b), suggesting that some aspects of the speech-related breathing control might derive from a previously selected ability to cortically prepare the volume and timing of particular breaths. In certain species, the ability to prepare inspirations according to locomotor and environmental context appears to be crucial. For example, marine mammals must coordinate inspirations with surfacing, sometimes with important timing constraints, such as during sustained rapid swimming in dolphins. Their ability to voluntarily control breathing for non-respiratory purposes has long been described (Ridgway et al., 1969), e.g., during bubble ring play (McCowan et al., 2000). Thirdly, speech must be fluently adapted to social interactions. Participation in conversation implies the possibility to prepare pre-phonatory breaths, to cue them from various signals, to adapt them to unplanned changes in speech programming, or to completely inhibit them to comply with the necessities of the inter-human exchange. This conversational ability requires excitatory-inhibitory interplays very similar to those described for locomotor and oculomotor movements. Our data suggest that this is indeed the case. Furthermore, the interactive race model has been tested on countermanding eye and hand tasks in previous studies (Papakostopoulos et al., 1975), and our results suggest that this model is also qualitatively effective in breath modality without any further modification.

While analyzing possible interactions between the voluntary and reflexive commands of expiration we did not find any statistical differences between reaction times at different points in the cycle of respiration: reaction times during expiration are not significantly longer than in inspiration. Because many factors could lead to an absence of significant variations our results only indicate that any possible interaction between the voluntary and reflexive command, if it exists, is weak in the context of our countermanding paradigm. In support of these findings, it has been shown that the spinal inspiratory motoneurons are hyperpolarized during expiration compared with inspiration (see for example Berger, 1979). Furthermore, it has been also been shown that the automatic inspiratory drive is sufficient to facilitate the response of the diaphragm to the cortical inputs generated by transcranial magnetic stimulation (Straus et al., 2004; Mehiri et al., 2006). It could therefore have been hypothesized that, in our present experiment, RTs would

have been shorter for voluntary inspirations initiated within the inspiratory phase of the automatic breathing cycle than within its expiratory phase as a consequence of bulbo-spinal facilitation. The fact that this was not the case is coherent with the notion that the expiratory disfacilitation of respiratory motoneurons can be overcome by corticospinal inputs in animals (see for example Planche, 1972) as in humans (Similowski et al., 1996). It also indicates that the excitatory corticospinal drive to inspiratory muscles not only bypasses the brainstem central pattern generators (Corfield et al., 1998) but can also be powerful enough to be beyond modulation by their output (under resting breathing conditions). The high values of  $\mu_{\text{stop}}$  and  $\beta_{\text{stop}}$  observed in most of the model sessions underline the necessity for the inhibition process to be very fast to be effective. These observations suggest that the automatic respiration cycle generated by the brainstem can be overridden by cortical outputs with the same efficiency at any moment of the respiratory cycle and that these cortical outputs can take full precedence over their subcortical counterparts.

## ETHICS STATEMENT

The research was carried out in accordance with the principles outlined in the Declaration of Helsinki. The subjects gave their written informed consent and the study received the ethical and legal approval of the appropriate external body (Comité de Protection des Personnes Paris Ile de France VI).

## AUTHOR CONTRIBUTIONS

EA, TP, and PP designed the experiments. EA and TP performed the experiments. EA, PP, and NW analyzed the data. EA, MR, NW, TS, and PP wrote the manuscript. NW computed model part. All authors reviewed the manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2018.00663/full#supplementary-material>

**FIGURE S1 |** Best interactive model predictions across all subjects. The conventions are the same as those of **Figure 5**. Thin lines represent simulated data, thick lines observed data. For each panel, the x axis is time (ms). **(a)** Cumulative latencies of NoStop trials. **(b)** Inhibition function. The y axis represents the error rate in response to Stop trials and the x axis the times of the stop signal delay (SSD). **(c)** Cumulative latencies of non-canceled Stop trials.

## REFERENCES

- Barrett, G., Shibasaki, H., and Neshige, R. (1986). Cortical potentials preceding voluntary movement: evidence for three periods of preparation in man. *Electroencephalogr. Clin. Neurophysiol.* 63, 327–339. doi: 10.1016/0013-4694(86)90017-9
- Berger, A. J. (1979). Phrenic motoneurons in the cat: subpopulations and nature of respiratory drive potentials. *J. Neurophysiol.* 42(1 Pt 1), 76–90. doi: 10.1152/jn.1979.42.1.76
- Boucher, L., Palmeri, T. J., Logan, G. D., and Schall, J. D. (2007a). Inhibitory control in mind and brain: an interactive race model of countermanding saccades. *Psychol. Rev.* 114, 376–397. doi: 10.1037/0033-295X.114.2.376

- Boucher, L., Stuphorn, V., Logan, G. D., Schall, J. D., and Palmeri, T. J. (2007b). Stopping eye and hand movements: are the processes independent? *Percept. Psychophys.* 69, 785–801. doi: 10.3758/BF03193779
- Butler, P. J. (1982). Respiratory and cardiovascular control during diving in birds and mammals. *J. Exp. Biol.* 100, 195–221.
- Corfield, D. R., Murphy, K., and Guz, A. (1998). Does the motor cortical control of the diaphragm 'bypass' the brain stem respiratory centres in man? *Respir. Physiol.* 114, 109–117. doi: 10.1016/S0034-5687(98)00083-8
- Deecke, L., Scheid, P., and Kornhuber, H. H. (1969). Distribution of readiness potential, pre-motion positivity, and motor potential of the human cerebral cortex preceding voluntary finger movements. *Exp. Brain Res.* 7, 158–168. doi: 10.1007/BF00235441
- Dubois, M., Chenivesse, C., Raux, M., Morales-Robles, A., Nierat, M. C., Garcia, G., et al. (2016). Neurophysiological evidence for a cortical contribution to the wakefulness-related drive to breathe explaining hypocapnia-resistant ventilation in humans. *J. Neurosci.* 36, 10673–10682. doi: 10.1523/JNEUROSCI.2376-16.2016
- Feldman, J. L., Del Negro, C. A., and Gray, P. A. (2013). Understanding the rhythm of breathing: so near, yet so far. *Annu. Rev. Physiol.* 75, 423–452. doi: 10.1146/annurev-physiol-040510-130049
- Gandevia, S. C., and Rothwell, J. C. (1987). Activation of the human diaphragm from the motor cortex. *J. Physiol.* 384, 109–118. doi: 10.1113/jphysiol.1987.sp016445
- Georges, M., Morawiec, E., Raux, M., Gonzalez-Bermejo, J., Pradat, P. F., Similowski, T., et al. (2016). Cortical drive to breathe in amyotrophic lateral sclerosis: a dyspnoea-worsening defence? *Eur. Respir. J.* 47, 1818–1828. doi: 10.1183/13993003.01686-2015
- Hanes, D. P., and Schall, J. D. (1995). Countermanding saccades in macaque. *Vis. Neurosci.* 12, 929–937. doi: 10.1017/S0952523800009482
- Heywood, P., Murphy, K., Corfield, D. R., Morrell, M. J., Howard, R. S., and Guz, A. (1996). Control of breathing in man; insights from the 'locked-in' syndrome. *Respir. Physiol.* 106, 13–20. doi: 10.1016/0034-5687(96)00060-6
- Kristeva, R., and Kornhuber, H. H. (1980). Cerebral potentials related to the smallest human finger movement. *Prog. Brain Res.* 54, 178–182. doi: 10.1016/S0079-6123(08)61622-2
- Launois, C., Attali, V., Georges, M., Raux, M., Morawiec, E., Rivals, I., et al. (2015). Cortical drive to breathe during wakefulness in patients with obstructive sleep apnea syndrome. *Sleep* 38, 1743–1749. doi: 10.5665/sleep.5156
- Lee, B. I., Lüders, H., Lesser, R. P., Dinner, D. S., and Morris, H. H. III (1986). Cortical potentials related to voluntary and passive finger movements recorded from subdural electrodes in humans. *Ann. Neurol.* 20, 32–37. doi: 10.1002/ana.410200107
- Lin, Y. C. (1982). Breath-hold diving in terrestrial mammals. *Exerc. Sport Sci. Rev.* 10, 270–307. doi: 10.1249/00003677-198201000-00009
- Logan, G. D., Cowan, W. B., and Davis, K. A. (1984). On the ability to inhibit simple and choice reaction time responses: a model and a method. *J. Exp. Psychol. Hum. Percept. Perform.* 10, 276–291. doi: 10.1037/0096-1523.10.2.276
- Macefield, G., and Gandevia, S. C. (1991). The cortical drive to human respiratory muscles in the awake state assessed by premotor cerebral potentials. *J. Physiol.* 439, 545–558. doi: 10.1113/jphysiol.1991.sp018681
- McCowan, B., Marino, L., Vance, E., Walke, L., and Reiss, D. (2000). Bubble ring play of bottlenose dolphins (*Tursiops truncatus*): implications for cognition. *J. Comp. Psychol.* 114, 98–106. doi: 10.1037/0735-7036.114.1.98
- McKay, L. C., Adams, L., Frackowiak, R. S., and Corfield, D. R. (2008). A bilateral cortico-bulbar network associated with breath holding in humans, determined by functional magnetic resonance imaging. *Neuroimage* 40, 1824–1832. doi: 10.1016/j.neuroimage.2008.01.058
- Mead, J., Peterson, N., Grimby, G., and Mead, J. (1967). Pulmonary ventilation measured from body surface movements. *Science* 156, 1383–1384. doi: 10.1126/science.156.3780.1383
- Mehiri, S., Straus, C., Arnulf, I., Attali, V., Zelter, M., Derenne, J. P., et al. (2006). Responses of the diaphragm to transcranial magnetic stimulation during wake and sleep in humans. *Respir. Physiol. Neurobiol.* 154, 406–418. doi: 10.1016/j.resp.2005.12.003
- Neshige, R., Lüders, H., and Shibasaki, H. (1988). Recording of movement-related potentials from scalp and cortex in man. *Brain* 111(Pt 3), 719–736. doi: 10.1097/00004691-198804000-00036
- Obeso, J. A., Rothwell, J. C., and Marsden, C. D. (1981). Simple tics in Gilles de la Tourette's syndrome are not prefaced by a normal premovement EEG potential. *J. Neurol. Neurosurg. Psychiatry* 44, 735–738. doi: 10.1136/jnnp.44.8.735
- Papakostopoulos, D., Cooper, R., and Crow, H. J. (1975). Inhibition of cortical evoked potentials and sensation by self-initiated movement in man. *Nature* 258, 321–324. doi: 10.1038/258321a0
- Planche, D. (1972). Effects of cerebral cortex stimulation on the activity of the phrenic nerve. *J. Physiol.* 64, 31–56.
- Pouget, P., Logan, G. D., Palmeri, T. J., Boucher, L., Paré, M., Schall, J. D. (2011). Neural basis of adaptive response time adjustment during saccade countermanding. *J. Neurosci.* 31, 12604–12612. doi: 10.1523/JNEUROSCI.1868-11.2011
- Raux, M., Straus, C., Redolfi, S., Morelot-Panzini, C., Couturier, A., Hug, F., et al. (2007). Electroencephalographic evidence for pre-motor cortex activation during inspiratory loading in humans. *J. Physiol.* 578(Pt 2), 569–578. doi: 10.1113/jphysiol.2006.120246
- Raux, M., Tyvaert, L., Ferreira, M., Kindler, F., Bardinet, E., Karachi, C., et al. (2013). Functional magnetic resonance imaging suggests automatization of the cortical response to inspiratory threshold loading in humans. *Respir. Physiol. Neurobiol.* 189, 571–580. doi: 10.1016/j.resp.2013.08.005
- Ridgway, S. H., Scronce, B. L., and Kanwisher, J. (1969). Respiration and deep diving in the bottlenose porpoise. *Science* 166, 1651–1654. doi: 10.1126/science.166.3913.1651
- Similowski, T., Straus, C., Coïc, L., and Derenne, J. P. (1996). Facilitation-independent response of the diaphragm to cortical magnetic stimulation. *Am. J. Respir. Crit. Care Med.* 154, 1771–1777. doi: 10.1164/ajrcm.154.6.8970369
- Smith, W. K. (1938). The representation of respiratory movements in the cerebral cortex. *J. Neurophysiol.* 1, 55–68. doi: 10.1152/jn.1938.1.1.55
- Straus, C., Locher, C., Zelter, M., Derenne, J. P., and Similowski, T. (2004). Facilitation of the diaphragm response to transcranial magnetic stimulation by increases in human respiratory drive. *J. Appl. Physiol.* 97, 902–912. doi: 10.1152/japplphysiol.00989.2003
- Tarkka, I. M., and Hallett, M. (1990). Cortical topography of premotor and motor potentials preceding self-paced, voluntary movement of dominant and non-dominant hands. *Electroencephalogr. Clin. Neurophysiol.* 75, 36–43. doi: 10.1016/0013-4694(90)90150-I
- Tremoureaux, L., Raux, M., Hudson, A. L., Ranohavimparany, A., Straus, C., and Similowski, T. (2014a). Does the supplementary motor area keep patients with Ondine's curse syndrome breathing while awake? *PLoS One* 9:e84534. doi: 10.1371/journal.pone.0084534
- Tremoureaux, L., Raux, M., Ranohavimparany, A., Morelot-Panzini, C., Pouget, P., and Similowski, T. (2014b). Electroencephalographic evidence for a respiratory-related cortical activity specific of the preparation of prephonatory breaths. *Respir. Physiol. Neurobiol.* 204, 64–70. doi: 10.1016/j.resp.2014.06.018
- Usher, M., and McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychol. Rev.* 108, 550–592. doi: 10.1037/0033-295X.108.3.550

**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 © 2018 Pouget, Allard, Poitou, Raux, Wattiez and Similowski. 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) and the copyright owner(s) 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.



# Role of Oscillations in Auditory Temporal Processing: A General Model for Temporal Processing of Sensory Information in the Brain?

Andreas Bahmer<sup>1\*</sup> and Daya Shankar Gupta<sup>2</sup>

<sup>1</sup> Comprehensive Hearing Center, ENT Clinic, University of Würzburg, Würzburg, Germany, <sup>2</sup> Biology Department, Camden County College, Gloucester Township, NJ, United States

## OPEN ACCESS

### Edited by:

Kathrin Ohla,  
Institut für Neurowissenschaften und  
Medizin, Forschungszentrum Jülich,  
Germany

### Reviewed by:

Raul Cristian Muresan,  
Transylvanian Institute of  
Neuroscience (TINS), Romania  
Lu Zhang,  
Georgia Institute of Technology,  
United States

### \*Correspondence:

Andreas Bahmer  
bahmer@ukw.de

### Specialty section:

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

**Received:** 04 March 2018

**Accepted:** 12 October 2018

**Published:** 31 October 2018

### Citation:

Bahmer A and Gupta DS (2018) Role  
of Oscillations in Auditory Temporal  
Processing: A General Model for  
Temporal Processing of Sensory  
Information in the Brain?  
Front. Neurosci. 12:793.  
doi: 10.3389/fnins.2018.00793

We review the role of oscillations in the brain and in the auditory system showing that the ability of humans to distinguish changes in pitch can be explained as a precise analysis of temporal information in auditory signals by neural oscillations. The connections between auditory brain stem chopper neurons construct neural oscillators, which discharge spikes at various constant intervals that are integer multiples of 0.4 ms, contributing to the temporal processing of auditory cochlear output. This is subsequently spatially mapped in the inferior colliculus. Electrophysiological measurements of auditory chopper neurons in different species show oscillations with periods which are integer multiples of 0.4 ms. The constant intervals of 0.4 ms can be attributed to the smallest synaptic delay between interconnected simulated chopper neurons. We also note the patterns of similarities between microcircuits in the brain stem and other parts of the brain (e.g., the pallidum, reticular formation, locus coeruleus, oculomotor nuclei, limbic system, amygdala, hippocampus, basal ganglia and substantia nigra), dedicated to the processing of temporal information. Similarities in microcircuits across the brain reflect the importance of one of the key mechanisms in the information processing in the brain, namely the temporal coupling of different neural events via coincidence detection.

**Keywords:** canonical microcircuits, cochlear nucleus, locus coeruleus, limbic system, amygdala, hippocampus, basal ganglia, substantia nigra

## 1. INTRODUCTION

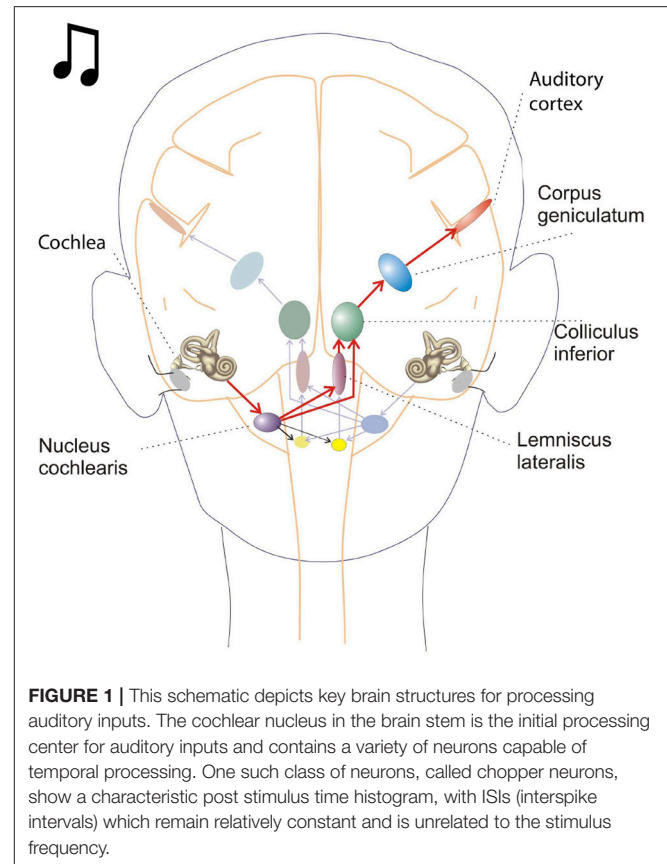
Oscillations are defined as periodic temporal changes in the state parameters of a system and characterize stable states in the non-linear neural dynamics of the brain. The study of oscillations in the human brain began in the early part of the last century, when neural oscillations were recorded by electroencephalography (EEG) in 1924 by Hans Berger at the University of Jena. Neural oscillations in the EEG recordings are classified according to their frequency in different bands. However, EEG signals are only the summed electrical activity of the brain, as they are measured at the surface of the skull. This averaged activity would mask mechanisms, subserved by oscillations in smaller subpopulations of neurons. Furthermore, invasive single unit recording (extracellular and intracellular) as well as the recording of local field potentials reveal the presence of oscillations.

## 2. THE ROLE OF OSCILLATIONS IN COUPLING NEURAL ACTIVITIES IN THE BRAIN

Neural oscillations are observed in various parts of the brain involving different sensory systems, such as the visual, olfactory, motor, and auditory system. In the midbrain, the presence of neural oscillations in electrophysiological recordings in the auditory system was discovered by Langner (1978), which led to the model of auditory temporal processing and neural oscillators by Langner (1981). Later, neural oscillations became a hot topic of research in the visual system. Studies of Gray and Singer (Gray and Singer, 1989; Gray, 1994), and others (Eckhorn et al., 1988) linked oscillations in the visual system to the binding of various percepts. It was shown that neural oscillations, resulting from the synchronization of spatially segregated retinal ganglion cells evoked by stationary and moving visual stimuli, are reliably transmitted by the lateral geniculate nuclei, which suggests the importance of maintaining temporal coupling of neural activities in processing the perception of global stimulus properties such as size and continuity of spatial features (Neuenschwander and Singer, 1996). The temporal coupling of peripheral neural activities between adjacent retinal ganglion cells is due to the presence of intercellular gap junctions (Roy et al., 2017). Studies of olfactory responses has also revealed temporal coupling of neuronal activities. Gilles Laurent and his colleagues observed that during an oscillatory response to odor in locusts, different neurons in the olfactory antennal lobe showed a higher probability of coincidental firing in a pair of neurons in some cycles but not in other cycles of the oscillatory response (Wehr and Laurent, 1996). Furthermore, neural oscillations play a pivotal role in various timing functions of the brain, including time perception (Buhusi and Meck, 2005; Gupta, 2014). In a recent study, recordings from the medial prefrontal cortex in monkeys, who produced different time-intervals using hand or eye movements, showed that the firing rate profiles were temporally scaled to match the produced intervals (Wang et al., 2018). This finding could be explained by the differences in the activation profiles of temporally-coupled subsets of neurons during the production of short and long intervals. Moreover, this study is consistent with the idea that the time course of the temporal coupling of neurons is responsible in part for the conscious time-interval production, while the scaling of the time course is correlated to the length of produced intervals.

## 3. OSCILLATIONS IN THE AUDITORY BRAIN STEM AS A TEMPORAL SCALE

Oscillations in the auditory pathways are observed in the cochlear nucleus and the inferior colliculus (Figure 1) among others. These oscillations are attributed to a class of neurons in the cochlear nucleus, called “chopper neurons” (see e.g., Blackburn and Sachs, 1989). Chopper neurons, which exhibit a unique response pattern, project to the inferior colliculus. They generate oscillations with a frequency, which is relatively independent of the changes of important stimulus parameters



(Pfeiffer, 1966; Blackburn and Sachs, 1989; Wiegand and Winter, 2001; Winter et al., 2001). The interspike interval (ISI) of chopper neurons exhibit a distribution pattern in different species, which is centered at integer multiples of 0.4 ms (Langner and Schreiner, 1988; Bahmer and Langner, 2006a). In Mandarin, a tonal language wherein word meanings change with the pitch, periods, which are integer multiples of 0.4 ms can be found in statistically preferred tones (Langner, 2015). Recently, the temporal constant of 0.4 ms was found in electrophysiological recordings of the cochlear nucleus in human auditory brain stem implant patients (Bahmer et al., 2017). Chopper neurons play a key role in pitch perception (Langner, 1981; Hewitt et al., 1992; Wiegand and Winter, 2001). Incoming acoustical stimuli contain information about the pitch in their temporal modulation. Information about the temporal modulation is transferred via the auditory nerve to the ascending auditory pathways. The tuning of the auditory nerve fibers alone is not sufficient to explain the precision with which humans distinguish between pitch differences (just noticeable differences are about 0.2%, Fastl and Weinberger, 1981). Therefore, in addition to the coarse spectral analysis of the incoming signals in the cochlea, a subsequent temporal analysis is mandatory. Especially for absolute listeners, an inherent scale (neural oscillations in clock mechanism) could explain their outstanding ability to determine absolute pitch. Candidates producing such scales would be the chopper neurons in the cochlear nucleus of



the brain stem. Chopper neurons have a significant role in the periodicity analyzing model introduced by Langner (1981, 1983) including the cochlear nucleus, inferior colliculus, and lemniscus lateralis. According to this model, a neuronal network including different types of oscillators (**Figures 2, 4**) correlates features of the input signal to each other or correlates the features of the input signal to neuronal oscillations. In both modes, chopper neurons provide the temporal scale (Oscillator circuit 1, **Figure 2**). The function of the network is based upon the correlation of undelayed (oscillator circuit 1) and delayed neuronal responses (oscillator circuit 2) of the depicted neurons (**Figure 2**) to envelopes of amplitude modulated (AM) signals. These responses converge at neurons acting as coincidence detectors. Each modulation period of an AM signal activates the trigger neuron, which in turn activates a rapid oscillation (oscillator potential with a predefined frequency). Via parallel processing, the integrator neuron responds to the same cycle of the modulation frequency but with a longer delay which corresponds to the integration period from the integrator-like function. Moreover, the coincidence neuron will be activated, despite different delay intervals of the two previous units, provided that the integration period equals the period of the AM signal. A coincidence neuron will respond more often, when its inputs are synchronized, i.e., when the spikes of the oscillator and of the integrator converge synchronously. Thus, modulation periods (periodicity;  $\tau_m$ ),  $m \times \tau_m$ , with  $m = 1, 2, \dots$ , which activate the oscillations and drive the coincidence unit, can be computed according to the following linear equation:

$$m \times \tau_m = n \times \tau_c - k \times \tau_k \quad (1)$$

where  $k, m, n$  are small integers.  $n \times \tau_c$  is the integration period, which consists of  $n$  carrier periods and after this interval the integrated input signal reaches a threshold.  $1/\tau_c$  is the carrier frequency of the AM signal,  $1/\tau_k$  the frequency of the auditory oscillations. Equation (1) will be referred here as coincidence equation. The parameter  $m$  takes into account the fact that coincidence neurons respond also to harmonics ( $m > 1$ ) of the modulation frequency of the AM signal, which implicates ambiguity of IC neurons with respect to harmonically related signals. A solution to this problem is proposed by an input from the inhibitor (anatomically attributed to the lemniscus lateralis, a spiral structure). Because of the cochlear frequency analysis, neurons respond strongest at a characteristic frequency (CF). In addition to the CF, the coincidence neuron is tuned to a certain periodicity, i.e., a certain modulation frequency of an AM signal, also called the best modulation frequency (BMF). Therefore, different trigger, oscillator, integrator, and coincidence units are incorporated to explain the range of periodicity of AM signals (Langner, 2015). A detailed simulation of the periodicity analyzing model introduced by Langner (1981, 1983) can be found in Borst et al. (2004) and Voutsas et al. (2005). An example of the simulation results of the periodicity model with and without inhibition is depicted in **Figure 3**.

For the peripheral auditory system, ISIs of neural oscillations are argued to serve as a temporal scale (Bahmer and Langner, 2005). Absolute listeners may use this temporal scale for their

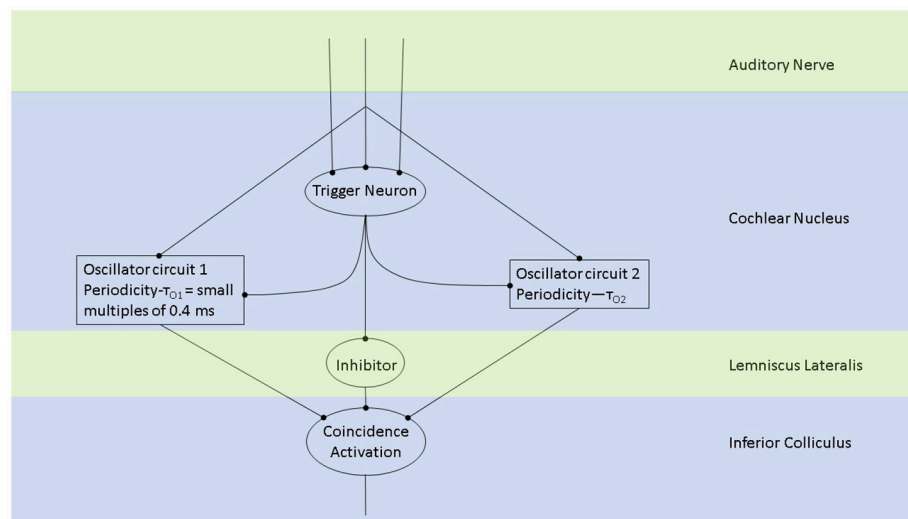
outstanding ability to determine absolute pitch of the incoming tonal acoustic signals.

In a work presented here, we show that by simulating chopper neurons with various oscillation frequencies these neurons may serve a scale for a subsequent temporal analysis as for pitch determination. Furthermore, we hypothesize that microcircuits found in the auditory system which are dedicated to temporal analysis are ubiquitous in the brain for an operation in the temporal domain.

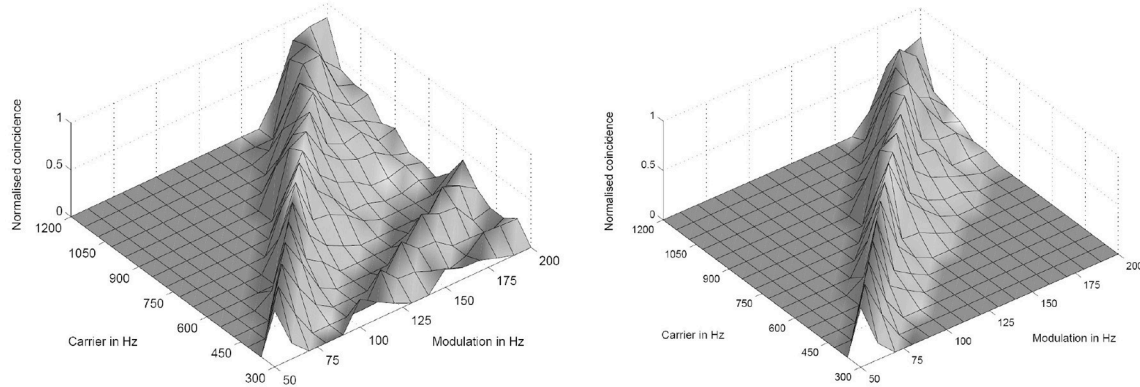
## 4. NEURONAL MODELING OF OSCILLATION IN THE AUDITORY BRAIN STEM

The simulation of the oscillatory neuronal network in the auditory brain stem from Bahmer and Langner (2006b) are performed in Matlab 2006 (The MathWorks, Inc., Natick) and NEURON (Hines and Carnevale, 1997). The differential equations are numerically realized by the Euler method in Matlab. Time steps of  $25 \mu\text{s}$  are sufficient for the relevant time scales of about 0.1 ms. Signal, onset neuron, and chopper neurons are implemented as script-files, and auditory nerve fiber response is calculated within a mex-file in Matlab. Programs were executed on a PC with 2.0 GHz and 512 MB RAM.

The inner ear, inner stereociliary hair cells and auditory nerve fibers were modeled according to Hemmert et al. (2003). A wave-digital filter model describes the vibrations of the basilar membrane on the basis of the passive inner ear hydrodynamics; it consists of 125 mass-spring resonators that are connected by a coupling-mass (Strube, 1985; Zwicker, 1986). To simulate the outer hair cell function, the amplitude of the vibration of the basilar membrane is amplified and the traveling-wave along the basilar membrane is sharpened at the low values of the amplitude. This is performed by the second order resonators that are added at the outputs of the cochlear filter bank. The quality factors of the resonators are altered in all iteration steps depending on the displacement of each resonator. Four stages of the resonators are cascaded to achieve physiologically plausible amplification and filter shapes. Bundles of stereocilia of sensory hair cells are deflected by fluid motion from the movements of the basilar membrane (Mountain and Cody, 1999). When bundles of stereocilia are deflected, ion channels open and  $K^+$ -ions diffuse into the sensory hair cells. The  $K^+$ -ion diffusion depolarizes the inner hair cell membrane. Due to the depolarization,  $Ca^{2+}$ -ions enter the cell through voltage activated Ca-channels. High  $Ca^{2+}$ -concentration within the cell leads to the fusion of synaptic vesicles with the cell membrane (Moser and Beutner, 2000; Beutner et al., 2001). Specific quanta of neurotransmitter release are required to trigger the action potential at the postsynaptic membrane. Since there is a depletion of vesicles with release, spiking probability of the auditory nerve diminishes after a strong stimulus (adaptation). The model also includes a refractory period of about 1 ms (Carney, 1993). The generation of the action potential is a stochastic process due to the implemented random vesicle fusion. A single inner hair cell is connected to 20 synapses of the auditory nerve. Physiological



**FIGURE 2 |** The periodicity analyzing neural model from Langner (2015) consists of two oscillators driven by the incoming auditory signal (Langner, 1981, 1983). Trigger neurons shown in this schematic are frequency-specific t-stellate cells/ chopper neurons, which give rise to tonic firing in response to phasic firing by the auditory nerve (Oertel et al., 2011). Successful periodicity analysis in the brain stem will result when coincidence activation takes place. Coincidence activation, projecting to the auditory cortex, can also account for sparse coding in the auditory cortex while there is tonic/phasic firing at the level of the brain stem and periphery.

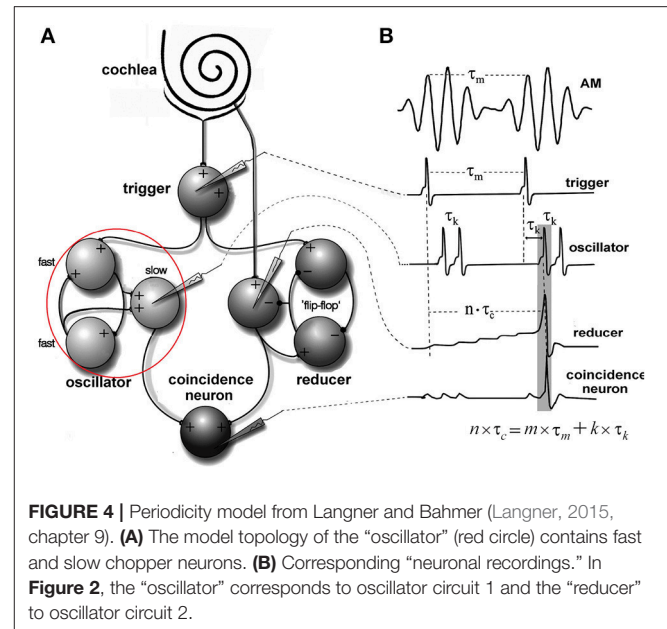


**FIGURE 3 |** Simulation results of the periodicity model without (Left) and with (Right) an inhibitory connection. The response results from 16 periodicity models tuned to characteristic frequencies (CF) and best modulation frequencies (BMF) with the ratio 6:1 (CF/BMF). Stimuli are 256 combinations of 16 carrier and 16 modulation frequencies. The carrier frequency axes corresponds to the CF of one periodicity model (from Voutsas et al., 2005).

and anatomical findings have led to the following simulation paradigm (Figure 4). (A) Two or three chopper neurons (fast) which are connected, can activate its subsequent neighbor, operate as a pace-maker, and project to other chopper neurons (slow) that have a longer refractory period. The fast neurons act as a pace-maker with a clock-rate of 0.4 ms. The slower chopper neurons which, due to longer refractory periods, skip short intervals while producing outputs at the long intervals, which are multiples of 0.4 ms. This reduces the number of the chopper neurons that are required to produce ISIs longer than 0.8 ms. (B) The first of two additional inputs are transmitted via five synapses from the auditory nerve fibers (Ferragamo et al., 1998a). (C) The additional input comes from the onset neuron and activates only one of the chopper neurons in the

circuit. The onset neuron (trigger) receives its broadband input from the auditory nerve and excites one chopper neuron (fast). Inputs from the auditory nerve depolarize the membrane of the chopper neurons. This change in the membrane voltage enables chopping but does not initiate it. The reason is that the weights of auditory nerve synapses are adjusted in such a way that the auditory nerve input alone cannot drive the membrane voltage to the threshold. Instead, the chopping is initialized by a spike from the trigger/onset neuron. The onset neuron is a simplified version of the model that was proposed by Rothman and Manis (2003) and is based on Hodgkin-Huxley (HH) equations. The model consists of a sodium ( $I_{Na}$ ), a low-threshold potassium ( $I_{LTK}$ ), an excitatory synaptic ( $I_E$ ) and a leakage ( $I_{lk}$ ) current. The low threshold of the potassium channel opening is responsible

for the onset neuron behavior (Rothman and Manis, 2003). Simulation parameters of the adapted HH-like onset neuron can be found in Bahmer and Langner (2006b). Chopper neurons are modeled as leaky integrate-and-fire neurons with synapses (Bleeck, 2000). The synapses are modeled as follows. The action potential in the presynaptic neuron leads to the fusion of vesicles, discharging neurotransmitters into the synaptic cleft. The emission of vesicles is simulated by use of a look-up table. The neurotransmitter molecules traveling in the cleft to the postsynaptic neuron is modeled by diffusion. The decay of neurotransmitter effect is simulated by a leaky integrator. The probability of open channels for certain ions increases as the concentration of neurotransmitter in the synaptic cleft becomes higher. Various ions produce either excitatory or inhibitory postsynaptic currents. A hyperbolic tangent function controls the channel conductance. A time delay with adjustable jitter (parameters: mean and standard deviation) that stands for the overall neurotransmitter diffusion time was integrated in the simulation. Details like the neuron and synapse model equations and simulation parameters can be found in Bahmer and Langner (2006b). The simulation of chopper neuron soma activity is based on a leaky integrate-and-fire model. The incoming postsynaptic currents from the synaptic inputs are integrated and build up a postsynaptic potential while a leakage current diminishes the input. When the potential reaches a predefined threshold, a spike is elicited, and the membrane potential is reset. The absolute and relative refractory period (exponentially decreasing) ensures that the spike generation is suppressed or needs a stronger input, respectively, for a given period of time. The time constant of the fast chopper neurons in the simulation is set to 0.8 ms to ensure a fast chopping; whereas the time constants of the slow chopper neurons is set to higher values according to their low chopping frequencies. The summed weight of the synapses of the nerve is on average eight times lower in the simulations than the weights of the synapses of the chopper and onset neuron. Excitatory postsynaptic potentials lead to the subthreshold depolarization of the membrane to enable chopping. This weak auditory nerve input does not mean that the overall response of the chopper neuron is low because the input from the network also contributes to the response. As an alternative to the leaky integrate-and-fire chopper neuron model described in the previous section, the HH-like chopper model of Rothman and Manis (2003) for the simulation environment NEURON was simulated (Bahmer and Langner, 2010). According to the results, the model has the disadvantage that it cannot reproduce *in vivo* data of subpopulations of chopper neurons showing small ISIs (e.g., 1.4 ms, Young et al., 1988). Moreover, the dynamic range of the spike rate of real chopper neurons is about 200–300 spikes/s in average (Frisina et al., 1990). If this physiologically dynamic range is applied to the simulation, the corresponding ISIs in the simulation span a range of about 5–23 ms, whereas *in vivo* values of ISIs differ much less with varying levels (e.g., Frisina et al., 1990). Therefore, the model was adapted by means of genetic algorithms (Bahmer and Langner, 2010) which resulted in cell parameters in a physiologically plausible range. For the simulation of the modified model, the currents are varied in NEURON and the corresponding voltage responses are saved.



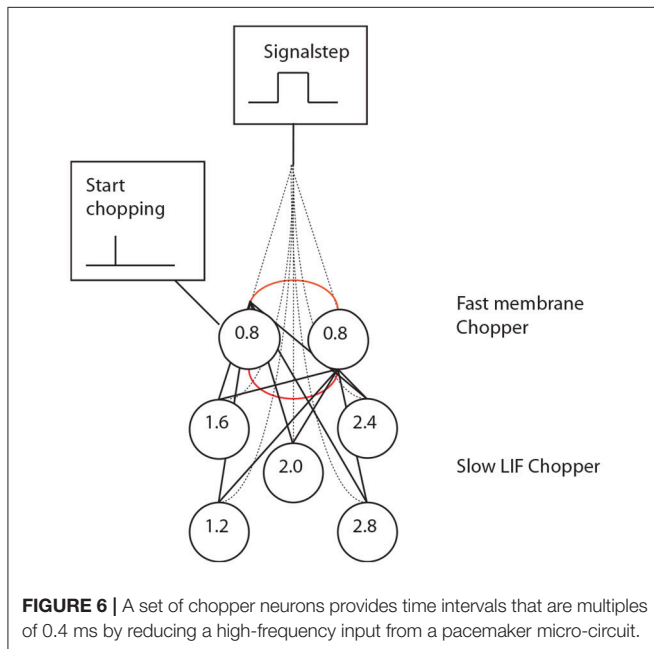
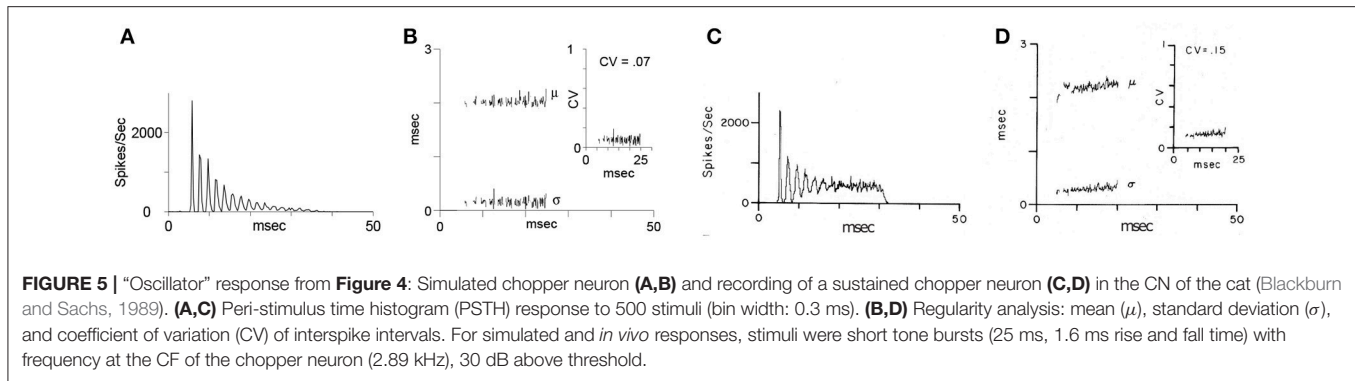
The voltage responses were then analyzed in Matlab and the ISIs were plotted versus the input strength. For the neuronal modeling II, the auditory nerve input is modeled as a signal step and the onset neuron is modeled as a single-spike generator.

## 5. SIMULATION OF A SMALL NETWORK OF FAST PACEMAKER NEURONS IN THE AUDITORY SYSTEM

Blackburn and Sachs (1989) classified (anterior ventral) cochlear nucleus neurons using regularity analysis of ISIs. Important parameters of this analysis were mean and standard deviation. The coefficient of variation value (CV, ratio: standard deviation to the mean of ISIs) enables a comparison of different units of chopper neurons and different stimulus levels. The CV is computed as a function of time. Sustained chopper neurons are a subtype of chopper neurons and classified by a small CV, indicating their highly regular ISIs. **Figure 5** shows the simulation results of the multi-oscillator and physiological data of a sustained chopper neuron in the CN (Bahmer, 2007). Firing rate and ratio of peak heights match their known physiological properties. The data obtained after the simulation, such as firing rate, number of peaks, and ratio of peak heights are similar to electrophysiological data. Even the regularity analysis could be matched to *i* data. In this simulation, a jitter (standard deviation 0.26 ms) is added to the synaptic delay of the interconnections of the fast chopper neurons.

## 6. SIMULATION OF A SMALL NETWORK OF SLOW PACEMAKER NEURONS IN THE AUDITORY SYSTEM

Simulation with the adapted model (**Figure 6**, see also Bahmer, 2007) shows oscillations with ISIs of 0.8 ms duration. Two of



these adapted neurons can mutually excite each other and act as pacemaker. This pacemaker projects to other chopper neurons that have slower time constants and, therefore, skip a certain number of spikes. Nevertheless, the skipping results in ISIs with are integer multiples of 0.4 ms (**Figure 6**). In the simulation, the post synaptic current (**Figure 7** left, PSC) drives the membrane voltage of the slow chopper neuron to the threshold but due to the refractory period several supra-threshold inputs are skipped. Only action potentials at every third supra-threshold input are elicited. Thus, action potentials are only elicited at every third supra-threshold input (ISI: 1.2 ms). For a set of slow chopper neurons, action potentials with various ISI (integer multiples of 0.4 ms) are generated which depends on the refractory period. Note that the refractory period is not necessarily an integer multiple of 0.4 ms, but is a continuous variable; however, ISIs are integer multiples of 0.4 ms, corresponding to the periodic inputs from the fast chopper neurons.

As it can be noted from equation 1, the solution for the correlation of the integration period of the carrier, the

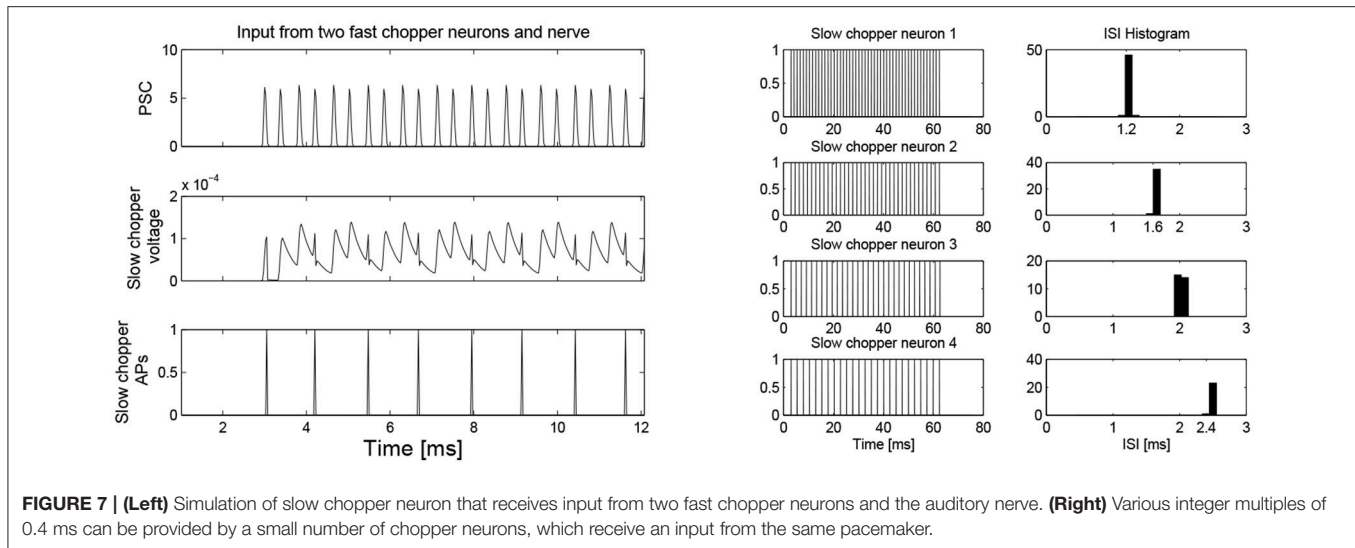
modulation frequency, and frequency of auditory oscillations is constrained by integer values of  $m$ ,  $k$  and  $n$ . The integer values of  $m$ ,  $k$  and  $n$  would represent the number of oscillations, which are reached in respective circuits before integration, a correlate of perception occurs. In fact, as discussed later, circuit patterns found in the auditory system for an effective analysis of high temporal informational content can be found throughout the entire brain (Oertel and Young, 2004; Langner, 2015). We review literature, which shows that many microcircuits, which employ coincidence detection mechanism to temporally couple neural events, are found across the brain.

## 7. INHIBITION OF THE SELF-EXCITING OSCILLATOR MICROCIRCUIT IN THE AUDITORY BRAIN STEM

The simulation of a cluster of chopper neurons shows that oscillations with precise ISIs can be generated with the help of a few neurons. Two or three interconnected fast chopper neurons act as a pacemaker with a smallest temporal resolution of 0.4 ms projecting to the slow chopper neurons. The slow neurons can skip supra-threshold inputs and generate outputs at longer ISIs. In physiological measurement ISIs span a wide range of durations (Young et al., 1988). In the simulation from Bahmer and Langner (2006b), chopper neurons can excite each other as observed in T-stellate cells (Ferragamo et al., 1998b). T-stellate cells also receive an inhibitory input from D-stellate cells. This input, in the presence of the input from the auditory nerve, can inhibit the self-excitation of the network. In the simulation from Bahmer and Langner (2006b), the offset at the end of the input from the auditory nerve was sufficient to stop the excitation of the network. In a future version, the input from D-stellate cells shall be included as excitation must be balanced by inhibition especially if the network contains more interconnected chopper neurons. Furthermore, a combination of inhibitory and excitatory inputs enhances the signal detection and provides means of gain control by reducing noise by inhibition (Casparly et al., 1994; Josephson and Morest, 1998).

For the fast chopper neurons, this input enables chopping; it is a condition for starting and stopping the chopper neurons and is necessary in a self-exciting network (Bahmer and Langner, 2007). But, in the context of the current model, this does not





seem to be necessary for the slow chopper neurons because this functional role is substituted by the projection of the fast chopper neurons. On the other hand, if an additional inhibition of chopper neurons is included (see above: functional role of inhibition of D-stellate cells) this input again seems reasonable. However, if the inhibition is strong enough to mute the circuit, the onset neuron would not activate the chopper neurons. With the help of the excitatory inputs from the auditory nerve, the inhibition is balanced, and the onset neuron is able to activate the chopper neurons. Moreover, the integration of inhibition in this model can plausibly enhance dynamic processing (Eguia et al., 2010).

## 8. TRANSFORMATION OF INCOMING AUDITORY INFORMATION INTO A SPARSE CODE

Psychoacoustical studies in the past have indicated that the perception of speech is not adequately accounted by place frequency mechanisms (Rosen, 1992). The temporal information represented in sounds is also important in the perception of speech (Rosen, 1992). Therefore, it is noteworthy that a recent theoretical work and a growing number of experimental studies indicate that time-dimension is an integral part of information processing underlying various perceptual functions (Gupta, 2014; Gupta and Chen, 2016). Most natural sounds are modulated in amplitude (Joris et al., 2004; Eguia et al., 2010), and, thus, they are represented by two frequencies: a fast frequency, which represents fine oscillations of sound waves and a slow frequency of the amplitude modulation. The oscillations of both frequencies, forming the structure of AM signals of natural sounds processed by cochlea, help to represent physical time-dimension (Gupta, 2014). The spike structure of the AM signals is phase locked to the movements of inner hair cells, which directly results from the pressure changes produced by amplitude-modulated sound waves. Thus, oscillatory structure

of AM signals inputs temporal information into neural circuits when they are processed by trigger neurons (Figure 2). Moreover, this is consistent with the discussion of equation 1, based on the periodicity analyzing model (Langner, 1981, 1983) which suggests that both the carrier frequency of sounds as well as its modulation frequency are responsible for the integration underlying perception. The coincidence detection (Figure 2), responsible for integration would result in a sparse code (Harris et al., 2011), which would be processed in the cortical auditory areas to create the perception of sound.

## 9. COINCIDENCE DETECTION VIA DISTRIBUTED MICROCIRCUITS IS A KEY MECHANISM FOR CONSCIOUS BRAIN FUNCTIONS

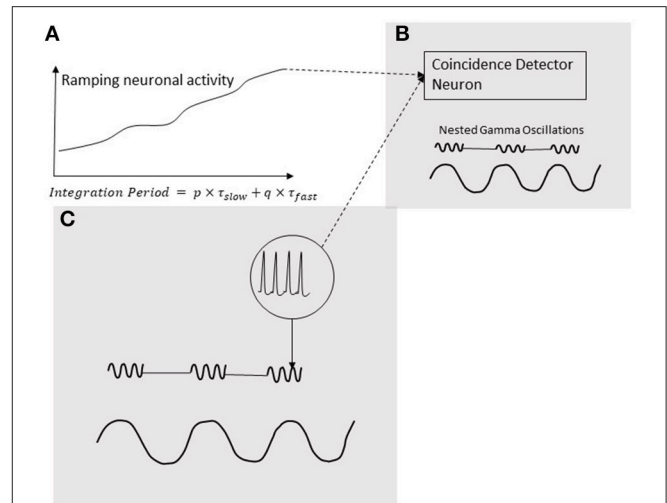
Neural oscillations are hypothesized to play a pivotal role in decoding the temporal information in ramping neuronal activities (Gupta, 2014) that are commonly observed in the cortex (Leon and Shadlen, 2003; Durstewitz, 2004; Lebedev et al., 2008; Schneider and Ghose, 2012; Narayanan, 2016). As discussed in the Introduction, temporal coupling of neural events is important for various cognitive functions of the brain. Moreover, the temporal coupling can be realized by coincidental activation of neural circuits. Furthermore, our models support the role of coincidence detection in the analysis of temporal information in auditory signals. Coincidence detection would play a key role in generating the information that produces a consciously timed behavior. According to the schematic in Figure 8, this information is processed when coincidence detector neuron is stimulated by both, excitatory presynaptic terminals controlled by gamma oscillations (Fries, 2015) as well an increasing excitatory input coming from a ramping neuronal activity. In this mechanism, the ramping activity of neurons resembles an integrator and the oscillators periodicity determine the limit of integration. A coincidence detection

model (**Figure 8**), based on the periodicity analyzing model for auditory signals proposed by Langner (2015), can provide a basis for decoding the information coded by the pattern of ramping activity. As argued by Gupta and Chen (2016), action and perception are temporally coupled by hierarchical neural oscillations. Consistent with this, a coincidence detection of three events is depicted in **Figure 8**. Two of these events are fast-(gamma) oscillations nested in the excitation phase of a slow-oscillation (**Figure 8C**). The third event is the ramping activity of a neuron (**Figure 8A**). The output of the neuron with the ramping activity stimulates the neuron (**Figure 8B**) in the brain area synchronized with the nested oscillation. The neuron in (**Figure 8B**) will be stimulated when ramping activity reaches the threshold, coinciding with the nested gamma oscillations. The time-period from the start of the ramping activity, called Integration Period, will encode the timing of the action.

Cross-frequency coupling allows discrete packets of high (gamma band) frequency oscillations to be formed across larger areas of the brain synchronized by low (alpha and beta bands) frequency oscillations (Buzsáki and Watson, 2012; Gupta and Chen, 2016). The excitatory phase of neural oscillations can increase the probability of coincidental firing of neurons leading to information processing via discrete circuits in a network. Furthermore, according to a leading modern theory of perception, predictive coding, there is an interaction between feedforward and feedback information (Friston, 2008). Cross-frequency coupling would lead to integration by climbing neuronal activities in the cortex during interaction between feedforward and feedback circuits. Experimental evidence and theoretical considerations, reviewed earlier (Bastos et al., 2012), suggest that feedforward connections, predominantly present in the superficial layers of the cortex, use higher frequency oscillation (gamma range), compared to alpha or beta frequency used by feedback connections in the deep cortical layers.

$$\text{Integration Period} = p \times \tau_{\text{slow}} + q \times \tau_{\text{fast}} \quad (2)$$

$\tau_{\text{slow}}$  and  $\tau_{\text{fast}}$  are periodicities of slow- and fast-oscillations, and  $p$  and  $q$  are integers. Ramping activities could also play an important role in the analysis of multiple inputs that underlies a decision process. Single cell recording from layer 5 in the primary motor cortex of rats had shown that there is a strong modulation of specific neuronal activity when there are unfamiliar movements, such as the right or left movements (Cohen and Nicolelis, 2004), which is a suggestive of a decision process. Moreover, the neurons in the cortical layer 5 send axons to the thalamus, basal nuclei, brain stem as well as the spinal cord to control motor movements (Crossmann and Neary, 2010). Since the primary motor cortex receives inputs from the prefrontal cortex and different sensory areas (Borra and Luppino, 2017; Kheradmand and Winnick, 2017), ramping activity may result from a variable balance of inputs from many of these areas, which would be the basis for the decision process.



**FIGURE 8 |** Coincidence detection of three events. Two of these events are gamma oscillations nested in the excitation phase of a low-frequency oscillation (**C**); the third event is a ramping activity of a neuron (**A**). The output of the neuron ramping activity stimulates the neuron (**B**) in the brain area synchronized with the nested oscillation. The neuron in (**B**) will be stimulated when ramping activity reaches a threshold coinciding with the nested gamma oscillations. The time-period from the start the ramping activity, called Integration Period, will encode the timing of the action.

## 10. ANATOMICAL SUBSTRATES FOR CANONICAL MICROCIRCUITS FOR TEMPORAL PROCESSING IN THE BRAIN

The auditory system has evolved by adapting its internal functional structures for a fast processing of incoming signals. As outlined in the Introduction, a periodicity analysis of incoming signals can be accomplished by simple neuronal elements (Langner, 1981). These elements resemble components like integrators, differentiators, and temporal coincidence detectors. Even the occurrence of harmonics in the periodicity analysis—the unwanted side effect of a correlation analysis see **Figure 3**—is suppressed by a helical structure located in the lemniscus lateralis (Ochse, 2004; Voutsas et al., 2005; Langner, 2015). Note that oscillations are ubiquitous in the brain as outlined in the Introduction. However, in contrast to their specific functional role as a temporal scale in the auditory brain stem, they are rather seen as an epiphenomenon in other brain areas, that is, no distinct meaning can be generally attributed to a certain oscillation frequency. Nevertheless, oscillations are a power tool for communication between neuronal networks (Gray and Singer, 1989; Gray, 1994; Fries, 2015). Given that temporal neuronal processing is enhanced by oscillations, it is not surprising to find similar canonical microcircuits in the brain (e.g., the cerebellum-like circuit pattern found in the dorsal cochlear nucleus and pallidum, see Oertel and Young, 2004). There are several parts of the brain that contain helical-like structures after reconstructing from sections, and resolved at the level of cells [**Figure 9**, ventral part of the lemniscus lateralis, locus coeruleus, oculomotor nuclei,

amygdala, hippocampus (cornu ammonis 3), and pars compacta and reticulata of the substantia nigra, Langner (2015)]. These structures provide plausible anatomical solutions for processing hierarchical oscillations as there could be at least two gradients of frequencies in ensembles of neurons: one from periphery to the center and the other between several turns of the helix (Langner, 2015).

## 11. OSCILLATIONS AS A TARGET FOR BRAIN-COMPUTER-INTERFACES

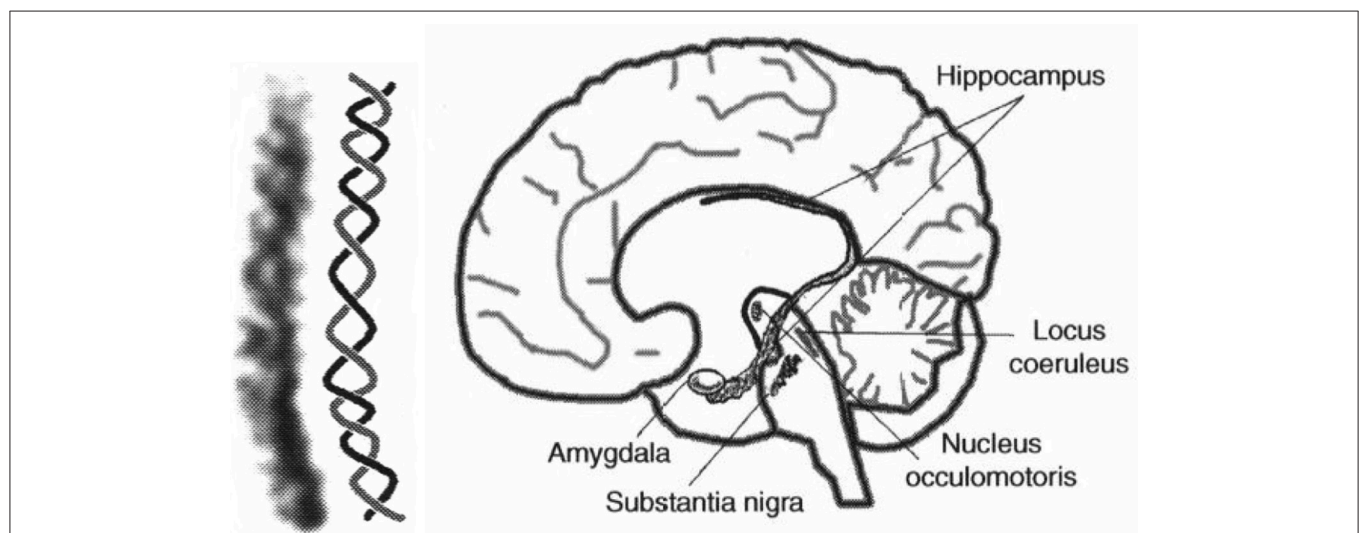
It has always been a vision to interface the brain with a computer to control brain functions. In the auditory system, computer-brain interfaces have already become reality with the development of cochlea, brain stem, and midbrain implants. Cochlea implants stimulate the auditory nerve in the cochlea with electrical impulses, brain stem implants are located in the cochlear nucleus, midbrain implants in the inferior colliculus. These implants are still undergoing further improvements through research, and understanding the role of the oscillations in the cochlear nucleus may be the key to further improvements. In addition, a resonance phenomenon may help to locate target structures for auditory brain stem implants. Ramsden et al. (2016) have postulated the existence of chopper neurons with a preference for certain oscillations periods (Bahmer and Langner, 2006a,b) as a target for electrical stimulation. Based on the idea of targeting certain neuronal networks, strategies have been proposed in electrical stimulation of neuronal networks for cochlear implants, auditory brain stem implants, auditory mid brain implants, as well as for deep brain stimulation (Bahmer et al., 2009; Bahmer, 2016, 2017; Bahmer and Schleich, 2016). These stimulation strategies and alternative pulse shapes (Bahmer et al., 2010; Bahmer and Baumann, 2016) may also

be useful for the deep brain stimulation in psychiatric diseases (Buzsáki and Watson, 2012).

## 12. OSCILLATIONS UNDERLYING AUDITORY STEADY STATE RESPONSES: IMPACT ON SCHIZOPHRENIA AND DEPRESSION

Studies have shown that the perception of sound waves is associated with an increased inter-hemispheric interaction via synchronization long-range gamma bands (Steinmann et al., 2014). Gamma oscillations could play a key role during the long-distance synchronization of local circuits in this inter-hemispheric interaction (Buzsáki and Watson, 2012; Fries, 2015). In each gamma cycle, there is a state of excitation, lasting 3 ms, which triggers an inhibition, lasting for the remainder of the gamma cycle (Fries, 2015). The precision of the 3 ms excitation in the gamma cycle may help to temporally align neural events via long-range gamma band synchronization (Steinmann et al., 2014) in circuits, subserving the perception of sound waves in two hemispheres. Thus, the perception of sounds could be causally related with the temporal coupling in cortical areas, which would result from the coincidence detection events, similar to the processing of auditory signals in the brain stem.

In schizophrenia, which is characterized by the impairment of the perceptual functions, patients often suffer from hallucinations. Thus, it is not surprising that a meta-analytic study finds that in schizophrenia, there is a reduction in the power as well as phase locking values of the 40 Hz gamma-range auditory steady state responses (ASSR) (Thuné et al., 2016). This is consistent with a reduction in the temporal coupling of neural



**FIGURE 9 | (Left)** Example for a helical-like structure in the hippocampus. **(Right)** Helical-like structures can be seen in various locations in the brain, such as the basal ganglia (substantia nigra), reticular formation (locus coeruleus), and limbic system (hippocampus, amygdala), both reproduced from Langner (2015) with permission.

activities, processing sound stimuli in schizophrenia, which would be responsible for the impairments in sound perception, contributing to auditory hallucinations. In addition, ASSR is also affected in bipolar disorder (Rass et al., 2010).

Depression is the most prevalent psychiatric disease (a roughly 20% lifetime incidence in Western populations) and the third largest amongst all illnesses in the world (Mathers et al., 2008). Abnormal differences in oscillations after auditory stimulation have been found between depressed patients versus controls (Iosifescu, 2011). Treatment options are restricted, and the medication success is often based on trial-and-error and a relevant question is whether a particular measure can predict the outcome of the treatment (Buzsáki and Watson, 2012). Interestingly, the loudness-dependence of auditory evoked potentials, can determine the responsiveness to serotonergic versus non-serotonergic antidepressants (Hegerl and Juckel, 1993; Iosifescu, 2011).

### 13. CONCLUSION

In this review, we discuss how temporal information in auditory signals can be accurately analyzed by means of the oscillating activity of chopper neurons in the brain stem. This analysis involves the activation of coincidence neurons, which detects the temporal coupling between the discharges by circuits of chopper neurons with a regular firing pattern, and the integrator neurons with a ramping activity pattern (Figure 4), which would project to the cortex as a sparse code. Moreover, neurons

with ramping activity, resembling the integrator neurons, are commonly found across the cortex. Mechanisms involving coincidence detection neurons, modulated by nested gamma oscillations may contribute to the information processing that decodes the activity of ramping neurons (Figure 8). Additionally, it should be noted that the coincident activation only detects spatiotemporal convergence of neural events; however, primary triggering events may be few milliseconds apart (Fries, 2015). Coincidence detection of neural events, is also likely to form the basis of a variety of perceptions, such as sensations of smell, sound, even the spatial perception of visual objects. As noted above, the impairments of temporal coupling could also contribute partly to the defects of conscious functions in schizophrenia, bipolar disorder, depression, just to name a few. Accordingly, the future investigations of the temporal coupling in the brain may help us develop new treatments of some of the most socially devastating ailments.

### AUTHOR CONTRIBUTIONS

AB and DG conception, revising, final approval, approving of publication and accountable for all aspects of the work. AB simulation, data analysis.

### ACKNOWLEDGMENTS

Parts of this work are published in the thesis of AB (Bahmer, 2007).

### REFERENCES

- Bahmer, A. (2007). *Computer Simulation of Chopper Neurons: Intrinsic Oscillations and Temporal Processing in the Auditory System*. Ph.D., Technical University Darmstadt.
- Bahmer, A. (2016). Temporal coding for hearing implants. US Patent 9,278,213.
- Bahmer, A. (2017). Auditory prosthesis using stimulation rate as a multiple of periodicity of sensed sound. US Patent 9,674,621.
- Bahmer, A., and Baumann, U. (2016). The underlying mechanism of preventing facial nerve stimulation by triphasic pulse stimulation in cochlear implant users assessed with objective measure. *Otol. Neurotol.* 37, 1231–1237. doi: 10.1097/MAO.0000000000001156
- Bahmer, A., Kroneis, T., and Shehata-Dieler, W. (2017). “An intrinsic time constant of the cochlear nucleus revealed by abi electrophysiological measurements,” in *2017 Abstracts of the Conference on implantable auditory prosthesis (CIAP)*, ed J. Arenberg (Lake Tahoe, CA), 73.
- Bahmer, A., and Langner, G. (2005). Modeling intrinsic oscillations in the auditory system: a neuronal mechanism for quantal pitch shifts and absolute pitch? *Ann. N. Y. Acad. Sci.* 1060.
- Bahmer, A., and Langner, G. (2006a). Oscillating neurons in the cochlear nucleus: I. Experimental basis of a simulation paradigm. *Biol. Cybern.* 95, 371–379. doi: 10.1007/s00422-006-0092-6
- Bahmer, A., and Langner, G. (2006b). Oscillating neurons in the cochlear nucleus: II. Simulation results. *Biol. Cybern.* 95, 381–392. doi: 10.1007/s00422-006-0091-7
- Bahmer, A., and Langner, G. (2007). “Simulation of oscillating neurons in the cochlear nucleus: a possible role for neural nets, onset cells, and synaptic delays,” in *Hearing - From Basic Research to Applications*, eds B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, et al. (Heidelberg; New York, NY: Springer), 126–132.
- Bahmer, A., and Langner, G. (2010). Parameters for a model of an oscillating neuronal network in the cochlear nucleus defined by genetic algorithms. *Biol. Cybern.* 102, 81–93. doi: 10.1007/s00422-009-0353-2
- Bahmer, A., Langner, G., Hemmert, W., and Baumann, U. (2009). “Cochlear implant: from theoretical neuroscience to clinical application,” in *Computational Neuroscience Conference Abstract: Bernstein Conference on Computational Neuroscience* (Frankfurt).
- Bahmer, A., Polak, M., and Baumann, U. (2010). Recording of electrically evoked auditory brainstem responses after electrical stimulation with biphasic, triphasic and precision triphasic pulses. *Hear. Res.* 259, 75–85. doi: 10.1016/j.heares.2009.10.003
- Bahmer, A., and Schleich, P. (2016). Auditory prosthesis stimulation rate as a multiple of intrinsic oscillation. EP Patent App. EP20,140,838,130.
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron* 76, 695–711. doi: 10.1016/j.neuron.2012.10.038
- Beutner, D., Voets, T., Neher, E., and Moser, T. (2001). Calcium dependence of exocytosis and endocytosis at the cochlear inner hair cell afferent synapse. *Neuron* 29, 681–690. doi: 10.1016/S0896-6273(01)00243-4
- Blackburn, C., and Sachs, M. (1989). Classification of unit types in the anteroventral cochlear nucleus: Pst histograms and regularity analysis. *J. Neurophysiol.* 62, 1303–1329. doi: 10.1152/jn.1989.62.6.1303
- Bleeck, S. (2000). *Holistische Signalverarbeitung in Einem Modell latenzverknuepfter Neuronen*. Ph.D. thesis, TU Darmstadt.
- Borra, E., and Luppino, G. (2017). Functional anatomy of the macaque temporo-parieto-frontal connectivity. *Cortex* 97, 306–326. doi: 10.1016/j.cortex.2016.12.007
- Borst, M., Langner, G., and Palm, G. (2004). A biologically motivated neural network for phase extraction from complex sounds. *Biol. Cybern.* 90, 98–104. doi: 10.1007/s00422-003-0459-x



- 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
- Buzsáki, G., and Watson, B. O. (2012). Brain rhythms and neural syntax: Inhibitory inputs modulate discharge rate within frequency receptive fields of anteroventral cochlear nucleus neurons. *J. Neurophysiol.* 72, 2124–2133. doi: 10.1152/jn.1994.72.5.2124
- Cohen, D., and Nicolelis, M. A. (2004). Reduction of single-neuron firing uncertainty by cortical ensembles during motor skill learning. *J. Neurosci.* 24, 3574–3582. doi: 10.1523/JNEUROSCI.5361-03.2004
- Crossmann, A. R., and Neary, D. (2010). *Neuroanatomy: An Illustrated Colour Text*. Edinburgh, NY: Churchill Livingstone.
- Durstewitz, B. (2004). Neural representation of interval time. *Neuroreport* 15, 745–749. doi: 10.1097/00001756-200404090-00001
- Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., et al. (1988). Coherent oscillations: a mechanism of feature linking in the visual cortex? *J. Physiol.* 60, 121–130. doi: 10.1007/BF00202899
- Eguia, M. C., Garcia, G. C., and Romano, S. A. (2010). A biophysical model for modulation frequency encoding in the cochlear nucleus. *J. Physiol. Paris* 104, 118–127. doi: 10.1016/j.jphysparis.2009.11.014
- Fastl, H., and Weinberger, M. (1981). Frequency discrimination for pure and complex tones. *Acustica* 49, 77–78.
- Ferragamo, M., Golding, N., and Oertel, D. (1998a). Synaptic inputs to stellate cells in the ventral cochlear nucleus. *J. Neurophysiol.* 79, 51–63. doi: 10.1152/jn.1998.79.1.51
- Ferragamo, M. J., Haresign, T., and Simmons, J. A. (1998b). Frequency tuning, latencies, and responses to frequency-modulated sweeps in the inferior colliculus of the echolocating bat, *Myotisotis fuscus*. *J. Compar. Physiol. A* 182, 65–79. doi: 10.1007/s003590050159
- Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034
- Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (1990). Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. *Hear. Res.* 44, 99–122. doi: 10.1016/0378-5955(90)90074-Y
- Friston, K. (2008). Hierarchical models in the brain. *PLoS Comput. Biol.* 4:e1000211. doi: 10.1371/journal.pcbi.1000211
- Gray, C. M. (1994). Synchronous oscillations in neuronal systems: mechanisms and functions. *J. Comput. Neurosci.* 1, 11–38. doi: 10.1007/BF00962716
- Gray, C. M., and Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 86, 1698–1702. doi: 10.1073/pnas.86.5.1698
- 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. (2016). Brain oscillations in perception, timing and action. *Curr. Opin. Behav. Sci.* 8, 161–166. doi: 10.1016/j.cobeha.2016.02.021
- Harris, K. D., Bartho, P., Chadderton, P., Curto, C., de la Rocha, J., Hollender, L., et al. (2011). How do neurons work together? lessons from auditory cortex. *Hear. Res.* 271, 37–53. doi: 10.1016/j.heares.2010.06.006
- Hegerl, U., and Juckel, G. (1993). Intensity dependence of auditory evoked potentials as an indicator of central serotonergic neurotransmission: a new hypothesis. *Biol. Psychiatry* 33, 173–187. doi: 10.1016/0006-3223(93)90137-3
- Hemmert, W., Holmberg, M., and Gerber, M. (2003). “Coding of auditory information into nerve-action potentials,” in *Fortschritte der Akustik*, ed G. K. Behler (Oldenburg: Deutsche Gesellschaft für Akustik e.V.), 770–771.
- Hewitt, M. J., Meddis, R., and Shakleton, T. M. (1992). A computer model of a cochlear-nucleus stellate cell: responses to amplitude-modulated and pure-tone stimuli. *J. Acoust. Soc. Am.* 91, 2096–2109. doi: 10.1121/1.403696
- Hines, M. L., and Carnevale, N. T. (1997). The neuron simulation environment. *Neural Comp.* 9, 1179–1209. doi: 10.1162/neco.1997.9.6.1179
- Iosifescu, D. V. (2011). Electroencephalography-derived biomarkers of antidepressant response. *Harvard Rev. Psychiatry* 19, 144–154. doi: 10.3109/10673229.2011.586549
- Joris, P. X., Schreiner, C. E., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiol. Rev.* 84, 541–577. doi: 10.1152/physrev.00029.2003
- Josephson, E. M., and Morest, D. K. (1998). A quantitative profile of the synapses on the stellate cell body and axon in the cochlear nucleus of the chinchilla. *J. Neurocyt.* 27, 841–864. doi: 10.1023/A:1006959532686
- Kheradmand, A., and Winnick, A. (2017). Perception of upright: multisensory convergence and the role of temporo-parietal cortex. *Acta Laryngol.* 8:552. doi: 10.3389/fneur.2017.00552
- Langner, G. (1978). The periodicity matrix. A correlation model for central auditory frequency analysis. *Verh Dtsch Zool Ges.* 71:194.
- Langner, G. (1981). Neuronal mechanisms for pitch analysis in the time domain. *Exp. Brain Res.* 44, 450–454. doi: 10.1007/BF00238840
- Langner, G. (1983). Evidence for neuronal periodicity detection in the auditory system of the guinea fowl: Implications for pitch analysis in the time domain. *Exp. Brain Res.* 52, 333–355. doi: 10.1007/BF00238028
- Langner, G. (2015). *The Neural Code of Pitch and Harmony*. Cambridge: Cambridge University Press.
- Langner, G., and Schreiner, C. (1988). Periodicity coding in the inferior colliculus of the cat: I. Neuronal mechanisms. *J. Neurophysiol.* 60, 1799–1822. doi: 10.1152/jn.1988.60.6.1799
- 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
- Mathers, C., Fat, D. M., Organization, W. H., and Boerma, J. T. (2008). *The Global Burden of Disease 2004 Update*. Geneva: World Health Organization.
- Moser, T., and Beutner, D. (2000). Kinetics of exocytosis and endocytosis at the cochlear inner hair cell afferent synapse of the mouse. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11883–11888. doi: 10.1073/pnas.97.2.883
- Mountain, D. C., and Cody, A. R. (1999). Multiple modes of inner hair cell stimulation. *Hear. Res.* 132, 1–14. doi: 10.1016/S0378-5955(99)00013-1
- Narayanan, N. S. (2016). Ramping activity is a cortical mechanism of temporal control of action. *Curr. Opin. Behav. Sci.* 8, 226–230. doi: 10.1016/j.cobeha.2016.02.017
- Neuenschwander, S., and Singer, W. (1996). Long-range synchronization of oscillatory light responses in the cat retina and lateral geniculate nucleus. *Nature* 379, 728–732. doi: 10.1038/379728a0
- Ochse, M. (2004). *Neuronale Kodierung von Tonhöhen und harmonischen Relationen im auditorischen Mittelhirn der Rennmaus (Meriones unguiculatus)*. PhD thesis, TU Darmstadt.
- Oertel, D., Wright, S., Cao, X. J., Ferragamo, M., and Bal, R. K. (2011). The multiple functions of t stellate/multipolar/chopper cells in the ventral cochlear nucleus. *Hear. Res.* 276, 61–69. doi: 10.1016/j.heares.2010.10.018
- Oertel, D., and Young, E. (2004). What a cerebellar circuit doing in the auditory system? *Trends Neurosci.* 27, 104–110. doi: 10.1016/j.tins.2003.12.001
- Pfeiffer, R. R. (1966). Classification of response patterns of spike discharges for units in the cochlear nucleus: tone-burst stimulation. *Exp. Brain Res.* 1, 220–235. doi: 10.1007/BF00234343
- Ramsden, R. T., Freeman, S. R. M., Lloyd, S. K. W., King, A. T., Shi, X., Ward, C. L., et al. (2016). Auditory brainstem implantation in neurofibromatosis type 2: Experience from the manchester programme. *Otol. Neurotol.* 37, 1267–1274. doi: 10.1097/MAO.0000000000001166
- Rass, O., Krishnan, G., Brenner, C. A., Hetrick, W. P., Merrill, C. C., Shekhar, A., et al. (2010). Auditory steady state response in bipolar disorder: relation to clinical state, cognitive performance, medication status, and substance disorders. *Bipolar Disord.* 12, 793–803. doi: 10.1111/j.1399-5618.2010.00871.x
- Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 336, 367–373. doi: 10.1098/rstb.1992.0070
- Rothman, J. S., and Manis, P. B. (2003). The roles potassium currents play in regulating the electrical activity of ventral cochlear nucleus neurons. *J. Neurophysiol.* 89, 3097–3113. doi: 10.1152/jn.00127.2002
- Roy, K., Kumar, S., and Bloofield, S. A. (2017). Gap junctional coupling between retinal amacrine and ganglion cells underlies coherent activity integral to global object perception. *Proc. Natl. Acad. Sci. U.S.A.* 114, E10484–E10493. doi: 10.1073/pnas.1708261114

- Schneider, B. A., and Ghose, G. M. (2012). Temporal production signals in parietal cortex. *PLoS Biol.* 10:e1001413. doi: 10.1371/journal.pbio.1001413
- Steinmann, S., Leicht, G., Erl, M., Andreou, C., Polomac, M., Westerhausen, R., et al. (2014). Conscious auditory perception related to long-range synchrony of gamma oscillations. *Neuroimage* 100, 435–443. doi: 10.1016/j.neuroimage.2014.06.012
- Strube, H. W. (1985). A computationally efficient basilar-membrane model. *Acustica* 58, 207–214.
- Thuné, H., Recasens, M., and Uhlhaas, P. J. (2016). The 40-Hz auditory steady-state response in patients with schizophrenia: a meta-analysis. *JAMA Psychiatry* 73, 1145–1153. doi: 10.1001/jamapsychiatry.2016.2619
- Voutsas, K., Langner, G., Adamy, J., and Ochse, M. (2005). A brain-like neural network for periodicity analysis. *IEEE Trans. Syst. Man. Cybern. B Cybern.* 35, 12–22. doi: 10.1109/TSMCB.2004.837751
- Wang, J., Narain, D., Hosseini, E. A., and Jazayeri, M. (2018). Flexible timing by temporal scaling of cortical responses. *Nat. Neurosci.* 21, 102–110. doi: 10.1038/s41593-017-0028-6
- Wehr, M., and Laurent, G. (1996). Odour encoding by temporal sequences of firing in oscillating neural assemblies. *Nature* 384, 162–166. doi: 10.1038/384162a0
- Wiegrefe, L., and Winter, I. M. (2001). Temporal representation of iterated rippled noise as a function of delay and sound level in the ventral cochlear nucleus. *J. Neurophysiol.* 85, 1206–1219. doi: 10.1152/jn.2001.85.3.1206
- Winter, I. M., Wiegrefe, L., and Patterson, R. D. (2001). The temporal representation of the delay of iterated rippled noise in the ventral cochlear nucleus of the guinea-pig. *J. Physiol.* 537, 553–566. doi: 10.1111/j.1469-7793.2001.00553.x
- Young, E. D., Robert, J. M., and Shofner, W. P. (1988). Regularity and latency of units in ventral cochlear nucleus: implications for unit classification and generation of response properties. *J. Neurophysiol.* 60, 1–29. doi: 10.1152/jn.1988.60.1.1
- Zwicker, E. (1986). A hardware cochlear nonlinear pre-processing model with active feedback. *J. Acoust. Soc. Am.* 80, 146–153. doi: 10.1121/1.394175

**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 © 2018 Bahmer and Gupta. 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) and the copyright owner(s) 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.



# The Temporal Dynamic Relationship Between Attention and Crowding: Electrophysiological Evidence From an Event-Related Potential Study

Chunhua Peng<sup>1,2</sup>, Chunmei Hu<sup>1,2</sup> and Youguo Chen<sup>3\*</sup>

<sup>1</sup> Laboratory of Emotion and Mental Health, Chongqing University of Arts and Sciences, Chongqing, China, <sup>2</sup> Collaborative Innovation Center for Brain Science, Chongqing, China, <sup>3</sup> 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:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

Joseph Charles Schmidt,  
University of Central Florida,  
United States  
Michael Herzog,  
École Polytechnique Fédérale  
de Lausanne, Switzerland

### \*Correspondence:

Youguo Chen  
ygchen246@gmail.com

### Specialty section:

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

**Received:** 06 December 2017

**Accepted:** 29 October 2018

**Published:** 22 November 2018

### Citation:

Peng C, Hu C and Chen Y (2018)  
The Temporal Dynamic Relationship  
Between Attention and Crowding:  
Electrophysiological Evidence From  
an Event-Related Potential Study.  
*Front. Neurosci.* 12:844.  
doi: 10.3389/fnins.2018.00844

Visual crowding is the difficulty experienced in identifying a target flanked by other objects within the peripheral visual field. Despite extensive research conducted on this topic, the precise relationship between attention and crowding is still debatable. One perspective suggests that crowding is a bottom-up and pre-attentive process, while another suggests that crowding is top-down and attentional. A third perspective proposes that crowding is a combination of bottom-up and top-down processes. To address this debate, the current study manipulated the attention and distance between targets and flankers, while simultaneously measuring event-related potentials, in human participants. Results indicated that, compared to uncrowded targets, crowded targets elicited more negative frontal N1 and P2 activity and a less negative occipital N1 activity, regardless of whether targets were attended or unattended, and a more positive occipital P2 activity when they were attended. Furthermore, the crowded minus uncrowded difference amplitude was more negative over the frontal region and more positive over the occipital region when the targets were attended, compared to when they were unattended during the N1 and P2 stages. This suggests that crowding, a concept that originates from Gestalt grouping, occurs automatically and can be modulated by attention.

**Keywords:** crowding, attention, Gestalt grouping, event-related potentials, temporal dynamic

## INTRODUCTION

The crowding effect is a visual phenomenon in which objects are easily identified in isolation, but become more difficult to identify when surrounded by other objects in the peripheral visual field (Pelli and Tillman, 2008). In a typical scenario, a letter can be identified when it is presented alone, but it cannot be recognized when it is flanked by other letters (Pelli et al., 2004). In addition to English letters, the identification of various other targets has been found to deteriorate in the presence of neighboring objects, including orientation signals (van den Berg et al., 2010), Chinese characters (Yeh et al., 2012; Peng et al., 2013; Zhou et al., 2016), and faces (Farzin et al., 2009).

The relationship between attention and crowding has been the focus of several recent studies. One perspective suggests that crowding is bottom-up and pre-attentive. This assertion is based on the notion that crowding occurs at a lower visual processing level. An early view suggested that the neural activity caused by flanking objects decreases neural activity related to the target due to lateral inhibition (Westheimer and Hauske, 1975). Another view, based on spatial pooling, posits that features from both targets and flankers are pooled, and that these features are compulsorily averaged (Parkes et al., 2001) or combined into a jumbled percept (Levi, 2008; Pelli and Tillman, 2008). Consistent with the pre-attentive account of crowding, Pöder (2006, 2007) demonstrated the important role of bottom-up salience in the binding of visual features, which determines the observed extent of crowding. Dakin et al. (2009) indicated that crowding does not reflect an attentional limit, and that crowding and attention rely on distinct neural mechanisms. Specifically, crowding persists even when people are completely unaware of the flankers, which suggests that conscious awareness and attention are not prerequisites for crowding (Ho and Cheung, 2011). Furthermore, Yong et al. (2014) examined the question of why individuals with posterior cortical atrophy (PCA) show excessive crowding in central vision and suggested that crowding in PCA can be regarded as a pre-attentive process that uses averaging to regularize the pathologically noisy representation of letter feature positions.

Another perspective of this topic suggests that crowding is top-down and attentional. This is based upon the idea that crowding occurs at a higher processing level. Therefore, while crowded targets can be perceived, the coarse attentional resolution in peripheral vision limits access of crowded targets to the consciousness (He et al., 1996, 1997). A probabilistic substitution model assumes that crowding results from binding a target and nearby distractors to incorrect spatial locations (Ester et al., 2014, 2015). This perspective has been supported by several studies. Attention improves performance at peripheral locations by enhancing spatial resolution (Yeshurun and Carrasco, 1998). Furthermore, attention reduces the critical target–flanker distance at which the flankers no longer interfere with target identification (Yeshurun and Rashal, 2010). Moreover, attention can be directly guided to various flankers. In a study examining this, attended flankers produced typical lateral interactions, while ignored flankers did not (Freeman et al., 2001). Attention modulates target–flanker integration, rather than just the processing of local flanker elements (Freeman et al., 2003). However, strong and specific attentional modulation of contour–integration mechanisms in early vision are sensitive to collinear configurations (Freeman et al., 2004). Therefore, covert attention to stimuli can increase the weights of their pooled features during crowding (Mareschal et al., 2010).

Electrophysiological studies suggest that attention plays a critical role in crowding. For instance, evidence from an attention-related N2pc component showed that attention functions to minimize interference from flankers at intermediate target–flanker distances (Hilimire et al., 2009, 2010; Bacigalupo and Luck, 2015). Additionally, evidence from a sustained posterior contralateral negativity (SPCN) study showed that

working memory may be recruited when attention fails to select the target at small target–flanker distances (Bacigalupo and Luck, 2015). The earliest ERP component, C1, which originates from V1 areas, is suppressed by crowded targets, whereas no suppression of C1 is found if the crowded targets are not attended. This indicates that attention-dependent V1 suppression contributes to crowding at a very early stage of visual processing (Chen et al., 2014). Chicherov et al. (2014) suggested that the P1 component reflects basic stimulus characteristics (i.e., flanker length), and N1 suppression reflects the occurrence of crowding when targets and flankers are grouped into wholes.

A third perspective suggests that crowding is a combination of bottom-up and top-down processes (i.e., crowding occurs automatically and can be modulated by attention). This perspective comes from the Gestalt grouping hypothesis of crowding. Increasing evidence shows that Gestalt grouping is critical for crowding (Malania et al., 2007; Saarela et al., 2009; Sayim et al., 2010). These findings are well explained by the hypothesis that crowding is strong when the flankers are grouped with the target and weaker when the target is segregated from the flankers (Manassi et al., 2012; Herzog et al., 2015; Herzog and Manassi, 2015). For instance, a cortical neural network model has been proposed that uses perceptual grouping and a novel segmentation process to account for several properties of visual crowding, such as effects of flanker length, the number of flanker lines, Gestalt effects, uncrowding effects, and similarity effects (Francis et al., 2017). While the relationship between Gestalt grouping and attention is well known, it is important to note that Gestalt grouping occurs automatically. Thus, visual stimuli that is irrelevant to a given task can be grouped without attention (Russell and Driver, 2005; Lamy et al., 2006), and the formation of visual object representations by grouping can occur outside the focus of voluntary attention (Müller et al., 2010). Electrophysiological evidence has shown that Gestalt stimuli automatically capture attention (Marini and Marzi, 2016). Further, Gestalt grouping has been shown to be modulated by attention. For example, grouping can be modulated by task relevance and attention as early as 100 ms after onset of sensory stimulation (Han et al., 2005), with this interaction between attention and grouping taking place as early in the perceptual process as the primary visual cortex (Wu et al., 2005; Khoe et al., 2006).

The relationship between attention and crowding needs further elucidation. First, additional evidence is necessary to test whether crowding can occur automatically. Behavior studies have shown that crowding is distinct from attention (Dakin et al., 2009) and occurs automatically (Ho and Cheung, 2011; Yong et al., 2014). Yet, several studies have emphasized that attention plays a crucial role in crowding, and that crowding may not occur completely automatically (Herzog et al., 2015; Francis et al., 2017). While N1 components have been found to be suppressed when observers discriminate crowded targets (Chicherov et al., 2014; Ronconi et al., 2016), no study to date has reported whether the N1 signal can be suppressed by crowded targets if they are unattended. Additionally, more direct evidence needs to be provided for the attentional modulation of crowding. Chicherov et al. (2014) reported that N1 suppression was much



stronger when the task was to discriminate the crowded Vernier as compared with the flankers length discrimination task. In this task, targets and flankers were in very close proximity, especially for crowded targets; thus, the Vernier may be attended to a certain extent, and this may lead to a slight N1 suppression in the length discrimination task. In the ideal situation, it is necessary to examine the N1 suppression elicited by crowded targets both when the targets are attended and when they are not.

The present study focused on the temporal dynamic relationship between crowding and attention. We combined a crowding paradigm (Yeh et al., 2012; Peng et al., 2013; Zhou et al., 2016) with a cross-modal delayed response oddball paradigm (Wei et al., 2002; Chen et al., 2010). The intermodal selective attention paradigm has been shown to effectively manipulate attention (Alho, 1992; Woods et al., 1992). The cross-modal delayed response paradigm can effectively control attention and minimize target effects (Wei et al., 2002). A series of auditory and visual stimuli are presented in sequence. The presentation of an attended visual stimulus (e.g., a crowded or uncrowded target) is followed by zero, one, or two unattended auditory stimuli (e.g., a tone) and a response signal. This order can also be reversed (e.g., the tone is the attended stimulus, and the crowded or uncrowded target is the unattended stimulus). Participants are required to identify signals of the attended modality, ignore those of the unattended modality, and wait to respond until presentation of the response signal.

Given that crowding is reflected by N1 suppression (Chicherov et al., 2014; Ronconi et al., 2016), three hypotheses on the relationship between attention and crowding can be tested. First, if crowding is bottom-up and pre-attentive, N1 should be suppressed by crowded targets when the targets are not unattended. Next, if crowding is top-down and attentional, N1 suppression would be stronger when the crowded targets are attended than when unattended. Finally, if crowding is a combination of bottom-up and top-down processes, the above two predictions would be observed simultaneously.

Further, additional evidence can be provided to test the Gestalt grouping hypothesis by a measurement of the P2 component. A study reported that P2 engages in grouping elements into a unitary object (Flevaris et al., 2013). If crowding originates from Gestalt grouping, similar results to the N1 stage would be observed during the P2 stage. In other words, a significant difference in P2 amplitude between crowded and uncrowded targets, both in the attended and unattended conditions, will be observed, and the crowded minus uncrowded difference amplitude will be modulated by attention.

## MATERIALS AND METHODS

### Participants

Eighteen right-handed undergraduate students (two males, 19–24 years of age) participated in this experiment. One participant (female, 23 years of age) was excluded because of excessive eye movement and blinking during the experiment. Participants were not taking any medications and did not suffer from any central nervous system abnormalities or injuries. All were naive

to the purpose of the experiment. The study was approved by the institutional review board of Southwest University. 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 20 Chinese single-character words, four Chinese pseudo-characters, a fixation, and a visual response signal. Ten Chinese words indicated animals (e.g., 象 means elephant), and ten indicated inanimate objects (e.g., 家 means home). Four pseudo-characters were made from stroke features using TrueType software. They were white, single-bodied, and had no semantic meaning (Peng et al., 2013). The visual angles of all the characters were  $1 \times 1^\circ$ . The fixation point was a white dot with a diameter of  $0.4^\circ$ . The visual response signal was a red square with a width of  $0.5^\circ$ . Visual stimuli were presented on a 22-inch Iiyama MA203DT D color monitor with a background screen color of medium gray (RGB color coordinates: 128, 128, 128). The refresh rate of the computer monitor was 85 Hz. The computer screen was placed approximately 80 cm in front of the participants' eyes.

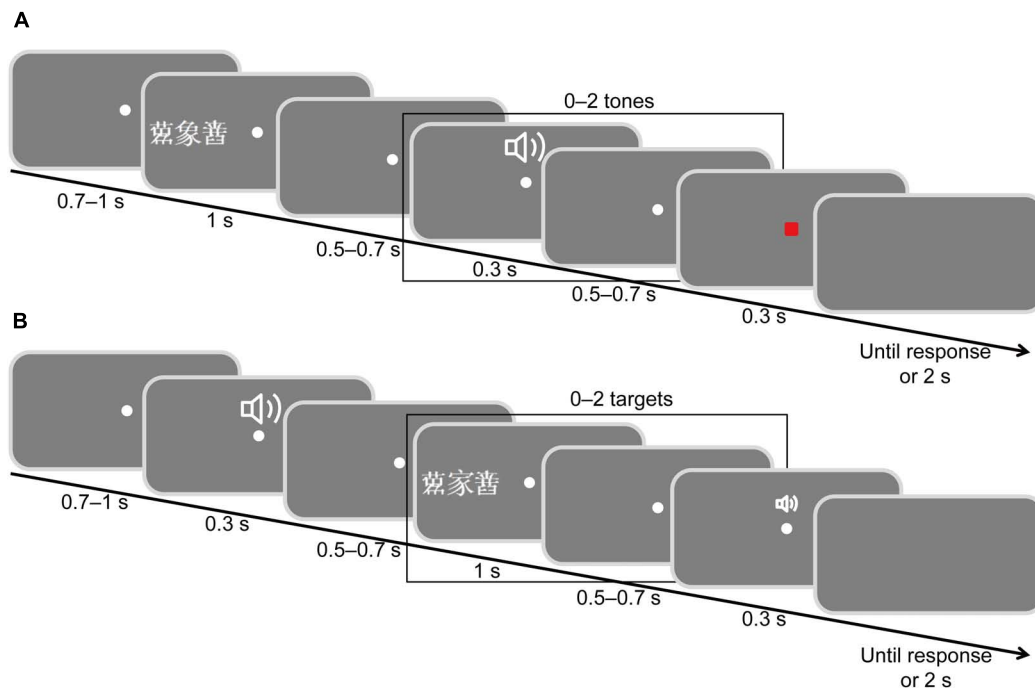
Auditory stimuli were two sinusoidal tones and an auditory response signal. The sinusoidal tones were delivered with 1000 or 800 Hz at 30 ms, 60 dB HL. The auditory response signal was a faint click at 500 Hz, 30 ms, 20 dB HL. All auditory stimuli were presented binaurally through earphones.

### Procedure

The experiment employed a cross-modal delayed response oddball paradigm (Figure 1; Wei et al., 2002; Chen et al., 2010). Participants were asked to fixate on the center of the screen and put on earphones. Each participant carried out two tasks. Task 1 involved attending to visual stimuli and ignoring auditory stimuli; Task 2 involved attending to auditory stimuli and ignoring visual stimuli. The orders of the two tasks (attending visual and attending auditory) were counterbalanced between participants. Both Task 1 and Task 2 included 800 trials. Participants were provided rest for 30 s after finishing 100 trials, and for 2 min after finishing a task. The experimental procedure was programmed with E-prime 1.1.

### Task 1: Attending to Visual Stimuli While Ignoring Auditory Stimuli

Participants were instructed to attend to visual signals and ignore auditory signals (Figure 1A). They were required to fixate on the center point throughout the study and to view targets only utilizing peripheral vision. At the beginning of each trial, a white fixation dot was presented at the center of the screen for 500–700 ms. A target word and two flankers (flanker, target, flanker) were randomly presented for 1000 ms in either the left or right visual field on the horizontal meridian. The eccentricity of the target was  $6^\circ$ , and the spacing between the target and the flankers was at either  $1^\circ$  (crowded trials) or  $4^\circ$  (uncrowded trials). The crowded and uncrowded targets were presented randomly. According to a typical oddball paradigm (Hillyard et al., 1973), we set animal targets as deviant stimuli with a small probability



**FIGURE 1 |** Schematic illustration of the stimulus display sequence. **(A)** Attending to visual stimuli while ignoring auditory stimuli. **(B)** Attending to auditory stimuli while ignoring visual stimuli. The Chinese character 象 means elephant and 家 means home.

(20%) and inanimate targets as standard stimuli with a large probability (80%). There were 160 animal targets, including 80 crowded and 80 uncrowded targets, and 640 inanimate targets, including 320 crowded and 320 uncrowded targets. Two flankers were selected randomly from the four pseudo-characters. After a randomized delay of 500–700 ms, 0–2 tones were presented. The duration of each tone was 30 ms. The time interval between the two tones was 500–700 ms. The frequency of 640 tones was 800 Hz, while 160 tones were 1000 Hz. Finally, a visual response signal (a small red square) was presented for 30 ms after a randomized inter-stimulus interval of 500–700 ms. Participants were required to judge the meanings of targets and to make responses by pressing one of the two mouse buttons with the thumb of either hand. Half of the participants were instructed to press the left mouse button if the meaning of target word was not an animal and to press the right mouse button if the target's meaning was an animal, whereas the other half of the participants were instructed to perform the opposite action. Once the small red square appeared, participants were required to respond as quickly and accurately as possible. The next trial was presented once the participants had responded; the maximum time interval for response was 2000 ms.

### Task 2: Attending to Auditory Stimuli While Ignoring Visual Stimuli

Participants were instructed to attend to auditory stimuli and ignore visual stimuli (**Figure 1B**). A tone (0–2 crowded or uncrowded words) and a faint click were presented successively. A randomized delay of 500–700 ms was inserted between the two

stimuli. Participants were asked to discriminate the tone pitches and hold their response until the response signal (the faint click) was presented at the end of the trial. Half of the participants were instructed to press the left mouse button if the pitch was 800 Hz and to press the right mouse button if the pitch was 1000 Hz, while the other half of the participants were instructed to do the opposite. Other details of the task were the same as those reported in Task 1.

### Electrophysiological Recording

Continuous electroencephalogram (EEG) was acquired from Ag/AgCl electrodes mounted on a Quick-Cap (Neuroscan Inc.). Sixty-four electrodes were positioned according to the extended 10–20 system. All EEG electrodes were referenced to the left mastoid. The horizontal electrooculogram (EOG) was acquired using a bipolar pair of 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 0.05–100 Hz and were stored for offline analysis. All electrode impedances were maintained below 5 k $\Omega$ .

### EEG Analysis

EEGLAB (Delorme and Makeig, 2004) and MATLAB (The MathWorks, Inc., Massachusetts, United States) were used for offline EEG data processing. Continuous EEG data were re-referenced to the average of the right and left mastoids and were digitally low-pass filtered at 40 Hz. ERPs were time-locked to the onset of the target words, with an average epoch of

700 ms, including a 100 ms pre-stimulus baseline. All trials, no matter whether the response was correct or not, were included in analysis.

Ocular artifacts were rejected using a two-step procedure (Woodman and Luck, 2003; Luck, 2005). In the first step, for each point in the epoch, the mean value of the preceding 100 ms and that of the subsequent 100 ms were determined, and a difference value between two mean values was calculated. After this action was performed for each point, the largest difference value was compared with a threshold to determine whether the trial should be rejected. The single-trial waveforms were checked by visual inspection to determine a threshold value for each individual participant. Using the threshold, all clearly visible artifacts were rejected without the rejection of large numbers of artifact-free trials. We also excluded any participant for whom more than 25% of the trials were rejected owing to eye movements (Woodman and Luck, 2003). One participant's data were excluded from the analysis because artifacts led the rejection of 46.8% of trials. On average, 11.4% of trials, ranging from 1.9 to 21.4% were rejected for the 17 remaining participants.

In the second step, the average horizontal EOG waveforms for left-target and right-target trials were calculated to assess the degree of residual eye movement activity. The average difference in voltage between left-target and right-target trials was less than  $2.7 \mu\text{V}$ , which corresponded to an average eye movement of less than  $0.2^\circ$  (Lins et al., 1993; Zhang and Luck, 2009). Thus, it was determined that subjects were able to maintain fixation on the central fixation point throughout the task.

P1 (peaking at about 90 ms), N1 (about 160 ms), and P2 (about 230 ms) components were elicited by both crowded and uncrowded targets in both attended and unattended conditions (Figure 3). As shown in Figure 3, ERP component amplitude was measured from the mean amplitude of the 40-ms window centered at the grand average ERP peak latency and was separately determined for each condition (Näätänen et al., 2004).

The regions of interest (ROIs) were chosen according to previous studies and topographic information regarding P1, N1, P2, and crowded minus uncrowded wave differences observed in the current study (Figures 4, 5). Previous studies revealed the functional significance of the occipital region in crowding (Chen et al., 2014; Chicherov et al., 2014). The P1, N1, and P2 were predominantly distributed over the frontal, central, or occipital regions (Figure 4). As shown in Figure 5, results are consistent with previous studies that positive occipital distribution is accompanied by a negative frontal distribution (Clark and Hillyard, 1996; Fleva et al., 2013). Thus, frontal and occipital electrodes were chosen as ROIs. ERP amplitudes at the F1, F2, F3, F4, F5, F6, and Fz electrode sites were averaged as measures of the frontal cluster, and those at the O1, O2, Oz, PO7, PO8, P7, and P8 electrode sites were averaged as measures of the occipital cluster (Zhang and Luck, 2009).

Planned comparisons were performed to address specific hypotheses. In order to assess whether crowding occurs automatically, each ERP component (P1, N1, and P2) was subjected to a paired samples *t*-test. These tests were conducted on the mean amplitude of ERP components to determine whether the means of the crowded and uncrowded conditions were equal.

Paired *t*-tests were conducted in the attended and unattended conditions over both the frontal and occipital regions (four tests for each ERP component). To obtain a family wise confidence level of 0.95, a Bonferroni correction was used to adjust each individual confidence interval of 0.9875, and the corresponding significance level was set at  $0.05/4 = 0.0125$  (Armstrong, 2014).

In order to assess whether attention modulates crowding, difference amplitudes were obtained by subtracting the amplitudes of uncrowded ERP components from that of crowded ERP components in the attended and unattended conditions, respectively. For each ERP component, a paired sample *t*-test was conducted on the difference amplitudes to test whether the means of the attended and unattended conditions were equal. Paired *t*-tests were conducted over both the frontal and occipital region (two tests for each ERP component). The corresponding significance level was  $0.05/2 = 0.025$ . Cohen's *d* was used to estimate the effect size of the *t*-tests.

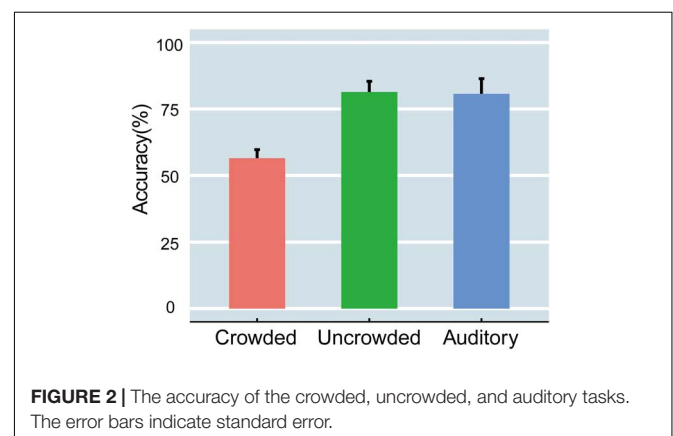
## RESULTS

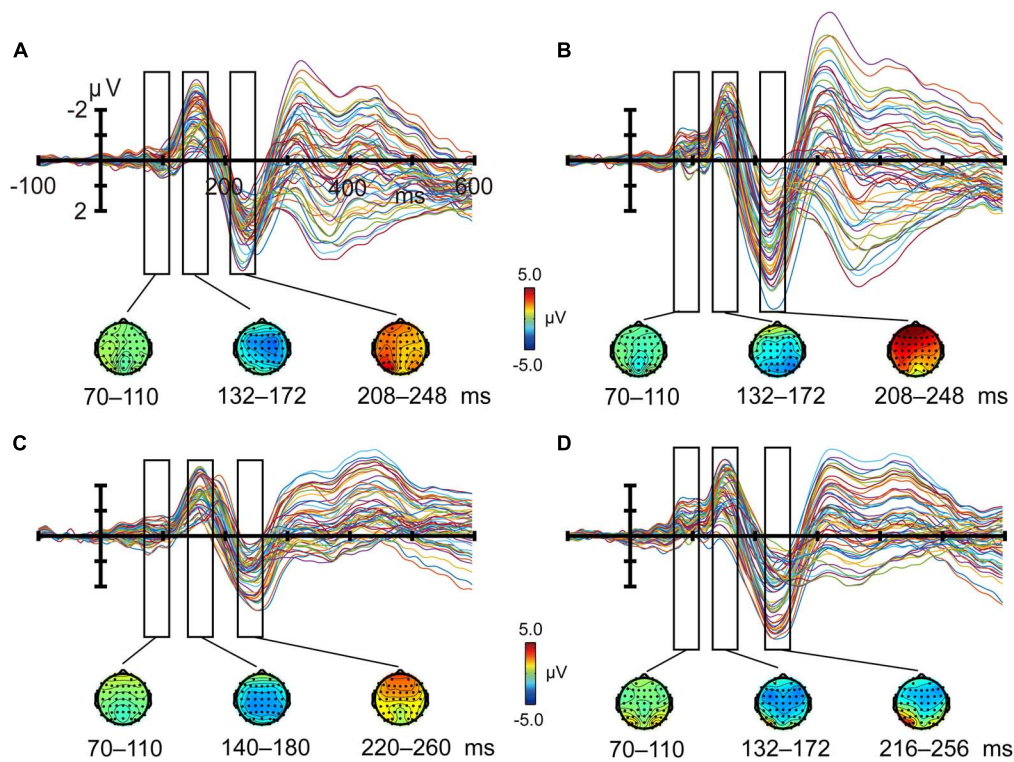
### Behavioral Data

Accuracy was computed for each participant in the crowded, uncrowded, and auditory conditions (Figure 2). A one-way, repeated measures analysis of variance (ANOVA) that was performed on accuracy scores revealed a significant main effect of condition [ $F(2,32) = 23.853$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.599$ ]. Specifically, accuracy was significantly lower in the crowded (ranging from 39 to 79%) than in the uncrowded condition (ranging from 53% to 96%) [ $t(16) = -7.396$ ,  $p < 0.001$ , Cohen's  $d = -1.794$ ]. However, the accuracy difference between the uncrowded and auditory (ranging from 58 to 98%) conditions was not significant [ $t(16) = -0.144$ ,  $p > 0.05$ , Cohen's  $d = 0.035$ ]. The results indicated that crowding led to a significant decline in performance, and that the identification of uncrowded targets had approximately the same level of difficulty as that observed in the auditory task.

### Event-Related Potential Data

Figures 3, 4 shows ERP waveforms elicited by crowded and uncrowded targets in both attended and unattended conditions.





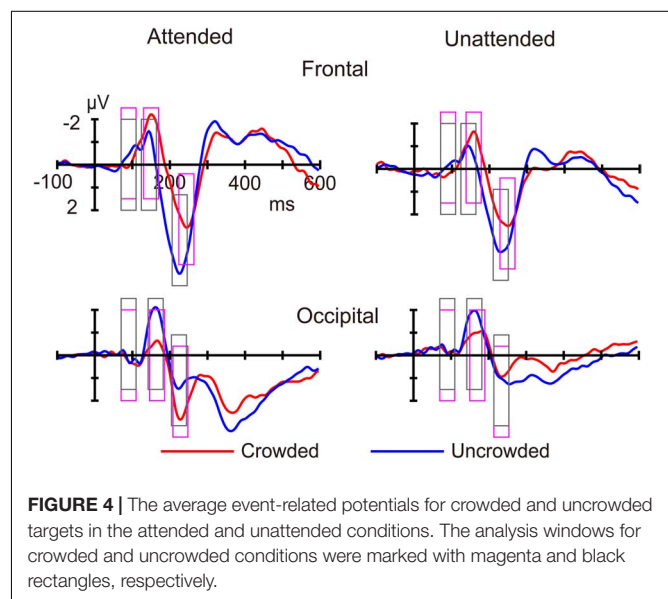
**FIGURE 3 |** Butterfly plots of grand average event-related potentials and topographies. Crowded (A) and uncrowded (B) targets in the attended condition; crowded (C) and uncrowded (D) targets in the unattended condition.

An obvious separation between the crowded and uncrowded targets appeared during the N1–P2 stage (Figure 4). Figure 5A shows wave amplitude differences obtained by subtracting uncrowded ERPs from crowded ERPs in the attended and unattended conditions. Figure 5B shows the topographic results of the crowded minus uncrowded difference waves in the attended and unattended conditions during the P1, N1, and P2 stages. Compared with difference in amplitude wave in the unattended group, the difference in the amplitude wave of the attended group was more negative over the frontal region and more positive over the occipital region during the N1 and P2 stages (Figure 5).

Planned comparisons showed that crowded targets elicited a more positive P1 amplitudes compared with uncrowded targets in the attended condition [ $t(16) = 3.132$ ,  $p < 0.01$ , Cohen's  $d = 0.760$ ] and in the unattended condition [ $t(16) = 4.094$ ,  $p < 0.01$ , Cohen's  $d = 0.993$ ] over the frontal region. However, the difference between the crowded and uncrowded conditions was not significant in the attended and unattended conditions over the occipital region ( $p$ -values  $> 0.05$ ; Figure 6A).

For the difference in wave amplitude (crowded – uncrowded) during the P1 stage, planned comparisons did not reveal any significant differences between the attended and unattended conditions over the frontal and occipital regions ( $p$ -values  $> 0.05$ ).

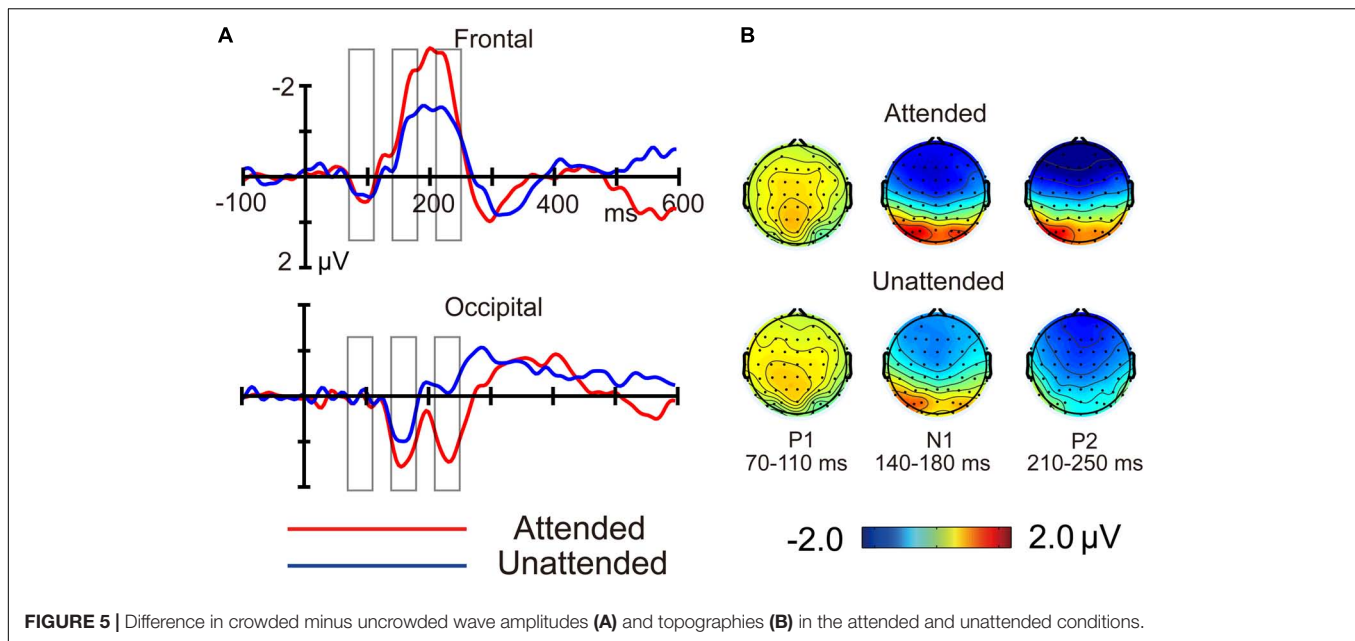
Planned comparisons showed that crowded targets elicited a more negative N1 amplitude compared with uncrowded targets



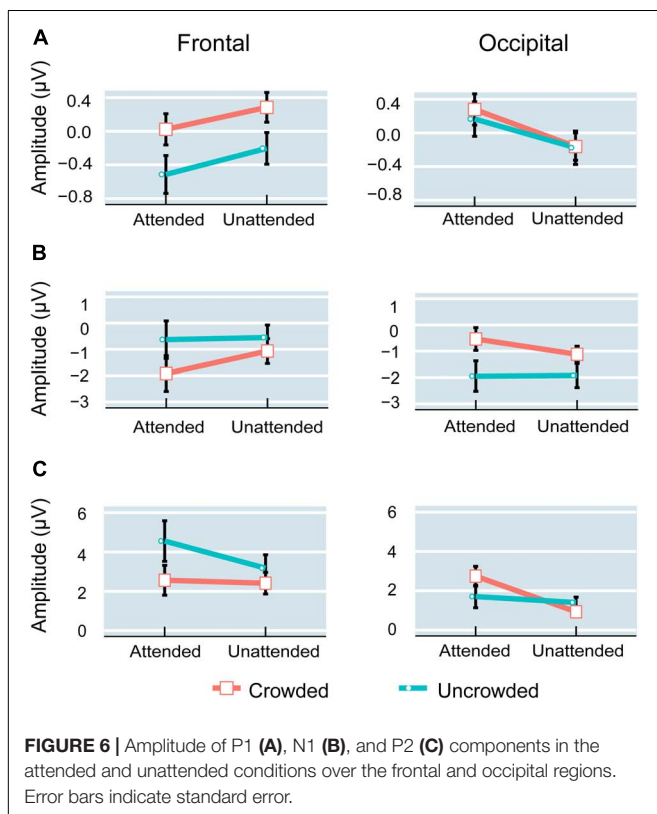
**FIGURE 4 |** The average event-related potentials for crowded and uncrowded targets in the attended and unattended conditions. The analysis windows for crowded and uncrowded conditions were marked with magenta and black rectangles, respectively.

in the attended [ $t(16) = -4.829$ ,  $p < 0.001$ , Cohen's  $d = -1.171$ ] and unattended [ $t(16) = -2.980$ ,  $p < 0.01$ , Cohen's  $d = -0.723$ ] conditions over the frontal region. The crowded targets elicited a less negative N1 amplitude compared to uncrowded targets in both the attended [ $t(16) = 5.159$ ,  $p < 0.001$ , Cohen's  $d = 1.251$ ]





**FIGURE 5 |** Difference in crowded minus uncrowded wave amplitudes (A) and topographies (B) in the attended and unattended conditions.



**FIGURE 6 |** Amplitude of P1 (A), N1 (B), and P2 (C) components in the attended and unattended conditions over the frontal and occipital regions. Error bars indicate standard error.

and unattended [ $t(16) = 3.394$ ,  $p < 0.01$ , Cohen's  $d = 0.823$ ] conditions over the occipital region (Figure 6B).

During the N1 stage, planned comparisons revealed that the difference amplitude was more negative in the attended ( $-1.289 \pm 0.267 \mu V$ ) than in the unattended condition ( $-0.516 \pm 0.173 \mu V$ ) over the frontal region [ $t(16) = -2.852$ ,

$p < 0.025$ , Cohen's  $d = -0.692$ ]. Further, the difference amplitude was more positive in the attended ( $1.410 \pm 0.273 \mu V$ ) than in the unattended condition ( $0.804 \pm 0.237 \mu V$ ) over the occipital region [ $t(16) = 2.665$ ,  $p < 0.025$ , Cohen's  $d = 0.646$ ].

Planned comparisons showed that the crowded targets elicited less positive P2 amplitudes compared with the uncrowded targets in the attended [ $t(16) = -3.706$ ,  $p < 0.01$ , Cohen's  $d = -0.899$ ] and unattended [ $t(16) = -3.461$ ,  $p < 0.01$ , Cohen's  $d = -0.839$ ] conditions over the frontal region. Over the occipital region, the crowded targets elicited a more positive P2 amplitude as compared with the uncrowded targets in the attended condition [ $t(16) = 3.701$ ,  $p < 0.01$ , Cohen's  $d = 0.898$ ], whereas there was no significant difference between the crowded and uncrowded targets in the unattended condition ( $p > 0.05$ ; Figure 6C).

In the P2 stage, the difference amplitude was more negative in the attended ( $-1.996 \pm 0.539 \mu V$ ) than in the unattended condition ( $-0.7648 \pm 0.221 \mu V$ ) over the frontal region [ $t(16) = -2.829$ ,  $p < 0.025$ , Cohen's  $d = -0.686$ ], whereas it was more positive in the attended ( $1.038 \pm 0.280 \mu V$ ) than in the unattended condition ( $-0.473 \pm 0.240 \mu V$ ) over the occipital region [ $t(16) = 5.255$ ,  $p < 0.001$ , Cohen's  $d = 1.275$ ].

## DISCUSSION

The present study combined a selective attention paradigm with a crowding paradigm to identify the relationship between attention and crowding. Consistent with previous studies (Pelli et al., 2004; Yeh et al., 2012; Peng et al., 2013), the ability to discriminate crowded targets dropped sharply compared to uncrowded targets (Figure 2). Additionally, the present study reproduced the N1 suppression in the crowding task (Chicherov et al., 2014; Ronconi et al., 2016). Furthermore, the current results

suggest that crowding indeed occurs automatically and can be modulated by attention.

We found that crowded targets evoked a more positive P1 component compared to uncrowded targets, irrespective of whether they were attended or unattended. Previous studies have shown that the P1 wave reflects early visual processing of low-level characteristics of stimuli, such as luminance, intensity, eccentricity, and size (Johannes et al., 1995; Busch et al., 2004; Schadow et al., 2007). Chicherov et al. (2014) reported that the P1 amplitude positively correlated with the length of flankers in a Vernier crowding task. Consistent with the findings of such previous studies, P1 was found to reflect the early, lower-level visual processing of stimulus characteristics in the present study.

We found that the N1 was largest over the frontal-central region or occipital region (**Figure 3**), which is consistent with a previous study that a posterior N150 was distributed over the occipitoparietal region, and an anterior N155 was distributed over the frontal-central region (Di Russo et al., 2002). A posterior N1 component is usually accompanied by a smaller frontal component with reverse polarity (Clark and Hillyard, 1996). It may be in part due to volume transmission (Hedge et al., 2015). However, the entire frontal N1 is not due to volume transmission of N1 from occipital areas, but an overlap of a small positive component and a negative N155, thus we did not observe a frontal N1 component with reverse polarity in the current study. It is consistent with a notion that ERP events can be a complex result of underlying neural phenomena, which are difficult to study (Luck, 2014). Only one thing can be certain is that the timing of the ERP event in occipital and frontal areas (Luck, 2014).

Additionally, the current study found that crowded targets elicited a more negative frontal N1 and a less negative occipital N1 compared with uncrowded targets, irrespective of whether they were attended or unattended (**Figure 4**). These results replicated the findings of the occipital N1 suppression of crowding (Chicherov et al., 2014; Ronconi et al., 2016) and were consistent with the previous finding that crowding is associated with a suppression of V1, regardless of whether targets were attended or unattended (Millin et al., 2014). These results are in line with the bottom-up and pre-attentive account of crowding. Furthermore, the current study found that the difference in amplitude of the crowded minus uncrowded wave during the N1 stage was more negative over the frontal and more positive over the occipital region, when the targets were attended (**Figures 5, 6**). These results are in line with the top-down and attentional account of crowding. Thus, the current study provided electrophysiological evidence that crowding occurs automatically, and that it can be modulated by attention. The lateral inhibition and spatial pooling hypotheses predict that crowding occurs automatically (Westheimer and Hauske, 1975; Parkes et al., 2001), while the attentional resolution hypothesis predicts that crowding can be modulated by attention (He et al., 1996, 1997). However, these crowding hypotheses cannot fully predict the relationship between attention and crowding. The Gestalt grouping principle is more applicable to the current study. When the flankers and targets were closer, they were grouped, and therefore, crowding occurred.

The Gestalt grouping hypothesis predicts that crowding is a combination of bottom-up (Müller et al., 2010; Marini and Marzi, 2016) and top-down (Wu et al., 2005; Khoe et al., 2006) processes. The current results most closely align with predictions made by the Gestalt grouping hypothesis of crowding.

In addition, the present findings of the P2 component are consistent with the Gestalt grouping hypothesis of crowding. Similar to the N1 component, the current study found that, crowded targets (compared to uncrowded targets) elicited a more negative frontal P2, irrespective of whether the targets were attended or unattended, as well as a less negative occipital P2 when targets were attended (**Figure 4**). Further, the difference in the crowded minus uncrowded wave amplitude during the P2 stage was more negative over the frontal and more positive over the occipital region when the targets were attended than when they were unattended (**Figures 5, 6**). It was also noted that the topographic elements of the difference waves were similar during the N1 and P2 stages (**Figure 5B**). A previous study reported that irrelevant probes superimposed on a moving image elicited an enhanced P2 component when the probes were contained within the boundaries of an object that was perceived as unitary, and that the topography of the P2 elicited by probes during object perception was distinct from that during fragment perception. These results indicate that the P2 wave is associated with grouping elements into unitary objects (Flevaris et al., 2013). Similar to the N1 component, P2 was found to be associated with Gestalt grouping in the present study.

The effects of target presentation time deserve further consideration. The presentation time of targets was 1 s in the present study, which was in line with the procedure followed in a previous study (Peng et al., 2013). Both the onset and offset of targets elicit neural activity (Baltzell and Billings, 2014). Using a long presentation time, we can avoid the overlapping of the neural activity of the offset with that of the cognitive process that we examined. However, more eye movement might be caused by visual stimuli with a longer presentation time. The current study rejected ocular artifacts using a two-step procedure (Woodman and Luck, 2003; Luck, 2005) and found that the remaining average eye movement was less than 0.2°. Thus, the effect of eye movement was excluded from ERP data. In addition, a longer presentation time may result in targets being able to access consciousness. The cross-modal delayed response of the oddball paradigm has been shown to control attention effectively; P300 component, an index of working memory and conscious perception (Salti et al., 2012), was only observed in the attended condition (Wei et al., 2002). This indicates that Wei et al.'s paradigm can exclude conscious access before and during the P300 stage in the unattended condition. The current study focused upon P1 (peaking at about 90 ms), N1 (about 160 ms), and P2 (about 230 ms) components. These components are earlier than P300, thus these components are not affected by consciousness in the unattended condition. Though we cannot exclude the possibility that the unattended targets access consciousness after the P300 stage, this does not affect the explanations of the findings on the P1, N1, and P2 components in the current study.

Finally, analytical methods used were checked to determine if they affected the conclusions of the current study. For instance, a two-step procedure to reject ocular artifacts (Woodman and Luck, 2003; Luck, 2005) was used. This procedure was utilized in a previous study on neural correlates of visual crowding (Chicherov et al., 2014). A separate study on the neural oscillatory correlates of crowding used Independent Component Analysis (ICA) to detect and correct ocular artifacts and removed epochs containing voltage deviation that exceeded  $\pm 75 \mu\text{V}$  (Ronconi et al., 2016). To the best of our knowledge, no previous research on this topic has checked whether the above two procedures are functionally equivalent. In addition, the current study used planned comparisons rather than ANOVA, because specific hypotheses that crowding occurs automatically and attention modulates crowding were maintained. It is necessary to determine whether similar statistical results can be obtained by using ANOVA. To address these issues, we conducted a supplementary analysis using ICA and ANOVA. Continuous EEG data were re-referenced, filtered, and segmented in the same manner described in Section 2.5 EEG analysis. Then, the  $\pm 75 \mu\text{V}$  ICA criterion was used to remove ocular artifacts. Similar ERP waveforms were obtained (**Supplementary Figures S1–S3**). Further, repeated measures ANOVA was conducted on amplitudes of P1, N1, and P2 waves. This analysis yielded similar statistical results (**Supplementary Results**). Therefore, the conclusions drawn in the current study do not appear to be affected by the analytical methods used.

## CONCLUSION

The present study employed a “cross-modal delayed response” oddball paradigm to investigate the relationship between attention and crowding. Previous findings that P1 reflects

the early low-level processing of stimuli characteristics were replicated. We revealed that the N1 and P2 components were associated with the concept of Gestalt grouping in crowding. Specifically, crowding-related neural activity was found to appear, regardless of whether the crowded targets were attended or unattended. Additionally, neural activities appeared to be modulated by attention during the N1 and P2 stages. These results suggest that crowding occurs automatically and can be modulated by attention. Our results are consistent with previous studies on Gestalt grouping, which supports the notion that crowding originates from Gestalt grouping.

## AUTHOR CONTRIBUTIONS

CP and YC designed the study. CP and CH performed the experiments. CP and YC decided on the final analyses and interpretation, and wrote the manuscript.

## FUNDING

This study was supported by a grant from the National Natural Science Foundation of China (Grant Nos. 31300845 and 31200855), Research Fund of Chongqing University of Arts and Sciences (Grant No. R2012JY23), and the Key Research Institute of Humanities and Social Science in Chongqing (Grant No. 16SKB008).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2018.00844/full#supplementary-material>

## REFERENCES

- Alho, K. (1992). Selective attention in auditory processing as reflected by event-related brain potentials. *Psychophysiology* 29, 247–263. doi: 10.1111/j.1469-8986.1992.tb01695.x
- Armstrong, R. A. (2014). When to use the Bonferroni correction. *Ophthalmic Physiol. Opt.* 34, 502–508. doi: 10.1111/opo.12131
- Bacigalupo, F., and Luck, S. J. (2015). The allocation of attention and working memory in visual crowding. *J. Cogn. Neurosci.* 27, 1180–1193. doi: 10.1162/jocn\_a\_00771
- Baltzell, L. S., and Billings, C. J. (2014). Sensitivity of offset and onset cortical auditory evoked potentials to signals in noise. *Clin. Neurophysiol.* 125, 370–380. doi: 10.1016/j.clinph.2013.08.003
- 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
- Chen, J., He, Y., Zhu, Z., Zhou, T., Peng, Y., Zhang, X., et al. (2014). Attention-dependent early cortical suppression contributes to crowding. *J. Neurosci.* 34, 10465–10474. doi: 10.1523/JNEUROSCI.1140-14.2014
- Chen, Y. G., Huang, X. T., Luo, Y. M., Peng, C. H., and Liu, C. X. (2010). Differences in the neural basis of automatic auditory and visual time perception: ERP evidence from an across-modal delayed response oddball task. *Brain Res.* 1325, 100–111. doi: 10.1016/j.brainres.2010.02.040
- Chicherov, V., Plomp, G., and Herzog, M. H. (2014). Neural correlates of visual crowding. *Neuroimage* 93, 23–31. doi: 10.1016/j.neuroimage.2014.02.021
- Clark, V. P., and Hillyard, S. A. (1996). Spatial selective attention affects early extrastriate but not striate components of the visual evoked potential. *J. Cogn. Neurosci.* 8, 387–402. doi: 10.1162/jocn.1996.8.5.387
- Dakin, S. C., Bex, P. J., Cass, J. R., and Watt, R. J. (2009). Dissociable effects of attention and crowding on orientation averaging. *J. Vis.* 9, 28.1–28.16. doi: 10.1167/9.11.28
- 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
- Di Russo, F., Martinez, A., Sereno, M. I., Pitzalis, S., and Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Hum. Brain Mapp.* 15, 95–111. doi: 10.1002/hbm.10010
- Ester, E. F., Klee, D., and Awh, E. (2014). Visual crowding cannot be wholly explained by feature pooling. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1022–1033. doi: 10.1037/a0035377
- Ester, E. F., Zilber, E., and Serences, J. T. (2015). Substitution and pooling in visual crowding induced by similar and dissimilar distractors. *J. Vis.* 15:4. doi: 10.1167/15.1.4
- Farzin, F., Rivera, S. M., and Whitney, D. (2009). Holistic crowding of Mooney faces. *J. Vis.* 9:18. doi: 10.1167/9.6.18

- Flevaris, A. V., Martínez, A., and Hillyard, S. A. (2013). Neural substrates of perceptual integration during bistable object perception. *J. Vis.* 13:17. doi: 10.1167/13.13.17
- Francis, G., Manassi, M., and Herzog, M. H. (2017). Neural dynamics of grouping and segmentation explain properties of visual crowding. *Psychol. Rev.* 124, 483–504. doi: 10.1037/rev0000070
- Freeman, E., Driver, J., Sagi, D., and Li, Z. (2003). Top-down modulation of lateral interactions in early vision: does attention affect integration of the whole or just perception of the parts? *Curr. Biol.* 13, 985–989. doi: 10.1016/S0960-9822(03)00333-6
- Freeman, E., Sagi, D., and Driver, J. (2001). Lateral interactions between targets and flankers in low-level vision depend on attention to the flankers. *Nat. Neurosci.* 4, 1032–1036. doi: 10.1038/nn728
- Freeman, E., Sagi, D., and Driver, J. (2004). Configuration-specific attentional modulation of flanker-target lateral interactions. *Perception* 33, 181–194. doi: 10.1068/p3481
- Han, S., Jiang, Y., Mao, L., Humphreys, G. W., and Qin, J. (2005). Attentional modulation of perceptual grouping in human visual cortex: ERP studies. *Hum. Brain Mapp.* 26, 199–209. doi: 10.1002/hbm.20157
- He, S., Cavanagh, P., and Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature* 383, 334–337. doi: 10.1038/383334a0
- He, S., Cavanagh, P., and Intriligator, J. (1997). Attentional resolution. *Trends Cogn. Sci.* 1, 115–121. doi: 10.1016/S1364-6613(97)89058-4
- Hedge, C., Stothart, G., Todd Jones, J., Rojas Frías, P., Magee, K. L., and Brooks, J. C. W. (2015). A frontal attention mechanism in the visual mismatch negativity. *Behav. Brain Res.* 293, 173–181. doi: 10.1016/j.bbr.2015.07.022
- Herzog, M. H., and Manassi, M. (2015). Uncorking the bottleneck of crowding: a fresh look at object recognition. *Curr. Opin. Behav. Sci.* 1, 86–93. doi: 10.1016/j.cobeha.2014.10.006
- Herzog, M. H., Sayim, B., Chicherov, V., and Manassi, M. (2015). Crowding, grouping, and object recognition: a matter of appearance. *J. Vis.* 15:5. doi: 10.1167/15.6.5
- Hilimire, M. R., Mounts, J. R. W., Parks, N. A., and Corballis, P. M. (2009). Competitive interaction degrades target selection: an ERP study. *Psychophysiology* 46, 1080–1089. doi: 10.1111/j.1469-8986.2009.00846.x
- Hilimire, M. R., Mounts, J. R. W., Parks, N. A., and Corballis, P. M. (2010). Event-related potentials dissociate effects of salience and space in biased competition for visual representation. *PLoS One* 5:e12677. doi: 10.1371/journal.pone.0012677
- Hillyard, S. A., Hink, R. F., Schwent, V. L., and Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science* 182, 177–180. doi: 10.1126/science.182.4108.177
- Ho, C., and Cheung, S.-H. (2011). Crowding by invisible flankers. *PLoS One* 6:e28814. doi: 10.1371/journal.pone.0028814
- Johannes, S., Münte, T. F., Heinze, H. J., and Mangun, G. R. (1995). Luminance and spatial attention effects on early visual processing. *Brain Res. Cogn. Brain Res.* 2, 189–205. doi: 10.1016/0926-6410(95)90008-X
- Khoe, W., Freeman, E., Woldorff, M. G., and Mangun, G. R. (2006). Interactions between attention and perceptual grouping in human visual cortex. *Brain Res.* 1078, 101–111. doi: 10.1016/j.brainres.2005.12.083
- Lamy, D., Segal, H., and Ruderman, L. (2006). Grouping does not require attention. *Percept. Psychophys.* 68, 17–31. doi: 10.3758/BF03193652
- Levi, D. M. (2008). Crowding—an essential bottleneck for object recognition: a mini-review. *Vision Res.* 48, 635–654. doi: 10.1016/j.visres.2007.12.009
- Lins, O. G., Picton, T. W., Berg, P., and Scherg, M. (1993). Ocular artifacts in EEG and event-related potentials. I: scalp topography. *Brain Topogr.* 6, 51–63.
- Luck, S. J. (2005). *An Introduction to the Event-Related Potential Technique*. London: The MIT Press.
- Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. 2nd Edn, Cambridge, MA: The MIT Press.
- Malania, M., Herzog, M. H., and Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds. *J. Vis.* 7:1. doi: 10.1167/7.2.1
- Manassi, M., Sayim, B., and Herzog, M. H. (2012). Grouping, pooling, and when bigger is better in visual crowding. *J. Vis.* 12:13. doi: 10.1167/12.10.13
- Mareschal, I., Morgan, M. J., and Solomon, J. A. (2010). Attentional modulation of crowding. *Vision Res.* 50, 805–809. doi: 10.1016/j.visres.2010.01.022
- Marini, F., and Marzi, C. A. (2016). Gestalt perceptual organization of visual stimuli captures attention automatically: electrophysiological evidence. *Front. Hum. Neurosci.* 10:446. doi: 10.3389/fnhum.2016.00446
- Millin, R., Arman, A. C., Chung, S. T., and Tjan, B. S. (2014). Visual crowding in V1. *Cereb. Cortex* 24, 3107–3115. doi: 10.1093/cercor/bht159
- Müller, D., Winkler, I., Roeber, U., Schaffer, S., Czinger, I., and Schröger, E. (2010). Visual object representations can be formed outside the focus of voluntary attention: evidence from event-related brain potentials. *J. Cogn. Neurosci.* 22, 1179–1188. doi: 10.1162/jocn.2009.21271
- Näätänen, R., Sysoeva, O., and Takegata, R. (2004). Automatic time perception in the human brain for intervals ranging from milliseconds to seconds. *Psychophysiology* 41, 660–663. doi: 10.1111/j.1469-8986.2004.00182.x
- Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., and Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nat. Neurosci.* 4, 739–744. doi: 10.1038/89532
- Pelli, D. G., Palomares, M., and Majaj, N. J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection. *J. Vis.* 4, 1136–1169. doi: 10.1167/4.12.12
- Pelli, D. G., and Tillman, K. A. (2008). The uncrowded window of object recognition. *Nat. Neurosci.* 11, 1129–1135. doi: 10.1038/nn.2187
- Peng, C. H., Zhang, Y., Chen, Y. G., and Zhang, M. (2013). Crowded words can be processed semantically: evidence from an ERP study. *Int. J. Psychophysiol.* 88, 91–95. doi: 10.1016/j.ijpsycho.2013.03.002
- Pöder, E. (2006). Crowding, feature integration, and two kinds of “attention”. *J. Vis.* 6:7. doi: 10.1167/6.2.7
- Pöder, E. (2007). Effect of colour pop-out on the recognition of letters in crowding conditions. *Psychol. Res.* 71, 641–645. doi: 10.1007/s00426-006-0053-7
- Ronconi, L., Bertoni, S., and Bellacosa Marotti, R. (2016). The neural origins of visual crowding as revealed by event-related potentials and oscillatory dynamics. *Cortex* 79, 87–98. doi: 10.1016/j.cortex.2016.03.005
- Russell, C., and Driver, J. (2005). New indirect measures of “inattentive” visual grouping in a change-detection task. *Percept. Psychophys.* 67, 606–623. doi: 10.3758/BF03193518
- Saarela, T. P., Sayim, B., Westheimer, G., and Herzog, M. H. (2009). Global stimulus configuration modulates crowding. *J. Vis.* 9:5. doi: 10.1167/9.2.5
- Salti, M., Bar-Haim, Y., and Lamy, D. (2012). The P3 component of the ERP reflects conscious perception, not confidence. *Conscious. Cogn.* 21, 961–968. doi: 10.1016/j.concog.2012.01.012
- Sayim, B., Westheimer, G., and Herzog, M. H. (2010). Gestalt factors modulate basic spatial vision. *Psychol. Sci.* 21, 641–644. doi: 10.1177/0956797610368811
- Schadow, J., Lenz, D., Thaeig, S., Busch, N. A., Fründ, I., Rieger, J. W., et al. (2007). Stimulus intensity affects early sensory processing: visual contrast modulates evoked gamma-band activity in human EEG. *Int. J. Psychophysiol.* 66, 28–36. doi: 10.1016/j.ijpsycho.2007.05.010
- van den Berg, R., Roerdink, J. B., and Cornelissen, F. W. (2010). A neurophysiologically plausible population code model for feature integration explains visual crowding. *PLoS Comput. Biol.* 6:e1000646. doi: 10.1371/journal.pcbi.1000646
- Wei, J. H., Chan, T. C., and Luo, Y. J. (2002). A modified oddball paradigm “cross-modal delayed response” and the research on mismatch negativity. *Brain Res. Bull.* 57, 221–230. doi: 10.1016/S0361-9230(01)00742-0
- Westheimer, G., and Hauske, G. (1975). Temporal and spatial interference with vernier acuity. *Vision Res.* 15, 1137–1141. doi: 10.1016/0042-6989(75)90012-7
- Woodman, G. F., and Luck, S. J. (2003). Serial deployment of attention during visual search. *J. Exp. Psychol. Hum. Percept. Perform.* 29, 121–138. doi: 10.1167/1.3.103
- Woods, D. L., Alho, K., and Algazi, A. (1992). Intermodal selective attention. I. Effects on event-related potentials to lateralized auditory and visual stimuli. *Electroencephalogr. Clin. Neurophysiol.* 82, 341–355. doi: 10.1016/0013-4694(92)90004-2
- World Medical Association (2013). *WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects*. Available at: <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>
- Wu, Y., Chen, J., and Han, S. (2005). Neural mechanisms of attentional modulation of perceptual grouping by collinearity. *Neuroreport* 16, 567–570. doi: 10.1097/00001756-200504250-00010



- Yeh, S. L., He, S., and Cavanagh, P. (2012). Semantic priming from crowded words. *Psychol. Sci.* 23, 608–616. doi: 10.1177/0956797611434746
- Yeshurun, Y., and Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature* 396, 72–75. doi: 10.1038/23936
- Yeshurun, Y., and Rashal, E. (2010). Precueing attention to the target location diminishes crowding and reduces the critical distance. *J. Vis.* 10:16. doi: 10.1167/10.10.16
- Yong, K. X., Shakespeare, T. J., Cash, D., Henley, S. M., Nicholas, J. M., Ridgway, G. R., et al. (2014). Prominent effects and neural correlates of visual crowding in a neurodegenerative disease population. *Brain* 137, 3284–3299. doi: 10.1093/brain/awu293
- Zhang, W., and Luck, S. J. (2009). Feature-based attention modulates feedforward visual processing. *Nat. Neurosci.* 12, 24–25. doi: 10.1038/nn.2223
- Zhou, J., Lee, C. L., and Yeh, S. L. (2016). Word meanings survive visual crowding: evidence from ERPs. *Lang. Cogn. Neurosci.* 31, 1167–1177. doi: 10.1080/23273798.2016.1199812

**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 © 2018 Peng, Hu and Chen. 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) and the copyright owner(s) 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.



# Why Do Durations in Musical Rhythms Conform to Small Integer Ratios?

Andrea Ravnani<sup>1,2,3\*</sup>, Bill Thompson<sup>1,2</sup>, Massimo Lumaca<sup>4†</sup> and Manon Grube<sup>4†</sup>

<sup>1</sup> Language and Cognition Department, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands, <sup>2</sup> Artificial Intelligence Lab, Vrije Universiteit Brussel, Brussels, Belgium, <sup>3</sup> Research Department, Sealcentre Pieterburen, Pieterburen, Netherlands, <sup>4</sup> Department of Clinical Medicine, Center for Music in the Brain, Aarhus University, Aarhus, Denmark

## OPEN ACCESS

### Edited by:

Dezhong Yao,  
University of Electronic Science and  
Technology of China, China

### Reviewed by:

Hugo Merchant,  
Universidad Nacional Autónoma de  
México, Mexico  
Daya Shankar Gupta,  
Camden County College,  
United States

### \*Correspondence:

Andrea Ravnani  
andrea.ravnani@gmail.com

<sup>†</sup>These authors have contributed  
equally to this work

**Received:** 28 February 2018

**Accepted:** 01 October 2018

**Published:** 28 November 2018

### Citation:

Ravnani A, Thompson B, Lumaca M  
and Grube M (2018) Why Do  
Durations in Musical Rhythms  
Conform to Small Integer Ratios?  
*Front. Comput. Neurosci.* 12:86.  
doi: 10.3389/fncom.2018.00086

One curious aspect of human timing is the organization of rhythmic patterns in small integer ratios. Behavioral and neural research has shown that adjacent time intervals in rhythms tend to be perceived and reproduced as approximate fractions of small numbers (e.g., 3/2). Recent work on iterated learning and reproduction further supports this: given a randomly timed drum pattern to reproduce, participants subconsciously transform it toward small integer ratios. The mechanisms accounting for this “attractor” phenomenon are little understood, but might be explained by combining two theoretical frameworks from psychophysics. The scalar expectancy theory describes time interval perception and reproduction in terms of Weber’s law: just detectable durational differences equal a constant fraction of the reference duration. The notion of categorical perception emphasizes the tendency to perceive time intervals in categories, i.e., “short” vs. “long.” In this piece, we put forward the hypothesis that the integer-ratio bias in rhythm perception and production might arise from the interaction of the scalar property of timing with the categorical perception of time intervals, and that neurally it can plausibly be related to oscillatory activity. We support our integrative approach with mathematical derivations to formalize assumptions and provide testable predictions. We present equations to calculate durational ratios by: (i) parameterizing the relationship between durational categories, (ii) assuming a scalar timing constant, and (iii) specifying one (of K) category of ratios. Our derivations provide the basis for future computational, behavioral, and neurophysiological work to test our model.

**Keywords:** rhythm, music perception, scalar expectancy theory, neural oscillations, integer ratio

## INTEGER RATIOS AND MUSICAL RHYTHM

What are *small integer ratios*, and what makes integer-ratio rhythms special? A *ratio* between two inter-onset-intervals (IOIs) is the division between two, usually adjacent durations. *Integer ratios* can be written as a fraction: 1.5 equals 15/10 or 3/2, but  $\sqrt{2}$  for instance cannot be written as a fraction. An integer ratio is *small* if the result of the division can be written as a small integer number divided by another small integer number e.g., 2/3, but not 23/51 (Pikovsky et al., 2003; Strogatz, 2003).

A *rhythm*, by definition as used here, is a pattern of durations (London, 2004, p. 4) characterized by the succession of event onsets over time, in other words a series of IOIs.

Auditory rhythms with small integer ratios between IOIs are common in the world's music (Essens and Povel, 1985; Toussaint, 2013; Savage et al., 2015). Psychological and neural research suggests that small integer-ratio rhythms allow a more accurate internal representation (Essens, 1986; Sakai et al., 1999), improved deviance detection (Jones and Yee, 1997; Large and Jones, 1999), enhanced memory (Deutsch, 1986; Palmer and Krumhansl, 1990) and reproduction (Povel and Essens, 1985; Essens, 1986), and better synchronization (Patel et al., 2005). The distortion of near-integer ratios toward integer ones (or their harmonics) reported in behavioral (Fraisse, 1982) and neurophysiological studies (Mozz et al., 2013) further supports the idea of small ratios acting as “attractors” (Gupta and Chen, 2016). This idea has recently received support from studies of iterated learning and reproduction. When humans reproduce an initially randomly-timed rhythmic sequence, and this process is repeated in a cascade fashion within one or across several individuals, the sequence is subconsciously reshaped to be composed of IOIs related by small integer ratios (Figure 1A; c.f. Polak et al., 2016; Ravignani et al., 2016, 2018; Jacoby and McDermott, 2017).

Why do rhythms (i.e., patterns of durations) tend to exhibit small integer ratios? Why are humans drawn to rhythms with such a peculiar mathematical property, in both perception and production? Does this property reflect a special quirk of music perception and/or motor sequencing, or could it be explained by domain-general aspects of cognition? Can we explore these alternatives through mathematical formalism? Here, we explore mathematically the possibility that the human bias toward small integer ratios may be explained by a combination of scalar expectancy and categorical perception.

We begin by outlining the relevant classical frameworks for human timing, and go on to summarize the evidence in support of the small-integer ratio bias in rhythm perception. We then present our proposal linking the frameworks to the bias through mathematical formalisms. Specifically, we draw on the scalar property of time interval estimation to formulate a simple model of categorical perception that may result in an integer ratio bias (Figure 1), and link this to neural oscillations. We conclude by briefly discussing the merits and limitations of our model and outlining future goals.

## PSYCHOPHYSICAL AND OSCILLATORY APPROACHES

Two major theoretical approaches, among several, have been suggested to account for the mechanisms behind human timing (Wing and Kristofferson, 1973a,b; Getty, 1975; Meck, 1996; Church, 1999; Grondin, 2001, 2010; Mauk and Buonomano, 2004; Karmarkar and Buonomano, 2007; Ivry and Schlerf, 2008; Allman et al., 2014; Merker, 2014). The most influential and empirically tested psychoacoustic model is the “scalar expectancy theory” (Wearden, 1991; Allman and Meck, 2011). Psychophysical research shows that human timing often follows Weber's law (Bizo et al., 2006): the error for an interval duration being timed is proportional to the duration of that interval. One perception-based formulation states that the ratio between

the just-noticeable difference (JND) and the duration of a reference stimulus is constant across stimulus length (Grondin, 2001). In another formulation, the coefficient of variation (standard deviation divided by mean) in estimating durations is constant across durations (Figure 1D; Gibbon, 1977).

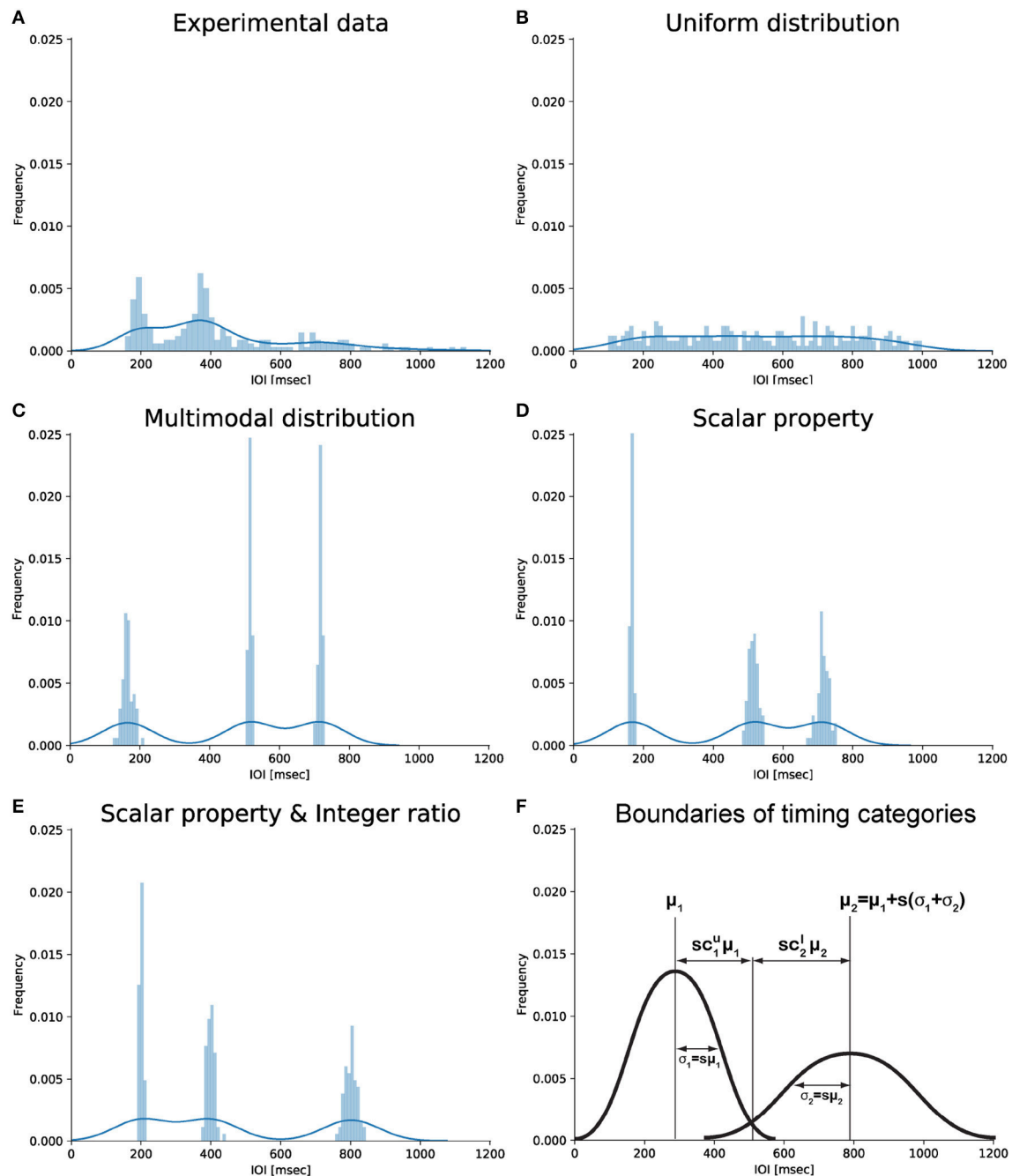
Another relevant approach to timing mechanisms comes from neuroscience and physics. It suggests that neural oscillations entrain (or even “resonate”) with the periodicity of external stimuli at multiple time-scales (Buzsaki, 2006; Large, 2008; Arnal and Giraud, 2012; Gupta, 2014; Aubanel et al., 2016; Celma-Miralles et al., 2016). Specifically, it states that phase and frequency of neural oscillations entrain with the phase and frequency of external events at multiple metrical levels. For instance, processing a metronome beat will induce low-frequency oscillations and/or power fluctuations in high-frequency oscillations following the periodicity of the beat, plus its multiples or divisors. Critically, the stability of the connection between two or more active neural oscillations, i.e., the “resistance” to external perturbations, depends on the ratio of their periods (e.g., 1:1, 2:1, 2:3). Small integer ratios typically confer greater stability. This may explain the perceptual advantage for integer-ratio stimuli over more complex metrical patterns (Large and Kolen, 1995). Other frameworks state that specific neurons or neural channels are tuned to particular durational intervals or tempi (Merchant et al., 2013; Bartolo et al., 2014).

## ITERATED DRUMMING EXPERIMENTS: SMALL INTEGER RATIOS AS COGNITIVE ATTRACTORS

Recent behavioral research investigated human priors for durations in rhythmic patterns (Ravignani et al., 2016, 2018; Jacoby and McDermott, 2017). Participants were given drumming sequences to reproduce to the best of their ability. The patterns produced were presented to the same or a new participant in an iterative procedure. Strikingly, “first-generation” participants were given completely random patterns, and “last-generation” participants produced rhythms exhibiting small integer ratios, in line with previous work on e.g., bimanual tapping (Peper et al., 1991, 1995a,b; Peper and Beek, 1998).

Specifically, participants were presented with sequences of IOIs sampled from a uniform distribution  $U$  (e.g., Figure 1B). As the patterns were transmitted through “chains of reproductions,” (Ravignani et al., 2016, 2018; Jacoby and McDermott, 2017), distribution  $U$  converged toward a distribution  $D$ : a human observer's posterior distribution of IOIs (e.g., Figure 1A). This distribution is multimodal, and the modes are related by small integer ratios, a universal property of human musical cultures (Ravignani et al., 2016; Jacoby and McDermott, 2017).

Here we aim to explain the distribution  $D$  via established psychophysical principles, none of which explicitly entail small-integer ratios. In other words, is the integer ratio bias a perceptual primitive in itself, or might it arise from the interaction of more



**FIGURE 1 |** Graphical representation of different types of IOI distributions. **(A)** Empirical distribution of drumming data showing two peaks (slightly below 200 and 400 ms) consistent with the notion of integer ratio categories. Data from the last experimental generation of chain 2 in Ravignani et al. (2016). **(B)** Uniform distribution from 100 to 1,000 ms. **(C)** Multimodal distribution based on 3 randomly chosen centroids without further assumptions. **(D)** Multimodal distribution around the same 3 centroids assuming the scalar timing property. **(E)** Multimodal distribution assuming the scalar timing property and showing small integer ratios. Data in panels **(B–E)** are simulated; they were randomly sampled from several normal distributions, with total sample size as in **(A)**. **(F)** Schematic representation of potential parameters linking scalar timing and small integer ratios. Panel **(F)** was produced without simulated or experimental data. Notice how the x-coordinate of the intersection point between the two Gaussians can be parameterized as to  $\mu_1 + sc_1^u \mu_1$  (first Gaussian) and  $\mu_2 - sc_2^l \mu_2$  (second Gaussian). For more than two Gaussians, the intersection can be parameterized as  $\mu_k + sc_k^u \mu_k$  (first Gaussian) and  $\mu_{k+1} - sc_{k+1}^l \mu_{k+1}$  (second Gaussian). This parameterization is used in the derivations below.

fundamental primitives? Jacoby and McDermott (2017) related a theoretically hypothesized prior with built-in integer ratios to an empirically estimated prior, showing that these were aligned.

Here, we investigate whether it is possible to derive a prior with similar properties by not building in the integer-ratio, but by combining empirically founded principles of timing with a



minimum of assumptions (and room for refinement by future testing).

## PROBABILISTIC INFERENCE FOR INTERVAL RATIO CATEGORIES

Our concrete question is: Under which conditions will a distribution  $G$  show small-integer ratios, without having built these ratios into our model?

Without any assumptions, distribution  $G$  would equal the uniform IOI distribution  $U$  *in expectation*. In other words which results on basic mechanisms of rhythm perception and production allow us to turn  $U$  into  $G$ ? Below, we make four assumptions based on psychophysical evidence and reduce the number of free parameters in the model drastically with little loss of generality. We begin by elaborating on previous formalizations to make relevant assumptions explicit and comparable.

### ASSUMPTION 1: CATEGORICAL TIMING

An  $n$ -event rhythm defines a sequence of IOIs  $\mathbf{d} = (d_1, \dots, d_{n-1})$  and of ratios  $\mathbf{r} = (r_1, \dots, r_{n-2})$ , such that  $r_i = d_{i+1}/d_i$ . Perception of a rhythm  $\mathbf{r}$  induces a representation  $\mathbf{z} = (z_1, \dots, z_{n-2})$ , with a strong tendency to categorize. The vector  $\mathbf{z}$  is a sequence of a small number of unique phenomenal interval-ratio *categories* that represent the observed data  $\mathbf{r}$ . More specifically, the notation  $z_i = k$  identifies that interval ratio  $r_i$  is attributed to phenomenal category  $k$  (Ravignani et al., 2018). Whilst not used explicitly in our calculations,  $\mathbf{z}$  formalizes the first key assumption: the processing of rhythmic sequences recruits a categorical interpretation of time intervals from a continuous stream of events (Clarke, 1987; Schulze, 1989; Desain and Honing, 2003). Behavioral evidence shows that also human motor timing is categorical: participants tapping produce IOI distributions with distinct peaks reflecting underlying durational *categories* (Collyer et al., 1994). This suggests that the distribution  $G$  can be approximated as a multimodal mixture of normal distributions (Figure 1C), rather than a uniform distribution (Figure 1B). A small number of durational categories naturally results in a small number of ratio categories. For the perception of a rhythmic sequence as a whole, we would argue that the perceived durations be transformed toward forming small ratios, as supported by iterated drumming experiments (Jacoby and McDermott, 2017), “ideally” into integer multiples of the smallest unit. Whilst categorical timing may appear to be a simplifying psychological concept (Schulze, 1989; Drake and Bertrand, 2001; Desain and Honing, 2003; ten Hoopen et al., 2006) based on behavioral observations, it may not be that far off neural reality. The notion of durational categories relate to basic durational tuning properties of premotor neurons recorded in non-human primates (Merchant et al., 2013). For instance, categories can be mapped to interval tuning in the premotor neurons of monkeys performing a synchronization continuation task (Merchant et al., 2013). Here, the distribution of preferred intervals could be viewed as a prior, although this distribution is multimodal, rather than bimodal as in Merchant et al. (2013). In addition, human

neuroimaging work showed specific activation patterns for the perceptual processing of integer interval ratios (Sakai et al., 1999). Moreover, sequences of small integer ratios may induce a metrical beat by the hierarchical organization of periodicity at two or more levels, i.e., the occurrence of an accent at a multiple small integer of the shortest time unit at the next higher level (Povel and Essens, 1985). Metrical structure is thus a higher, multi-level demonstration of the psychological prior toward small-integer ratios, that affords accurate reproduction (Povel and Essens, 1985). Moreover, the perceptual timing of rhythms with such a metrical beat is more accurate, their subjective percept “catchier” and their recognition more robust against temporal scaling, i.e., speeding up or slowing down the tempo, as the pattern is processed as one coherent whole rather than a series of time intervals, in contrast to rhythms that feature small integer ratios but no metrical beat (Grube and Griffiths, 2009).

### ASSUMPTION 2: BAYESIAN INFERENCE OVER GAUSSIAN CATEGORIES

A general assumption in rhythm research is that perceptual timing can be described as a process combining prior beliefs with sensory input. One way to capture this mathematically is to model time perception as Bayesian inference (Jazayeri and Shadlen, 2010; Cicchini et al., 2012; Merchant et al., 2013; Pérez and Merchant, 2018). Whilst our analysis relies on the nature of the prior rather than how it is deployed during perceptual interpretation, taking a Bayesian viewpoint is useful. It lets us express a prior distribution as an inductive bias (Thompson et al., 2016) and has been successfully applied in previous models of time interval estimation (e.g., Jazayeri and Shadlen, 2010; Cicchini et al., 2012). Employing Bayesian inference, we can characterize participant behavior as attributing a categorical representation to interval ratio  $r_i$  according to the distribution  $p(z_i = k | r_i) \propto p(r_i | z_i = k)p(z_i = k)$ . Our focus is the prior distribution over categories,  $p(z_i = k)$ , equivalently  $G$ . Alternatively, it would be possible to model learners’ assumptions about a likelihood distribution as a source of bias (e.g., Jazayeri and Shadlen, 2010; Cicchini et al., 2012).

Jacoby and McDermott (2017) recently modeled  $n$ -interval rhythms as single points in the  $n-1$  dimensional simplex, and formulated a multivariate-mixture prior over this space, assuming Gaussian models to underlie each of the mixtures. Namely, they formulated a multivariate  $p(\mathbf{z})$  directly. Our approach to the prior is closely related. Like Jacoby and McDermott (2017), we express the prior as a mixture of Gaussian components. However, our formulation treats an  $n$ -interval rhythm as a set of  $n-1$  independent samples from a *univariate multimodal distribution*, rather than a single multivariate sample. The two approaches essentially represent minor variants of the model for covariance of interval ratio categories. The assumption that the distribution  $p(\mathbf{z})$  has a Gaussian form should be tested in

future work, but is in line with existing work and a fair first approximation.

We write the prior as a  $K$ -dimensional Gaussian mixture of interval ratio categories, and the data likelihood as i.i.d. Gaussian underlying these categories, such that the marginal distribution of interval ratios has the form:

$$p(\mathbf{r}) = G(\mathbf{r}) = \prod_{i=1}^{n-1} \sum_{k=1}^K \varphi_k N(d_i; \mu_k, \sigma_k) \quad (1)$$

Here, the prior assigns to each Gaussian  $k = 1, \dots, K$  a weight in the mixture,  $\varphi_k$ , which determines its relative prominence as a category; a category mean  $\mu_k$ , which specifies the expected interval ratio underlying this category; and a category variance  $\sigma_k$ . The assumption we make is that weights are constant:  $\varphi_k = K^{-1}$  (corresponding to an equal number of observations in the Gaussians in **Figures 1C–E**). Whilst we hope to examine this assumption empirically in the future, we proceed under the most neutral assumption: no interval-ratio category is privileged.

### ASSUMPTION 3: A SMALL NUMBER OF SUB-SECOND CATEGORIES

Assuming that our indexing of categories under the prior is strictly ordered by the category means, such that  $\mu_j < \mu_k \Leftrightarrow j < k$ , we can immediately express our second empirical constraint on distribution  $G$ : only few categories exist (Desain and Honing, 2003; Motz et al., 2013; Ravignani et al., 2016, 2018).  $K$  is naturally limited by our approach to only model components for *small* integer ratios, and these are limited in number. Furthermore, we bound the range of category means  $\mu_k$  from 200 ms (London, 2004, p. 35) to 1,000 ms (Shaffer, 1983; Desain and Honing, 2003; Buhusi and Meck, 2005). This constraint limits  $K$  to the largest number of categories such that no category mean exceeds 1,000 ms:

$$K = \operatorname{argmax}_k \mu_k \text{ s.t. } \mu_k \leq 1000 \text{ for } k = 1, \dots, K. \quad (2)$$

### ASSUMPTION 4: SCALAR TIMING

So far, our assumptions constrain neither category means  $\mu_k$  nor standard deviations  $\sigma_k$ . Our final, perhaps most central assumption is that timing exhibits *scalar properties* in the sub-second time range considered here (Gibbon, 1977; Matell and Meck, 2000). Scalar timing drastically reduces the number of free parameters describing distribution  $G$ , by expressing category variances as a function of category means. The standard deviation of each category  $\sigma_k$  equals the mean  $\mu_k$  multiplied by a constant, dimensionless factor  $s$  (**Figure 1E**):

$$\sigma_k = s \mu_k. \quad (3)$$

Previous empirical reports estimated  $s$  to approximate 0.025 (Friberg and Sundberg, 1995; Madison and Merker, 2004).

## LINKING CATEGORICAL PERCEPTION AND SCALAR TIMING: HOW CLOSE CAN WE GET TO INTEGER RATIO INTERVALS?

All four assumptions are empirically based and independent of each other. Now,  $G$  can be further characterized by the degree of overlap between Gaussians composing the mixture. To formalize this, we assume each category  $k$  to intersect with its adjacent neighbors  $k-1$  and  $k+1$  at a distance proportional to  $c_k^l$  and  $c_k^u$  away from its mean  $\mu_k$  (**Figure 1F**), which is a constant proportion of the standard deviation  $\sigma_k$ .  $c_k^l$  and  $c_k^u$  parameterize the overlap between categories: they express how many standard deviations away from its mean  $\mu_k$  the cluster  $k$  intersects the cluster  $k+1$ , and how many standard deviations away from its mean  $\mu_{k+1}$  the cluster  $k+1$  intersects the cluster  $k$  (**Figure 1F** shows an example for  $k = 1, 2$ ).

Combining this idea of a parameterized overlap with scalar properties, each cluster  $k$  extends from  $\mu_k - s c_k^l \mu_k$  to  $\mu_k + s c_k^u \mu_k$ . Under these assumptions, the distance between the means of two adjacent distributions (**Figure 1F**) can be written as

$$\mu_{k+1} - \mu_k = s c_{k+1}^l \mu_{k+1} + s c_k^u \mu_k, \quad (4)$$

and their ratio as

$$r_k = \mu_{k+1} / \mu_k. \quad (5)$$

Substituting (5) into (4) provides

$$r_k \mu_k - \mu_k = s c_{k+1}^l r_k \mu_k + s c_k^u \mu_k, \quad (6)$$

which can be simplified and rewritten as

$$r_k = (1 + s c_k^u) / (1 - s c_{k+1}^l). \quad (7)$$

Equation (7) requires, to be well-defined, that its right side is positive, namely

$$0 < c_{k+1}^l < \frac{1}{s}. \quad (8)$$

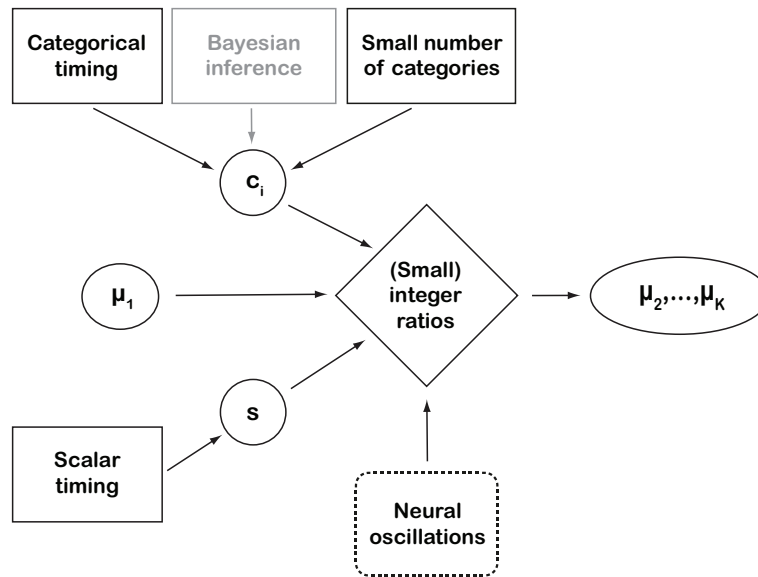
Operationally, the category means following from the constraints on  $G$  can be calculated using the recursion equation:

$$\mu_{k+1} = r_k \mu_k. \quad (9)$$

The constraints structure the space of component Gaussians in the prior such that, by specifying  $\mu_1$ , we can compute  $\mu_k$  for all  $k \leq K$  using Equation (9) (**Figure 1E**).

These quantitative tools enable the formulation of several questions. Given our *post-hoc* knowledge that the prior is characterized by categories centered at small integer ratios, do the constraints we laid out structure the prior such that integer-ratio clusters are predicted by setting  $\mu_1$  to the smallest possible integer ratio?

An alternative approach might be to assume that one ratio is e.g.,  $\frac{1}{2}$ , and ask whether our equations imply small integer



**FIGURE 2 |** Schematic representation of the perspective introduced by this paper. Black solid-line boxes represent empirically supported assumptions. “Bayesian inference” is outlined in gray to indicate that it is used here as a working assumption and conceptual framework, rather than an empirically supported assumption on cognitive processes (Shi et al., 2013). “Neural oscillations” are dashed because they represent observed neural process whose connection with the other behavioral concepts has not been proven (yet). The quantitative parameters are: category means  $\mu_i$ , a scalar constant  $s$ , and  $c_i$ , which is the abbreviation of  $c_i^u$  and  $c_i^l$ , parameterizing the overlap between categories. The proposed way of representing rhythmic structure depends, among other factors, on the constancy of  $r_k$  (see main text). A deviation from this constancy would result in larger integer ratios, with the deviation accumulating over the categories when iterating equation (8). Empirical work (e.g., Ravignani et al., 2016; Jacoby and McDermott, 2017) has tried to operationalize the connection between the “mathematical perfection” of integer ratios and their empirical counterpart in a number of alternative ways. This perspective does not address how and when a real number is perceived as an integer ratio, leaving this as an empirical question for psychophysics research. In general, large integer ratios, and even irrational-number ratios, can be perceived as small integer ratios if close enough to one. For instance,  $2^{7/12} \approx 1.498307$  is irrational (Coxeter, 1968) but close to  $3/2$ . Virtually all pianos, today, employ this irrational number (1.498307) in their well-tempered tuning, which is “close enough” for human hearing to the integer ratio  $3:2$ . At the same time, the “catchiness” of a rhythm also depends on small deviations from the integer ratios. For instance, delayed occurrences of expected beats even at varying levels of deviation from the underlying rhythms (together with the compensatory temporary speed-ups) are perceived as interesting, while a strictly regular rhythm will quickly appear dull.

ratios for the remaining cluster centers. More generally, do the constraints laid out impose an integer ratio structure on the prior without assuming an integer ratio for any of the clusters, simply by setting  $c_k$  in a certain way?

## HOW DO $c_k^u$ AND $c_k^l$ RELATE TO $\mu_k$ ?

The x-coordinates for the intersection point, expressed as  $\mu_k - sc_k^l \mu_k$  and  $\mu_k + sc_k^u \mu_k$ , can be substituted in the respective Gaussian probability density functions, equated to impose the condition of intersection on the y-axis (Figure 1F):

$$2 * \log(s\mu_k) + \frac{((\mu_k + sc_k^u \mu_k) - \mu_k)^2}{s^2 \mu_k^2} = 2 * \log(s\mu_{k+1}) + \frac{((\mu_{k+1} - sc_{k+1}^l \mu_{k+1}) - \mu_{k+1})^2}{s^2 \mu_{k+1}^2} \quad (10)$$

which simplifies as:

$$(c_k^u)^2 - (c_{k+1}^l)^2 = 2 \log(\mu_{k+1}/\mu_k). \quad (11)$$

Equation (11) means that the difference of squares between  $c$ 's is proportional to the logarithm of the ratio of the two means.

To make an example with actual numbers, if one substitutes  $\mu_k = \mu_1 = 100 \text{ ms}$  and  $\mu_{k+1} = \mu_2 = 200 \text{ ms}$  in (11), the equation becomes  $(c_k^u)^2 - (c_{k+1}^l)^2 = 2 \log(2)$ . Hence  $r_1 = \frac{\mu_{k+1}}{\mu_k} = 2$ ,  $c_1^u \approx 2.5$  and  $c_2^l \approx 2.2$  are two approximate solutions (among the infinite possible ones) of this particular example.

As the right side of Equation (11) is always strictly positive,  $c_k^u$  can never equal  $c_{k+1}^l$ . While this does not constitute a mathematical contradiction with our formulation (still leaving an infinite number of mathematically possible  $c$ 's), it is admittedly difficult to interpret psychophysically.

## SUGGESTED EXPERIMENTS: MODELING AND PSYCHOPHYSICS

Equations (7, 9) support a potential link between scalar timing and integer ratios, as they include the integer ratios  $r_k$  and the scalar constant  $s$  (Figure 2). These generative formulas can be implemented in computational simulations to explore the shape of the parameter space. Given specific values for parameters  $s$ ,  $c_k^u$  and  $c_k^l$ , the equations will return a unique set of ratios: are

these small integer ratios? Likewise, given one single integer ratio  $\mu_1$ , all other  $\mu_k$  are determined by Equation (9): which values of  $\mu_1$  result in  $r$  being integer ratios and  $s$ ,  $c_k^u$  and  $c_k^l$  being psychophysically plausible values?

The perspective we offer here creates the basis for expanding not only into theoretical but also empirical work on  $s$ ,  $c_k^u$  and  $c_k^l$ . Experimental research can advance this approach by estimating  $s$ ,  $c_k^u$  and  $c_k^l$  via Equation (7) or (11). Here, we treated the parameter  $s$  as an a priori known, one-valued constant ( $s = 0.025$ ). To improve the model further, the variance of  $s$  might be estimated by replications of previous psychophysical experiments such as those by Friberg and Sundberg (1995) and Madison and Merker (2004). Values for  $c_k^u$  and  $c_k^l$  can be estimated from experiments testing the perception (and misattribution) of durational categories.

## LIMITATIONS, DISCUSSION, AND CONCLUSIONS

We explore quantitative links between scalar timing and the human bias toward small integer ratios. The arguments we provide reduce the explanatory space to a few hypotheses. One possibility is that integer ratios are not a human cognitive primitive, but rather a simple by-product of other cognitive constraints, and their interaction.

Alternatively, the scalar timing framework might not be the most suitable one to explain the integer-ratio phenomenon of human rhythm. If one adopts oscillatory frameworks, integer ratios might simply arise from the oscillatory properties of brain activity, and so can scalar properties and categorical perception. Small integer ratios in particular would just reflect epiphenomena of harmonics of one oscillator or the interaction between two or more oscillators (Collyer et al., 1994; Strogatz, 2003; Buzsaki, 2006; Gupta, 2014; Merker, 2014; Gupta and Chen, 2016). Neural resonance to musical rhythm (Large, 2008), interval tuning

(Merchant et al., 2013; Bartolo et al., 2014), and population clocks (Crowe et al., 2014; Gouvêa et al., 2015; Bakhurin et al., 2016; Merchant and Averbeck, 2017) present alternative timing mechanisms, documented by *in-vivo* recordings of neural populations and compatible with the observed small integer bias.

In any case, scalar timing and oscillatory theories are simplifications, i.e., approximate descriptions derived from confined experimental set-ups. Neurally and behaviorally, the dissociation or compatibility between scalar timing and oscillatory theories is more complex than it may appear in higher level cognitive theories, and only detailed neural models will enable us to define the actual underlying mechanisms.

## AUTHOR CONTRIBUTIONS

AR and BT conceived the idea and performed the mathematical derivations. All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

## FUNDING

AR was supported by funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 665501 with the research Foundation Flanders (FWO) (Pegasus<sup>2</sup> Marie Curie fellowship 12N5517N awarded to AR). AR and BT were also supported by a visiting fellowship in Language Evolution from the Max Planck Society and ERC grant 283435 ABACUS (awarded to Bart de Boer).

## ACKNOWLEDGMENTS

We are grateful to the editor and the reviewers for their support and helpful comments on earlier versions of this manuscript.

## REFERENCES

- Allman, M. J., and Meck, W. H. (2011). Pathophysiological distortions in time perception and timed performance. *Brain* 135, 656–677. doi: 10.1093/brain/awr210
- 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
- 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
- Aubanel, V., Davis, C., and Kim, J. (2016). Exploring the role of brain oscillations in speech perception in noise: intelligibility of isochronously retimed speech. *Front. Hum. Neurosci.* 10:430. doi: 10.3389/fnhum.2016.00430
- Bakhurin, K. I., Mac, V., Golshani, P., and Masmanidis, S. C. (2016). Temporal correlations among functionally specialized striatal neural ensembles in reward-conditioned mice. *J. Neurophysiol.* 115, 1521–1532. doi: 10.1152/jn.01037.2015
- 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
- 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. Process.* 71, 201–210. doi: 10.1016/j.beproc.2005.11.006
- Buhusi, C. V., and Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6:755. doi: 10.1038/nrn1764
- Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press.
- Celma-Mirallas, A., de Menezes, R. F., and Toro, J. M. (2016). Look at the beat, feel the meter: top-down effects of meter induction on auditory and visual modalities. *Front. Hum. Neurosci.* 10:108. doi: 10.3389/fnhum.2016.00108
- Church, R. M. (1999). Evaluation of quantitative theories of timing. *J. Exp. Anal. Behav.* 71, 253–256. doi: 10.1901/jeab.1999.71.253
- Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M., and Burr, D. C. (2012). Optimal encoding of interval timing in expert percussionists. *J. Neurosci.* 32, 1056–1060. doi: 10.1523/JNEUROSCI.3411-11.2012
- Clarke, E. F. (1987). "Categorical rhythm perception: an ecological perspective," in *Action and Perception in Rhythm and Music*, ed A. Gabrielsson (Stockholm: Royal Swedish Academy of Music), 19–33.
- Collyer, C. E., Broadbent, H. A., and Church, R. M. (1994). Preferred rates of repetitive tapping and categorical time production. *Attent. Percept. Psychophys.* 55, 443–453. doi: 10.3758/BF03205301



- Coxeter, H. S. M. (1968). Music and mathematics. *Math. Teach.* 61, 312–320.
- 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
- Desain, P., and Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception* 32, 341–365. doi: 10.1068/p3370
- Deutsch, D. (1986). A musical paradox. *Music Percept.* 3, 275–280. doi: 10.2307/40285337
- Drake, C., and Bertrand, D. (2001). The quest for universals in temporal processing in music. *Ann. N. Y. Acad. Sci.* 930, 17–27. doi: 10.1111/j.1749-6632.2001.tb05722.x
- Essens, P. J. (1986). Hierarchical organization of temporal patterns. *Percept. Psychophys.* 40, 69–73. doi: 10.3758/BF03208185
- Essens, P. J., and Povel, D. (1985). Metrical and nonmetrical representations of temporal patterns. *Percept. Psychophys.* 37, 1–7. doi: 10.3758/BF03207132
- Fraisse, P. (1982). “Rhythm and tempo,” in *The Psychology of Music*, ed D. Deutsch (New York, NY: Academic Press), 149–180. doi: 10.1016/B978-0-12-213562-0.50010-3
- Friberg, A., and Sundberg, J. (1995). Time discrimination in a monotonic, isochronous sequence. *J. Acoust. Soc. Am.* 98, 2524–2531.
- Getty, D. J. (1975). Discrimination of short temporal intervals: a comparison of two models. *Percept. Psychophys.* 18, 1–8.
- 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
- Gouvêa, T. S., Monteiro, T., Motiwala, A., Soares, S., Machens, C., and Paton, J. J. (2015). Striatal dynamics explain duration judgments. *Elife* 4:e11386. doi: 10.7554/eLife.11386
- Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychol. Bull.* 127:22. doi: 10.1037/0033-2909.127.1.22
- Grondin, 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
- Grube, M., and Griffiths, T. D. (2009). Metricality-enhanced temporal encoding and the subjective perception of rhythmic sequences. *Cortex* 45, 72–79. doi: 10.1016/j.cortex.2008.01.006
- 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. (2016). Brain oscillations in perception, timing and action. *Curr. Opin. Behav. Sci.* 8, 161–166. doi: 10.1016/j.cobeha.2016.02.021
- 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
- Jacoby, N., and McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Curr. Biol.* 27, 359–370. doi: 10.1016/j.cub.2016.12.031
- Jazayeri, M., and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nat. Neurosci.* 13:1020. doi: 10.1038/nn.2590
- Jones, M. R., and Yee, W. (1997). Sensitivity to time change: the role of context and skill. *J. Exp. Psychol. Hum. Percept. Perform.* 23, 693–709. doi: 10.1037/0096-1523.23.3.693
- 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
- Large, E. W. (2008). Resonating to musical rhythm: theory and experiment. *Psychol. Time* 189–232. doi: 10.1016/B978-0-08046-977-5.00006-5
- Large, E. W., and Jones, M. R. (1999). The dynamics of attending: how we track time varying events. *Psychol. Rev.* 106, 119–159. doi: 10.1037/0033-295X.106.1.119
- Large, E. W., and Kolen, J. (1995). Resonance and the perception of musical meter. *Connect. Sci.* 6, 177–208. doi: 10.1080/09540099408915723
- London, J. (2004). *Hearing in Time*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195160819.001.0001
- Madison, G., and Merker, B. (2004). Human sensorimotor tracking of continuous subliminal deviations from isochrony. *Neurosci. Lett.* 370, 69–73. doi: 10.1016/j.neulet.2004.07.094
- 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
- Mauk, M. D., and Buonomano, D. V. (2004). The neural basis of temporal processing. *Annu. Rev. Neurosci.* 27, 307–340. doi: 10.1146/annurev.neuro.27.070203.144247
- Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Cogn. Brain Res.* 3, 227–242.
- Merchant, H., and Averbeck, B. B. (2017). The computational and neural basis of rhythmic timing in medial premotor cortex. *J. Neurosci.* 37, 4552–4564. doi: 10.1523/JNEUROSCI.0367-17.2017
- Merchant, H., Pérez, O., Zarco, W., and Gámez, 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
- Merker, B. (2014). Groove or swing as distributed rhythmic consonance: introducing the groove matrix. *Front. Hum. Neurosci.* 8:454. doi: 10.3389/fnhum.2014.00454
- Motz, B. A., Erickson, M. A., and Hetrick, W. P. (2013). To the beat of your own drum: cortical regularization of non-integer ratio rhythms toward metrical patterns. *Brain Cogn.* 81, 329–336. doi: 10.1016/j.bandc.2013.01.005
- Palmer, C., and Krumhansl, C. L. (1990). Mental representations for musical meter. *J. Exp. Psychol.* 16:728. doi: 10.1037/0096-1523.16.4.728
- Patel, A. D., Iversen, J. R., Chen, Y., and Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Exp. Brain Res.* 163, 226–238. doi: 10.1007/s00221-004-2159-8
- Peper, C. E., and Beek, P. J. (1998). Distinguishing between the effects of frequency and amplitude on interlimb coupling in tapping a 2:3 polyrhythm. *Exp. Brain Res.* 118, 78–92. doi: 10.1007/s002210050257
- Peper, C. E., Beek, P. J., and van Wieringen, P. C. (1995a). Frequency-induced phase transitions in bimanual tapping. *Biol. Cybernet.* 73, 301–309.
- Peper, C. E., Beek, P. J., and Van Wieringen, P. C. W. (1991). “Bifurcations in polyrhythmic tapping: in search of Farey principles,” in *Tutorials in Motor Neuroscience*, eds J. Requin and G. E. Stelmach (Dordrecht: Springer), 413–431.
- Peper, C. E., Beek, P. J., and van Wieringen, P. C. (1995b). Coupling strength in tapping a 2:3 polyrhythm. *Hum. Mov. Sci.* 14, 217–245.
- Pérez, O., and Merchant, H. (2018). The synaptic properties of cells define the hallmarks of interval timing in a recurrent neural network. *J. Neurosci.* 38, 4186–4199. doi: 10.1523/JNEUROSCI.2651-17.2018
- Pikovsky, A., Rosenblum, M., and Kurths, J. (2003). *Synchronization: A Universal Concept in Nonlinear Sciences*, Vol. 12. Cambridge: Cambridge University Press.
- Polak, R., London, J., and Jacoby, N. (2016). Both isochronous and non-isochronous metrical subdivision afford precise and stable ensemble entrainment: a corpus study of malian jembe drumming. *Front. Neurosci.* 10:285. doi: 10.3389/fnins.2016.00285
- Povel, D. J., and Essens, P. (1985). Perception of temporal patterns. *Music Percept.* 2, 411–440. doi: 10.2307/40285311
- Ravignani, A., Delgado, T., and Kirby, S. (2016). Musical evolution in the lab exhibits rhythmic universals. *Nat. Hum. Behav.* 1:0007. doi: 10.1038/s41562-016-0007
- Ravignani, A., Thompson, B., Grossi, T., Delgado, T., and Kirby, S. (2018). Evolving building blocks of rhythm: how human cognition creates music via cultural transmission. *Ann. N.Y. Acad. Sci.* 1423, 176–187. doi: 10.1111/nyas.13610
- Sakai, K., Hikosaka, O., Miyauchi, S., Takino, R., Tamada, T., Iwata, N. K., et al. (1999). Neural representation of a rhythm depends on its interval ratio. *J. Neurosci.* 19, 10074–10081. doi: 10.1523/JNEUROSCI.19-22-10074.1999
- Savage, P. E., Brown, S., Sakai, E., and Currie, T. E. (2015). Statistical universals reveal the structures and functions of human music. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8987–8992. doi: 10.1073/pnas.1414495112
- Schulze, H. H. (1989). Categorical perception of rhythmic patterns. *Psychol. Res.* 51, 10–15. doi: 10.1007/BF00309270
- Shaffer, L. H. (1983). Timing in musical performance. *Ann. N. Y. Acad. Sci.* 423, 420–428.

- Shi, Z., Church, R. M., and Meck, W. H. (2013). Bayesian optimization of time perception. *Trends in Cognitive Sciences*, 17, 556–564. doi: 10.1016/j.tics.2013.09.009
- Strogatz, S. H. (2003). *Sync: The Emerging Science of Spontaneous Order*. New York, NY: Hyperion Books.
- ten Hoopen, G., Sasaki, T., Nakajima, Y., Remijn, G., Massier, B., Rhebergen, K. S., et al. (2006). Time-shrinking and categorical temporal ratio perception: evidence for a 1: 1 temporal category. *Music Percept.* 24, 1–22. doi: 10.1525/mp.2006.24.1.1
- Thompson, B., Kirby, S., and Smith, K. (2016). Culture shapes the evolution of cognition. *Proc. Natl. Acad. Sci. U.S.A.* 113, 4530–4535. doi: 10.1073/pnas.1523631113
- Toussaint, G. T. (2013). *The Geometry of Musical Rhythm: What Makes a Rhythm “Good”?* New York, NY: Chapman and Hall; CRC Press.
- Wearden, J. H. (1991). Do humans possess an internal clock with scalar timing properties? *Learn. Motiv.* 22, 59–83. doi: 10.1016/0023-9690(91)90017-3
- Wing, A. M., and Kristofferson, A. B. (1973a). Response delays and the timing of discrete motor responses. *Atten. Percept. Psychophys.* 14, 5–12.
- Wing, A. M., and Kristofferson, A. B. (1973b). The timing of interresponse intervals. *Atten. Percept. Psychophys.* 13, 455–460.
- 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 © 2018 Ravignani, Thompson, Lumaca and Grube. 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) and the copyright owner(s) 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.



# The Musical Structure of Time in the Brain: Repetition, Rhythm, and Harmony in fMRI During Rest and Passive Movie Viewing

Dan Lloyd\*

Department of Philosophy and Program in Neuroscience, Trinity College, Hartford, CT, United States

## OPEN ACCESS

### Edited by:

Daya Shankar Gupta,  
Camden County College,  
United States

### Reviewed by:

David Papo,  
Université de Lille, France  
Xi-Nian Zuo,  
Institute of Psychology, Chinese  
Academy of Sciences, China

### \*Correspondence:

Dan Lloyd  
Dan.Lloyd@trincoll.edu

**Received:** 21 June 2018

**Accepted:** 23 December 2019

**Published:** 21 January 2020

### Citation:

Lloyd D (2020) The Musical  
Structure of Time in the Brain:  
Repetition, Rhythm, and Harmony  
in fMRI During Rest and Passive  
Movie Viewing.  
Front. Comput. Neurosci. 13:98.  
doi: 10.3389/fncom.2019.00098

Space generally overshadows time in the construction of theories in cognitive neuroscience. In this paper, we pivot from the spatial axes to the temporal, analyzing fMRI image series to reveal structures in time rather than space. To determine affinities among global brain patterns at different times, core concepts in network analysis (derived from graph theory) were applied temporally, as relations among brain images at every time point during an fMRI scanning epoch. To explore the temporal structures observed through this adaptation of network analysis, data from 180 subjects in the Human Connectome Project were examined, during two experimental conditions: passive movie viewing and rest. The temporal brain, like the spatial brain, exhibits a modular structure, where “modules” are intermittent (distributed in time). These temporal entities are here referred to as themes. Short sequences of themes – motifs – were studied in sequences from 4 to 11 s in length. Many motifs repeated at constant intervals, and are therefore rhythmic; rhythms, converted to frequencies, were often harmonic. We speculate that the structure and interaction of these global oscillations underwrites the capacity to experience and navigate a world which is both recognizably stable and noticeably changing at every moment – a temporal world. In its temporal structure, this brain-constituted world resembles music.

**Keywords:** fMRI, graph theory, time, temporality, oscillation, rhythm, harmony, music

## INTRODUCTION

In science, “ontology” denotes the determination of the relevant categories and objects available to observation and hypothesis formation (Chakravartty, 2017). Historically, scientific ontologies have divided space first, and only later, time. This is vivid in neuroscience: From the Latin-fluent anatomists to Brodmann to the physical Connectome (Sporns et al., 2005; Hagmann et al., 2007), neuroscience continues to deploy a rich spatial ontology. Temporal ontologies in neuroscience are comparatively recent, but equally rich. An exemplary building block of temporal ontology rests on the exploration of oscillatory neural signals. Fourier analysis affords a powerful descriptive vocabulary which has been abundantly employed in neuroscience. (The references are too many to list in one place, but will appear throughout this paper). The Fourier Transform (FT) has moreover inspired a family of wavelet transforms (WT). These techniques have been abundantly exploited in the study of the temporal brain (see section “Oscillation, Information Broadcasting, and Maintenance” for discussion). However, does the FT/WT framework provide a complete

temporal ontology for the brain? Are there other temporal structures, beyond the FT/WT package, that are detectable in brain image series? Here, we will borrow a few basic concepts from a domain where time is centrally important: music. Many musical concepts are essentially temporal, involving order, duration, and temporal relationships in their definition. Centuries of music theory and musicology (including cognitive musicology) afford reasonable criteria for their measurement (Sethares, 1998, 2007; Lloyd, 2011; Huron, 2014). Might they or their analogs apply to neuroimaging time series?

New departures in method necessarily involve exploratory data analysis, and new ontologies particularly involve starting afresh. There are therefore many possible directions for this paper to take (Hutchison et al., 2013; Allen et al., 2014; Calhoun et al., 2014; Kopell et al., 2014). Some plausible starting points and assumptions are highly negotiable. Their motivation will be reviewed in the methods section, but are also discussed in sections “What now? What’s next?”; “Oscillation, Information Broadcasting and Maintenance”; “Rhythm”; and “Harmony”. Here, we first develop a “temporal parcellation” of the imaging data to be examined. That is, we extract a preliminary differentiation of the temporal landscape, just as spatially rooted dynamics rests on a spatial parcellation as its foundation. Graph theory is one method (among many) that can be applied here, but with a pivot from space to time. In its spatial (standard) application, graph theory begins with a set of spatially discrete entities, called nodes (brain regions, usually), and some measure of linkage (edges) between them (often correlation) (Sporns et al., 2005; Hagmann et al., 2007; Bullmore and Sporns, 2009; Sporns, 2011a,b; cf. Stanley et al., 2013; Fallani et al., 2014). The links are thresholded in order to define adjacency among the nodes. Working from the adjacency graph, various communities of nodes can be distinguished, along with other network properties of interest. These go by different names: modules, networks, communities, clubs, or cliques, among others. They can be defined in various ways, but one common measure of modularity discovers groups that have many interconnections among the nodes within the group, but only sparse connections between the groups (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010; Sporns, 2011a, 2012).

Network analyses have usually been employed spatially or spatiotemporally. Spatial network analyses begin with spatially delineated regions (nodes). The time series of activation at all nodes are correlated over the entire time course of an experimental condition for one or more subjects, forming the basis of the resulting graph or network. Spatiotemporal modularity posits that the functional relationships among nodes are variable over time. For example, the same correlational measure might be applied along a sliding temporal window to generate a sequence of modular parcellations, a dynamic functional connectome (Hutchison et al., 2013; Allen et al., 2014; Calhoun et al., 2014). This temporal sensitivity nonetheless rests on an initial spatial parcellation or a sequence of spatial parcellations.

In contrast to both these applications, in this study the graph-theoretic analysis is exclusively based on temporal features in data. Or in other words, there is no initial spatial segregation;

the region of interest is simply the brain in its entirety, and the similarity measures are applied exclusively along the spatial dimensions. Instead of regional/spatial nodes, the foundational entities are temporal, namely, individual whole brain images, captured via fMRI, at each moment in time in the series of images. These fully temporal “nodes” might well be called “moments.”

Despite the application of graph theory, the complete pivot toward time translates the spatial concepts inherent in graph theory as spatial metaphors for relations in time. Adjacency among moments is measured by their spatial correlation (across all the voxels of each image, compared image to image), rather than temporal correlation of time series recorded at spatially distinct sites. By this measure moments that are separated by long intervals might nonetheless be adjacent. The equivalent of a module, then, might be distributed in time, and such modules might interweave; the spatial connotations of the term “module” is misleading in this context. We propose instead to refer to these collections of correlated moments as themes. One theme might be present for a single uninterrupted interval, or it may be distributed temporally among other themes. In short, a theme comprises timepoints where patterns of global brain activity are similar, and divisions among themes are determined by the modularity algorithm. The sequence of thematic instantiations comprises an overall thematic profile of an image series. Subsets of the overall thematic profile, short continuous sequences of thematic instantiations, are motifs. [Note that this usage differs from “motif” in network analysis (Milo et al., 2002)] The strategy for this analysis is sketched in **Figure 1**, with a division between a space-first parcellation (A) and a temporal parcellation (B).

This complete pivot toward time foregrounds the structure of time in the brain, without assumptions about spatial structure. The first and fundamental question, then, is simply: Is there temporal structure in fMR image series? This will offer a data-driven ontology of temporality in the brain. It will be useful, however, only if we can meaningfully describe observed temporal organization. There are many paths to follow, some of them to be discussed in specific contrasts with the methods here; the path in this paper, toward quasi-musical properties, does not exclude other approaches. If there is an anatomy of time, this global temporal parcellation can then provide a data-driven clue to the spatial divisions most relevant to the dynamic functional connectome (Zuo et al., 2010).

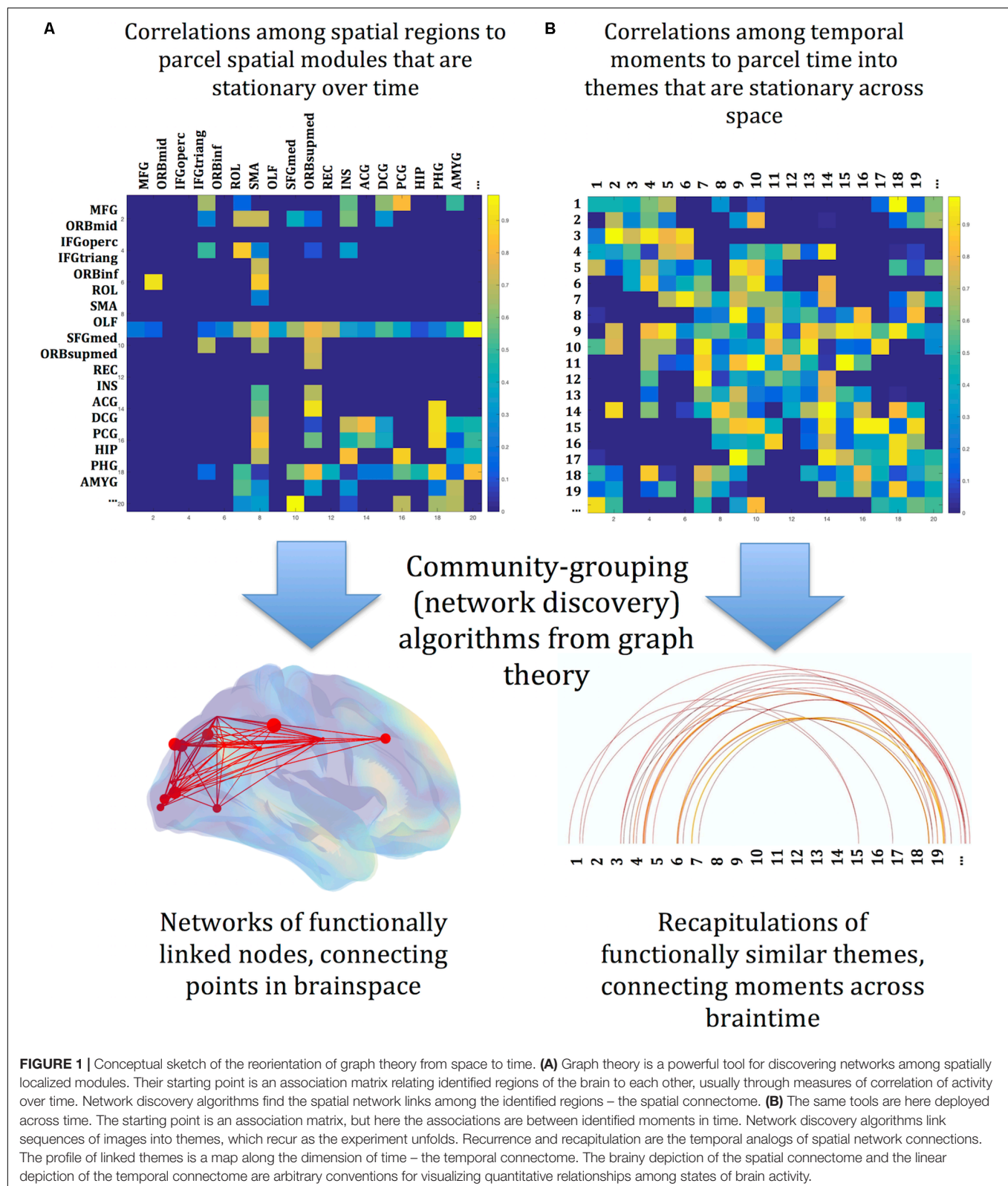
## MATERIALS AND METHODS

For this analysis, data from 180 adult subjects (ages 20–35, 108 Female) in two scanning conditions were downloaded from the Human Connectome Project 1200 Subjects Data Release, May 2018 (HCP<sup>1</sup>; data repository<sup>2</sup>; Marcus et al., 2011; Van Essen et al., 2013; Hodge et al., 2016). Subjects were scanned on a Siemens MAGNETOM 7T MR scanner housed at the Center for Magnetic Resonance (CMRR) at the University of Minnesota in

<sup>1</sup><https://www.humanconnectome.org/>

<sup>2</sup><https://db.humanconnectome.org>





Minneapolis, MN, using a Nova32 32-channel Siemens receiver head coil. Whole-brain sequence gradient-echo EPI images were acquired with the following parameters: TR 1000 ms; TE 22.2 ms;

flip angle 45 degrees; FOV 208 × 208 mm (RO × PE); Matrix 130 × 130 (RO × PE); Slice thickness 1.6 mm; 85 slices; 1.6 mm isotropic voxels; Multiband factor 5; Image Acceleration

factor(iPAT) 2; Partial Fourier (pF)sampling 7/8; Echo spacing 0.64 ms; BW 1924 Hz/Px<sup>3</sup>.

The two conditions studied here are: (1) resting with eyes open, 900 images (15 min), using the first of four imaging sessions with each subject, and (2) passive movie viewing, 900 images, using the first of four imaging sessions in that condition (Smith et al., 2013). The audio/visual movie was a compilation of short excerpts from Vimeo videos available under Creative Commons licensing<sup>3</sup>.

## Preprocessing

Images were preprocessed using HCP minimal preprocessing pipelines (Glasser et al., 2013; Marcus et al., 2013). This includes three structural pipelines: PreFreeSurfer, to create an aligned and undistorted structural volume space for each subject and register subjects' spaces to MNI space; FreeSurfer, to parcel volumes into predefined structures, reconstruct white and pial cortical surfaces, and register surfaces to FreeSurfer's surface atlas, fsaverage; PostFreeSurfer, performing individual surface registration using multimodal surface matching (MSM), based on areal features including sulcal depth, myelin, and functional connectivity maps.

Then, the fMRI Surface pipeline mapped the cortical gray matter voxels onto cortical surface vertices, and subcortical volume voxels, to standardize the surface and subcortical "grayordinate" space for all subjects. [Among other advantages, mapping via surface vertices greatly reduces the number of datapoints needed to express 7T images, making further computational analysis feasible (Glasser et al., 2013).] These data were smoothed with a surface algorithm to 2 mm FWHM (Glasser et al., 2013). Finally, 7T rfMRI 4D volume and grayordinate (surface vertices + subcortical voxels) Data, and 7T movie data, were further preprocessed to remove structured artifacts using FSL's FIX (FMRIB's ICA-based Xnoiseifier, Salimi-Khorshidi et al., 2014). For details, see Glasser et al. (2013, 2016) and the HCP 1200 Subjects Data Release Reference Manual<sup>3</sup>. The preprocessed data sets were downloaded from HCP during May 2018.

## Graphs

Association matrices were constructed using partial correlations between all voxels in each image series, while controlling for their mean activation (Marrelec et al., 2007; Smith et al., 2011; Hutchison et al., 2013; Varoquaux and Craddock, 2013; Epskamp and Fried, 2018). Each matrix was converted to binary (undirected) connection graph by thresholding the matrix to preserve the top 5% of inter-node values (Rubinov and Sporns, 2010). These binary association matrices were the basis for all subsequent analyses. In traditional network analysis, these steps would be applied to spatially distinct regions (usually anatomically defined regions of interest). Here, the entities to be linked are not spatial but temporal. Instead of measuring correlations of time series between regions, we measure correlations of spatial patterns between time points. The

overall analysis is sketched in **Figure 1**, with a division between a space-first parcellation (A) and a temporal parcellation (B). Each whole-brain image is a node (a "moment") in a temporally connected network, to be analyzed with graph theoretic methods. All the measures described herein were assessed for significance by contrast with baseline random images that preserve the degree distribution of the original data ("null-hypothesis networks," Rubinov and Sporns, 2010). That is, the distribution of the numbers of links originating from nodes remains constant while the pairings of the connection matrix are varied randomly (Maslov and Sneppen, 2002). The randomizing function and other functions are found in the Brain Connectivity Toolbox (BCT<sup>4</sup> Rubinov and Sporns, 2010). In all cases except where noted 100 baseline association matrices were generated for each subject to be contrasted, subject by subject, with the actual data. Appropriate corrections for multiple comparisons were applied using the Bonferroni method.

## Theme and Modularity

Modularity in network analysis denotes the subdivision of network nodes into non-overlapping groups where similarity is greatest within the group, and minimized between the groups (Rubinov and Sporns, 2010). The temporal analog of the module is the theme, which comprises moments of similarity among the full brain images, as assessed through the association matrix described above. A temporal theme is conceptually quite different from a spatial module; nonetheless various measures of modularity can be applied. Modularity was measured using two standard measures (Girvan and Newman, 2002; Newman, 2006; Reichardt and Bornholdt, 2006; Blondel et al., 2008), also implemented through BCT. Modularity assessments with these two algorithms were very similar, and so further analysis was based in the Louvain group method of Blondel et al. (2008). In addition to providing the optimal group divisions, the modularity algorithms calculate a statistic quantifying the degree to which the network can be subdivided into groups with high ingroup similarity and low outgroup similarity (Kaiser, 2011). We compare this statistic for each subject with the same statistic generated for 18,000 surrogate association matrices (100 per subject). Similar analyses, including baseline contrasts, were conducted for both the rest and movie conditions.

This method is conceptually similar to several recent proposals for deriving network architecture from time series data (Voss et al., 2004; Lacasa et al., 2008; Shirazi et al., 2009). Lacasa et al. (2008) describe a "visibility graph" from which some of the properties of classical graph theory can be observed, along with scaling properties. Shirazi et al. (2009) regard particular binned values in a time series as node identifications, and represent the transitions from each time point to the next as an internodal link with probabilities derived from the original series (Shirazi et al., 2009). These methods begin with a time series of a single variable. The target in this paper is distinct in two ways. First, each moment in the fMRI time series is a vector of ~96,000 continuous variables, while the methods just mentioned begin from time series of one variable. More

<sup>3</sup>[https://www.humanconnectome.org/storage/app/media/documentation/s1200/HCP\\_S1200\\_Release\\_Reference\\_Manual.pdf](https://www.humanconnectome.org/storage/app/media/documentation/s1200/HCP_S1200_Release_Reference_Manual.pdf)

<sup>4</sup><https://sites.google.com/site/bctnet/>

important, in the application here the goal is not the recovery of a physical network, but rather a compact representation of dynamical (temporal) properties of the image series. We move from one temporal series to another, in effect reducing the dimensionality of time series data. (Accordingly, other methods for dimensionality reduction could also be applied, e.g., Principal component analysis, Independent component analysis, Cluster analysis, among others). The application of graph theory probes the sequence of images as an oscillation among distinct themes (continuing with fully temporal descriptive language).

## Thematic (Temporal) Profiles

Following the application of these techniques, we can consider time series as dynamic temporal profiles, rather than as the expression of fixed spatial networks. Methods going forward, thus, are somewhat novel. This section will introduce them, along with their rationale.

We consider three features of the data sets: repetition, rhythm, and harmony. These are nested: Where repetitions recur with a constant interval between them, there is rhythm. Where there are multiple rhythms and the frequencies of rhythmic repetitions are in integer relations to one another, there is harmony.

### Repetition

The modular analysis decomposed the time series of global brain activity patterns into a small number of themes, reducing the experiment to a sequence of themes with various durations and alternations. We examined short image sequences (motifs) to see if particular sequences repeat over the image series. We considered sequences of length  $n$ , where  $n$  ranged from 4 to 11 s. Each sequence (1: $n$ , 2: $n + 1$ , etc.) was compared to all other sequences in each subject's thematic profile, and exact matches counted.

In the initial analysis, these counts were compared with similar analyses of thematic profiles derived from 100 surrogate association matrices for each subject, as described above. For each sequence length (from 4 to 11 s), repetitions in the data were compared with 100 surrogate data profiles using a one-tailed  $t$ -test, corrected for multiple comparisons. Then these measures were compared between the two stimulus conditions, rest and movie viewing.

### Rhythm

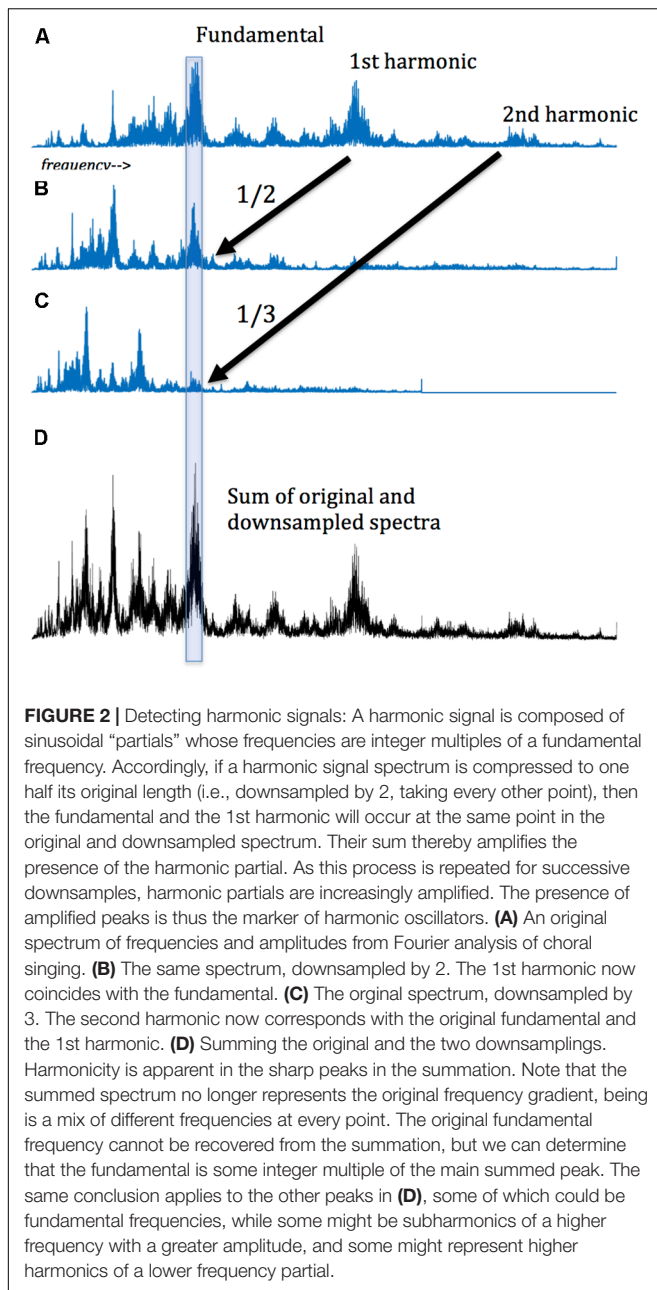
Even in domains where rhythm seems apparent, like music, an algorithm for rhythm detection can be elusive – computers can find the beat only very imperfectly (Sethares, 2007). The noisy signals of fMRI are even more difficult, compounded by the absence of intuitive or perceptible rhythms in the data. Here we deploy a continuous measure of “rhythmicity,” understood as the tendency for events that repeat to be separated by a constant interval. It follows from the analysis of repetition, just described. As the repetitions were counted, the intervals between each occurrence of any repeating sequence (motif) as each length (from 4 to 11 images) were also recorded. Where motifs recur more than twice, we compare the intervals between repetitions. If these intervals are equal, then that motif is recurring rhythmically. The presence of

these congruent intervals as a proportion of all intervals then serves as one index of rhythmicity. Note that as certain intervals recur more frequently, then other intervals become relatively more rare. This tilt toward rhythmicity is therefore reflected in the standard deviation of the set of numbers of occurrences of each interval. This value can then be compared to standard deviations of random surrogates derived from null network variations generated for each subject. These values then are compared across the two experimental conditions.

### Harmony

The search for harmony rests on a more tentative approach. A harmonic signal essentially comprises power at a fundamental frequency and/or at integer multiples of that fundamental. Together these higher frequencies form the harmonics of the fundamental. (These are also called overtones or partials). A signal with this spectral structure, then, is harmonic. In principle, harmonicity is easy to detect: the peak amplitude frequencies should be separated in frequency by constant differences. However, with these data neither the fundamental nor the harmonic frequencies are known, and certainly not apparent from the noisy Fourier spectrum. Instead, we exploit the rhythm information just collected: the repeating intervals are easily converted into frequencies, and thus the histogram of intervals transforms into the histogram of frequencies. In effect, this is an alternative form of signal spectrum (not based in Fourier analysis), where numbers of occurrences of each interval converts to amplitude at each frequency.

Then, we adapted the amplitude spectra to amplify the hidden harmonics. Specifically, the spectrum for each subject was downsampled by factors of 2 through 7, and the resultant vectors added to the original. (Downsampling decreases the sampling rate by integer factors. For example, a vector downsampled by a factor of two comprises every second element of the original). The downsampling of a harmonic signal spectrum preserves peaks at the same point in each downsampling. (In effect, each harmonic peak is moved left by an integer factor, so lower and higher harmonics coincide). This is shown schematically in **Figure 2**. Adding the original and downsampled vectors amplifies the magnitude of the harmonics. This method is similar to “harmonic product spectrum” methods for pitch detection (Cuadra et al., 2001), with the difference that here we sum, rather than multiply, the downsampled vectors. Using this method a maximum value for the summed vectors (original and downsampled transforms) can be calculated. (The position in the vector of this maximum is often interpreted as the fundamental frequency, but this is not necessary for the present analysis). Here we are interested in the maximum magnitude of this compounded amplitude. We identify the presence of harmonics by comparing amplitude at each point in the summed/downsampled vectors to similar points in 100 surrogate datasets, using a one-tailed  $t$ -test, corrected for multiple comparisons (Bonferroni method). The technique is applied to sequences of each tested length (4 through 11 images). In effect, this analysis considers repeating sequences as oscillators, and groups oscillators by the length of the sequences that repeat. Thus, we cast the net broadly in the hopes that harmonic



oscillation can emerge from the background of the inharmonic. These values were compared in the two experimental conditions.

Several standard approaches are the obvious foils: These derive from Fourier analysis and include Wavelet decomposition and measurements of phase synchrony. Fourier analysis construes a signal as a superposition of sinusoids at various frequencies and phases. Or in other words, the basis set for Fourier analysis is a series of sine/cosine functions. Periodic signals are built on that basis, stretched and slipped along the time axis. The second approach takes a brief basis function (the “mother wavelet”), stretches it to various lengths, and matches it to the target signal at every time point. The wavelet decomposition thus construes the signal as a moment by moment superposition of

the chosen basis functions, something like a short-segmented or “windowed” Fourier analysis. Phase synchrony between two signals is calculated from instantaneous phase measurements (Lachaux et al., 1999; Varela et al., 2001; Laird et al., 2002; Glerean et al., 2012).

The methods in this study, in contrast to Fourier and Wavelet approaches, make fewer assumptions about the basis function to be tested against the target signal (in this case, a thematic profile). We use segments of the thematic profile itself as basis functions, and analyze the entire signal against each of the thematic profile segments. Thus, every short sequence extracted from the signal is tested along a sliding window as a potential basis function for the whole signal itself. Thus, multiple basis functions are tested, derived from the data itself. For each sequence, we measure its repetition, and from the timing of repetitions we calculate rhythm and harmony. As in the initial parcellation, significance is tested via the permutations derived from the null networks. In this study, the segments tested ranged from four to eleven images (seconds). This window was selected because, in general, the analysis is computationally intensive, requiring some selectivity in what can be feasibly explored. Sequences of shorter than 4 s repeated densely in both the data and the permutations, rendering the comparison moot. At greater than 11 s, repetitions occurred only rarely, attenuating the comparison with the null permutations.

The measurement of rhythm follows a similar strategy, namely, examining the intervals between repetitions of repeating sequences. That periodicity in turn directly determines frequency of the revealed rhythms. These can be tested for harmonic relations, as described above. In general, then, the methods here are both open-ended and data-driven, resting on the sequences that occur in the data, and therefore afford more opportunities for discovering temporal regularities even if transient. In contrast, both Fourier analysis and Wavelet analysis make assumptions about the basis functions. For the FT this is of course the sine function; Wavelets can have many different shapes, but in all cases the analyst specifies the wavelet prior to the analysis. Arguably these assumptions could miss regularities that the methods here might detect. On the other hand both of the standard methods use basis functions stretched to various scales, and so in this sense are more receptive to regularities at multiple scales than the more constrained methods of this study.

Phase synchrony is a powerful marker of functional relatedness (Varela et al., 2001; Laird et al., 2002; Buzsáki and Draguhn, 2004; Glerean et al., 2012; Gupta and Chen, 2016; Gravel et al., 2018). However, it too rests on an *a priori* decision, namely the band-pass filtering of the target signals, necessary for determining instantaneous phase. The methods in the current study identify a range of frequencies involved in multiple rhythms and thus multiple harmonic relationships. Once again, the data is doing the driving.

One final rationale for the methods here is the analogy with music. The concepts of rhythm and harmony employed here are strict analogs of standard usage in musicology. If there is something to be made of the comparison of brain activation and music, these are among the measures we would hope might apply (Lloyd, 2011, 2013).



In summary, the stages of the analysis are these: The starting point is the full pattern of all voxels for each of 900 brain images (for each subject, in two experimental conditions). These are grouped according to a modularity algorithm into themes. The resultant vector, or thematic profile, is the basis for subsequent analysis. To measure repetition, we examine short excerpts or sequences (motifs) drawn from each thematic profile, separately considering all sequences from 4 to 11 s in length, counting the number of repeated occurrences of each sequence. To measure rhythm, we examine the intervals between repetitions of repeating sequences, counting the number of times specific inter-sequence intervals occur. To measure harmonicity, we examine the frequency of rhythmic repetitions, using the downsampling procedure described above to identify integer ratios between frequencies.

## RESULTS

In this study, we observe the presence of the temporal counterpart to modularity (see **Table 1**). Within and across subjects, two widely used measures of modularity agreed that the temporal connectome has a modular architecture. Following the Louvain group method, the modularity statistic, summarizing the degree to which the network can be subdivided into groups with high ingroup similarity and low outgroup similarity, averaged 0.1740 (SD 0.05) for the movie-viewing scans and 0.1905 (SD 0.05) for the rest-state scans. Randomized null networks averaged 0.1206 (SD 0.0261). In comparative terms, subject data/surrogate data, the subjects are approximately 1.5 times greater in this statistic than the surrogates. Adjusting for multiple comparisons, 96% of subjects in the movie viewing condition displayed significant modular organization along the time dimension, and 94% in the rest condition. 99% displayed significant modularity in at least one of the two conditions. Pivoted toward time, these module-analogs are called themes. Analysis identified 7.5 and 7.2 themes on average, respectively, in each subject in the two conditions, movie and rest, compared to approximately 12 themes in

randomized surrogates. These fundamental observations indicate that the progression of themes is structured, analogous to the modular architectures discovered when graph theory is applied to spatial networks.

Subsequent questions, then, probe the origin and structure of the apparent temporal dynamics, collected in **Table 2**. Repetition was measured by simply comparing all the subsequences of 4–11 images in each thematic profile, and counting exact matches throughout the profile. This analysis found highly significant repetition for at least some subjects at all of these durations. Overall, during the rest condition 81% of subjects displayed at least one sequence with repetitions greater than the surrogates. On average, each subject displayed significant repetition for 6 sequence lengths. The sequence length exhibiting repetition in the largest subset of subjects was 6 s. During the movie viewing condition, 87% of subjects showed repetition for at least one sequence length; on average, each manifested repetition for 6 different sequence lengths. Sequence lengths of 6 or 7 s were the most frequent repeaters.

Rhythmicity was measured by examining the intervals between repetitions and counting the number of recurrences of each interval. As with the repetitions, this was separately examined for sequences of lengths 4 through 11 s. These tabulations were compared with similar tabulations for 100 surrogate data sets, and significant deviations recorded (as always, correcting for multiple comparisons). Overall 88 and 91%, respectively, of movie and rest condition subjects showed significant rhythmicity at least one sequence length. Sequences of 6 or 7 s were most often rhythmic, displayed by around 86% of subjects in both conditions.

Harmonicity was measured by summing the original and six downsampled spectra for each subject. By shrinking the spectra by an integer factor, the downsampling preserves peaks in relationships of integer multiples – i.e., harmonics. (Due to the downsampling, the peaks of the summed downsampled spectra cannot be assigned to specific frequencies). This test yielded highly significant harmonics for all subjects in the two conditions. Sequence lengths of 4–7 s were harmonically organized in nearly all subjects, with all subjects viewing movies displaying harmony for 5 s sequences, and all subjects in the rest condition displaying harmonics for sequences of 5, 6, and 7 s. In both conditions over 40 specific peaks exceeded baseline surrogate measures in both conditions.

## DISCUSSION

### What Now? What's Next?

Very generally, animal brains face a dual computational demand: they must sense (and interpret) what is immediately present; and they must predict what will happen next (over a future from milliseconds to hours to years) (Friston and Stephan, 2007; Clark, 2013, 2016; Hohwy, 2013). In computational terms, all animals need a capacity to continuously maintain representations of past and future environmental (and bodily) conditions, while at the same time continuously refreshing, updating, and modifying these representations as new information arrives.

**TABLE 1 |** Global modularity measures for brain images collected during two experimental conditions.

Global modularity measures, 180 Ss × 900 images	Movie	Rest
N	180	180
Mean modularity ( $q$ ) ± SD	0.1740 (0.0546)	0.1905 (0.0542)
(Mean surrogate data)	0.1143 (0.0297)	0.1269 (0.0261)
% significant subjects	96	94
Modal node (theme) count	5	4
Surrogate modal node (theme) count	9	9

*The modularity statistic ( $q$ ) was computed using the Louvain group method (Blondel et al., 2008). Most subjects displayed community-grouping structures over time among recurrent themes, as assessed through contrasts with modularity measures derived from null networks (Rubinov and Sporns, 2010). Actual data tended toward network reconstructions with around five distinct themes. The surrogate models oscillated among more distinct themes.*

**TABLE 2 |** Repetition, rhythm, and harmony.

Repetition, Rhythm, and Harmony								
Passive movie viewing								
	Sequence length (s)							
	4	5	6	7	8	9	10	11
<b>Repetition</b>								
% of Ss with significant values	67	75	78	81	79	82	78	77
Repeats of themes/length of image series	1.70	1.09	0.73	0.48	0.33	0.23	0.17	0.12
Std	0.78	0.54	0.41	0.31	0.24	0.19	0.15	0.12
<b>Rhythm</b>								
% of Ss with significant values	76	85	85	86	82	72	63	59
Rhythmicity: std of significant interval counts	1.60	1.48	1.26	1.06	0.86	0.68	0.55	0.47
Mean difference from random surrogates	0.78	0.80	0.71	0.63	0.57	0.47	0.39	0.36
<b>Harmony</b>								
% of Ss with significant values	99	100	99	98	97	90	87	84
Median # of significant freqs per S	48.0	65.0	62.5	57.5	50.0	44.0	37.5	34.5
Harmonicity: mean ratio data vs. surrogates, power at downsampled frequencies	3.40	12.95	15.34	8.21	6.60	7.17	8.35	8.48
± SEM	0.08	0.88	0.90	0.30	0.20	0.21	0.28	0.33
<b>Rest</b>								
<b>Repetition</b>								
% of Ss with significant values	61	68	72	73	69	67	63	61
Repeats of themes/length of image series	1.82	1.18	0.78	0.50	0.34	0.23	0.16	0.11
Std	0.84	0.70	0.49	0.35	0.29	0.23	0.18	0.14
<b>Rhythm</b>								
% of Ss with significant values	78	84	87	83	72	67	56	53
Rhythmicity: std of significant interval counts	1.71	1.55	1.34	1.05	0.80	0.64	0.48	0.39
Mean difference from random surrogates	0.71	0.75	0.67	0.56	0.49	0.43	0.32	0.28
<b>Harmony</b>								
% of Ss with significant values	99	100	100	100	97	96	93	92
Median # of significant freqs per S	42.00	53.00	56.00	54.00	47.00	43.00	35.00	30.00
Harmonicity: mean ratio data vs. surrogates, power at downsampled frequencies	2.09	10.84	10.91	12.90	13.54	8.60	6.42	6.71
± SEM	0.05	1.23	0.59	0.57	0.79	0.41	0.22	0.26

*The temporal connectome as it arises in the brain is characterized by high degrees of repetition, the presence of regular intervals between repetitions (i.e., rhythm), and the integer multiple relations among the frequencies of rhythms, as determining by the downsampled sum of spectra described in the text and **Figure 2**.*

This sketch of cognition is subject to three general constraints: First, to be useful, information that is generated at one source must be available to modify information elsewhere. Global availability leads to the second constraint: For information to be effectively integrated across the brain, it must be transmitted with minimum confusion. In effect, channels must converge and diverge without crosstalk. Third, all this temporal mixing and matching must work quickly, to keep up with a dynamic world.

## Oscillation, Information Broadcasting, and Maintenance

How might repetition, rhythm, or harmony enable these computational ends? Oscillations are everywhere at frequencies from less than 1 to 150 Hz (Biswal et al., 1995; Linkenkaer-Hansen et al., 2001; Buzsáki and Draguhn, 2004; Buzsáki, 2006; Zuo et al., 2010; Kopell et al., 2014), and most likely originate in

neural activity (Zuo et al., 2010). Along with their observation we find a cornucopia of proposals for their function. Many of these posit interactions among oscillations at different frequencies, where frequency bands have distinct functions (Glassman, 1999; Onslow et al., 2011; Aru et al., 2015; Wiener and Kanai, 2016; Lundqvist et al., 2018; Zhang et al., 2018). For example, Friston et al. (2015) have proposed that theta and gamma oscillations signal from the periphery up while beta oscillations provide feedback from the top down in the visual system. Other researchers propose that oscillations perform a gating function, for synchronizing signals (Buzsáki and Draguhn, 2004; Fries, 2005, 2015; Maris et al., 2016). Canolty et al. (2010) propose that oscillation is the electrophysiological signature of Hebbian cell assemblies at work (Canolty et al., 2010). In general, however, oscillation implies a capacity for information maintenance (Buzsáki and Draguhn, 2004). This may be important for maintaining attention and working

memory, among other functions (Buzsáki and Draguhn, 2004; Hipp et al., 2011; Aru et al., 2015; Fries, 2015; Gregoriou et al., 2015; Gupta and Chen, 2016). Instantaneous coupling of signals (“microstates”) are another variation on the communicative role of oscillations (Schack, 2004; Dimitriadis et al., 2013). These are generally derived from EEG data, with the exception of Ville et al. (2010) and Hipp et al. (2011), who found scale-free dynamics including the frequencies observed through fMRI, and Zuo et al. (2010) who explored frequency bands below 0.1 Hz, to identify regions of the brain where oscillations within particular frequency bands exhibited greater amplitude.

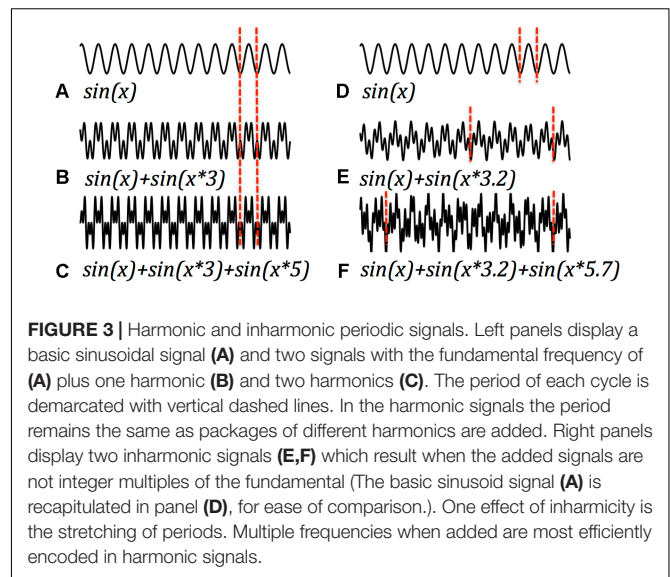
## Rhythm

The rhythms described in this study are particularly apt candidates for temporal holding patterns, in that the methods here define the elements in rhythmic repetition as sequences of thematic moments. What is repeating is itself a temporally extended pattern (4–11 s), drawn from an alphabet of around seven themes. The many rhythms found in both the rest and movie conditions invite a more detailed study of these patterns, in addition to research into their frequency of oscillation (Zuo et al., 2010).

## Harmony

Among the many discussions of oscillations in the brain, discussions of harmonic relationships among frequencies are rare. [Atasoy et al. (2016) probes spatial frequencies in harmonic relations, but not temporal harmonics]. Yet harmonic partials are rampant in the data in the present study – at least 30 partials were discovered in the frequencies of sequence occurrence at all sequence lengths. What could be the functional significance of this widespread observation?

A speculative argument could begin with functional distinctions between the stages of signal processing in any system, the brain included. A full mechanistic account of such a system must explain how signals are generated, how they are transmitted, and how they are received/interpreted. Abundant research explores how neuronal oscillations are generated; a fairly large literature considers how signals propagate over space and time; but there is little consideration of how a received signal is processed. Fourier analyses are computationally intensive, and require many signal samples to be precise – it seems unlikely that the brain computes in this way. In a periodic signal, however, the minimum interpretable packet is one cycle. Accordingly, in principle the fastest processing is most feasible when cycle time, the interval between repetitions of the periodic signal, is shortest. In general, harmonic signals offer a useful combination of multiple superimposed frequencies and short cycle times. This follows from the definition of harmonics, which are signals whose frequencies are in integer ratios, but might also be illustrated with an example. The left panel of **Figure 3** illustrates different harmonic signals, the pure signal at one frequency,  $\sin(x)$ , and several composite signals,  $\sin(x) + \sin(x \times C)$  where  $x$  is a monotonic vector (of time points) and  $C$  is an integer. Accordingly, the composite signals on the left side of the figure are harmonic signals of various lengths, with the same fundamental frequency. For each, one cycle is bracketed. The



right panel presents examples of inharmonic signals and their cycle times. It is apparent that the harmonic signals have the shortest cycles, as is indeed implied by the integer relationships of their frequencies. Thus, if a signal is composed of multiple frequencies, the smallest package (quickest, easiest, most efficient) that delivers all the frequency information employs frequencies in harmonic relationships. The example introduces just two harmonic partials, but this hypothetical computational process can accommodate more complex harmonic signals as well. The presence and absence of harmonics afford the system a binary code, albeit one of modest capacity. Such a system gets off the ground without Fourier analysis, and packages its message in a minimum interval (Glassman, 2000).

Harmonic signals, in short, afford rapid “unpacking” by their receivers, a property that might make them adaptive for natural selection.

Could the brain implement a computational process that can extract harmonics from a mixed signal? Encouragement comes from the real biological analogy of hearing, which sorts fundamentals and overtones through a combination of cochlear shape and specialized sensory neurons. It does so quickly, but not in a single cycle. Glassman considered harmonic information at the timescale captured by EEG, proposing that short term memory with its famous capacity limitation might be embodied in a harmonic resonance of a complex marker or cue for each memorized item (Glassman, 1999). In each octave of any resonator the number of subharmonics is limited [subharmonics are frequencies of higher harmonics dropped from their octaves (halved repeatedly) into a single octave]. He hypothesized that the harmonics could keep separate the items while binding them in a stable system of resonances, amenable to extraction at the time of recall. In the current paper the time resolution of fMRI limits the analysis to very low frequencies. A single cycle affords ample neural computation time. Indeed, it’s most likely that any real neural process that exploits harmonics is operating at higher frequencies, leaving only subharmonics in the cycles discernible to fMRI (Buzsáki, 2006).

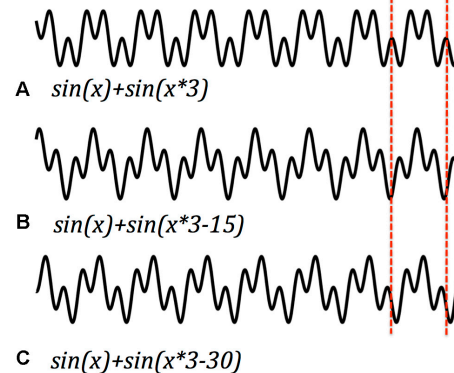
How might signal analysis work at the neural level? Timing in many animals may be supported by harmonic properties of the interaction of oscillations from particular brain regions (Gupta and Chen, 2016). The “striatal beat frequency” model of timing posits multiple oscillators at frequencies in harmonic relationships (Matell and Meck, 2004; Buhusi and Meck, 2005; Merchant et al., 2011; Kononowicz and van Wassenhove, 2016). For example, suppose there are three distinct oscillators with periods of 2, 3, and 5 s. At various moments their oscillations will coincide. The 2 s and 3 s oscillations reinforce at every 6 s, while 2 and 5 converge every 10 s. All three reinforce every 30 s. To time intervals of 6, 10, or 30 s a system needs simply to detect these convergences. Timing in this example is a punctate response. Temporal perception emerges when we imagine continuous relationships among harmonics. If the “beats” are rising and falling gradients, as might emerge from a time-varying harmonic signal, their mix could provide continuous temporal information.

Frequency is only part of the information that could be encoded by harmonics. Harmonic signals also differentiate by the relative phase of their component frequencies, which is the offset of the zero-crossings of their cycles. Since amplitude is additive from moment to moment in any signal, in-phase and out-of-phase signals have very different overall shapes (despite similar spectra). This offers another feature with a capacity to carry temporal information (Gupta and Chen, 2016; Hakim and Vogel, 2018). The phase of a periodic signal is set at the origin of the signal, or (more likely) reset by an event that interrupts continuing oscillation. If different events reset different signals at different frequencies, and if those frequencies are harmonically related, then the ongoing signal encodes the interval between the initiating events. **Figure 4** illustrates an example. Each is the sum of the same fundamental and first harmonic. However, in each panel the origin of the harmonic sinusoid has been time-shifted by a different interval. As in the earlier examples, the cycle time remains the same, and the overall structure of peaks within each period. But the relative magnitude of the peaks shifts with the mismatched phases. One cycle of a harmonic signal, it seems, can signal phasic differences, another available and quick vehicle for usable information. I’ve suggested that the event of initiation of a resonant wave could be what determines its phase. If that’s so, then the package carries a rough representation of sequence. Once again, harmonics can be added, each with a different phase (see also Onslow et al., 2011; Maris et al., 2016; Zhang et al., 2018).

To summarize this discussion, the computational constraints of temporal context, separability of signals, and speed might be efficiently met in a resonating system that is wired for rhythm and harmony.

## Limitations and Future Directions

Recent literature has emphasized the importance of confirming the reliability of the many measures typical of fMRI research (Zuo and Xing, 2014; Poldrack et al., 2017; Zuo et al., 2019a,b). This has not been undertaken in this paper. Therefore, the actual statistical power of the analysis awaits further study. Meanwhile, however, several features of the data and its analysis indicate that the main results of this study do rest on reliable observations. First, the effects described are large, as is the size of the data



**FIGURE 4 |** Harmonic signals with the same frequency but different phase offset differ in their wave form for each cycle (e.g., **A–C**). Phase offsets and wave forms. Harmonic signals with the same frequency but different phase offset differ in their wave form for each cycle. Dashed lines demarcate one cycle. Phasic information could be recovered from relative amplitude peaks within a single cycle.

set, with 180 subjects examined. For example, the modularity measure used here, when compared to the 100 null networks for each subject and task, has an effect size greater than 1 by Hedges’ *g* (Hedges, 1981). These observations are consistent across the two tasks examined, similar to the report of O’Connor et al. (2017) comparing four experimental conditions. More important, the Human Connectome Project has had replicability as a major goal (Marcus et al., 2013), and its reliability is supported in Termenon et al. (2016). Scanning at 7 Tesla, using standardized preprocessing pipelines, and the application of cortical surface coordinates to localize brain activity all increase confidence in the consistency of the scans (but see Cremers et al., 2017; Poldrack et al., 2017; Smith and Nichols, 2018). Studies of the reliability of HCP subjects have been less frequent, although the HCP records more than 500 phenotypic features for each subject, again indicating the care the researchers are bringing to their task. Nonetheless the analysis here should be regarded as provisional, pending a future examination of this issue.

## CONCLUSION

### Temporal Processing in the Brain

This paper offers an initial attempt to outline novel aspects of temporality in the human brain, at the low temporal resolution afforded by fMRI. The first step was to create a matrix of relationships among moments, an association matrix determined by similarity of patterns. This immediately revealed that the brain patterns are not simply counting off the seconds, where each image is most similar to its nearest temporal neighbor. Instead, the brains of our subjects moved among a small number of states. These clusters can be established by many different methods (Lloyd, 2002, 2004, 2012). Here we employed graph theoretic measures of modularity to outline the stable clusters we’ve called themes. The “network” discovered in this way is a sequential alternation of themes, a thematic profile. This served as a first



approximation of temporal processing analyzed independently from the general framework of Fourier and wavelet analysis.

The observation of modular temporal structure motivated a search for intrinsic temporal anatomy, regularities within the thematic structures projected by graph theory. There are many avenues of exploration possible. Here, we looked at the regular manifestations of repetition, rhythm, and especially harmony, a distinctively useful configuration of oscillations that seems to exploit discernible features that only harmonic signals possess. The analysis here suggested that there are harmonic relationships underlying the oscillations of the thematic profiles; their computational uses suggest that harmonic signaling might be a useful adaptation. Harmonic signals provide short repeating temporal motifs from which separate frequencies can be extracted. The motifs resonate at their fundamental frequency and within each period harmonics oscillate in regular patterns. These patterns are further modified by the relative phases of the component frequencies.

In this package of harmonics we may discern a basic structure of temporality. It is a resonating holding pattern which carries information originating before the interval of each repetition. The broad similarities in the temporal properties in two very different experimental conditions, movie viewing and rest, imply that the observed rhythms and harmonies are fundamental to informational processes in the brain. From this vantage point we can't determine if the global brain patterns we observe originate from separate sources with distinct frequency profiles, or from a single harmonic resonator, but we can observe the relationships among the frequencies that can be extracted. Moreover, in this analysis we haven't examined relative phase for the extracted harmonics, but here too this information could come from a single global source or multiple sources. The answers to these questions would pivot back toward the spatial, for example, measuring the amplitudes of oscillations in various brain regions (Zuo et al., 2010). But strictly within the temporal realm, brain dynamics are orderly and perseverating. The observations here suggest a human capacity to spread out from the immediate present tense of sensation, toward an overall temporal landscape. The brains examined here show signs of a present inflected by a past that resonates and possibly holds information about sequence and interval, and thus can also encode expectations of the immediate future. There is, of course, much that is speculative here; these proposals are offered as a starting point for further study.

## The Music of Thought

Certain concepts seem apt for redeployment in the study of temporality: theme and modularity, repetition, rhythm, and especially harmony. These of course are familiar through music. In other works, I've suggested that the analogies between brain dynamics and musical form should be taken literally, or at any rate as literally as language in the hypothetical "language of thought" (Lloyd, 2011). Musical concepts have one common feature that makes them especially useful here: time is essential to all of them, and so they are properties appropriate to the observations following the pivot from space to time (Lloyd, 2013). Such concepts *could* apply, but it is an empirical question whether

they do apply. Are there in fact rhythm, theme and refrain, and harmony in fMRI signals? In the experiment reviewed here, the answer is a (tentative) yes. Indeed, the fMRI analyses here point to the pervasive presence of repetition, rhythm, and especially harmony. Among human artifacts, only music approaches this density and structure of repetition (Huron, 2014). In sharp contrast, these properties are at best weakly present in language, which has often been proposed as the model for cognition and ultimately brain function (Fodor, 1979; Lloyd, 2011).

One attractive topic for further exploration is the scaling behavior of the rhythms observed here (He et al., 2010; Kello et al., 2010; Ville et al., 2010; Hardstone et al., 2012). Since frequency here has a novel definition, scaling behaviors would require a distinct test, a topic for future study [especially since music exhibits well-known scaling laws (Manaris et al., 2005; Lloyd, 2011; González-Espinoza et al., 2017)].

The literal connection to music may seem implausible. After all, music is a cultural artifact and art form seemingly incidental to the serious business of survival and reproduction – "auditory cheesecake," in Pinker's (1997) memorable phrase. But several considerations suggest that sidelining music is a mistake. In many ways music is fully parallel to language in its intimacy with the human condition: Music is universal to human cultures (Patel, 2010); World musical systems almost universally share certain features, including the use of scales and limited rhythmic patterns (Huron, 2014); Music is old. [The oldest instrument, a carefully crafted flute, was made more than 40,000 years ago (Higham et al., 2012)]. Music is potentially advantageous for social cohesion or sexual selection (Levitin, 2007; Patel, 2010). Music, unlike language, lacks the power to denote specific objects and scenes (Hanslick, 1891; Kivy, 1990; Lloyd, 2011). But what it might represent, by analogy, is the dynamic operation of the brain itself. Musicians may be improvising a model of mind in sound. In that case, musical concepts are not externals brought to bear on brain dynamics, but rather the natural, intuitive expression of that dynamic, a "music of thought."

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

## FUNDING

Data were provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research and the McDonnell Center for Systems Neuroscience at Washington University.

## ACKNOWLEDGMENTS

I would like to thank the referees for this paper, and the editor of this special issue of *Frontiers*, Daya Gupta.

## REFERENCES

- Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* N. Y. 24, 663–676. doi: 10.1093/cercor/bhs352
- Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., et al. (2015). Untangling cross-frequency coupling in neuroscience. *Curr. Opin. Neurobiol.* 31, 51–61. doi: 10.1016/j.conb.2014.08.002
- Atasoy, S., Donnelly, I., and Pearson, J. (2016). Human brain networks function in connectome-specific harmonic waves. *Nat. Commun.* 7:10340. doi: 10.1038/ncomms10340
- Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541. doi: 10.1002/mrm.1910340409
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* 2008:10008.
- 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
- Bullmore, E., and Sporns, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575
- Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press.
- Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745
- Calhoun, V. D., Miller, R., Pearson, G., and Adali, T. (2014). The chronnectome: time-varying connectivity networks as the next frontier in fMRI data discovery. *Neuron* 84, 262–274. doi: 10.1016/j.neuron.2014.10.015
- Canolty, R. T., Ganguly, K., Kennerley, S. W., Cadieu, C. F., Koepsell, K., Wallis, J. D., et al. (2010). Oscillatory phase coupling coordinates anatomically dispersed functional cell assemblies. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17356–17361. doi: 10.1073/pnas.1008306107
- Chakravartty, A. (2017). *Scientific Ontology: Integrating Naturalized Metaphysics and Voluntarist Epistemology*. Oxford: Oxford University Press.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behav. Brain Sci.* 36, 181–204. doi: 10.1017/S0140525X12000477
- Clark, A. (2016). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford: Oxford University Press.
- Creemers, H. R., Wager, T. D., and Yarkoni, T. (2017). The relation between statistical power and inference in fMRI. *PLoS One* 12:e0184923. doi: 10.1371/journal.pone.0184923
- Cuadra, P. D., Master, A. S., and Sapp, C. S. (2001). “Efficient pitch detection techniques for interactive music,” in *Proceedings of the International Computer Music Conference*, (San Francisco, CA).
- Dimitriadis, S. I., Laskaris, N. A., and Tzelepi, A. (2013). On the quantization of time-varying phase synchrony patterns into distinct functional connectivity microstates (FCμstates) in a multi-trial visual ERP paradigm. *Brain Topogr.* 26, 397–409. doi: 10.1007/s10548-013-0276-z
- Epskamp, S., and Fried, E. I. (2018). A tutorial on regularized partial correlation networks. *Psychol. Methods* 23, 617–634. doi: 10.1037/met0000167
- Fallani, F. D. V., Richiardi, J., Chavez, M., and Achard, S. (2014). Graph analysis of functional brain networks: practical issues in translational neuroscience. *Philos. Trans. R. Soc. B Biol. Sci.* 369:20130521. doi: 10.1098/rstb.2013.0521
- Fodor, J. A. (1979). *The Language of Thought*. Cambridge, MA: Harvard University Press.
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474–480. doi: 10.1016/j.tics.2005.08.011
- Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034
- Friston, K. J., Bastos, A. M., Pinotsis, D., and Litvak, V. (2015). LFP and oscillations—what do they tell us? *Curr. Opin. Neurobiol.* 31, 1–6. doi: 10.1016/j.conb.2014.05.004
- Friston, K. J., and Stephan, K. E. (2007). Free-energy and the brain. *Synthese* 159, 417–458.
- Girvan, M., and Newman, M. E. J. (2002). Community structure in social and biological networks. *Proc. Natl. Acad. Sci. U.S.A.* 99, 7821–7826.
- Glasser, M. F., Coalson, T. S., Robinson, E. C., Hacker, C. D., Harwell, J., Yacoub, E., et al. (2016). A multi-modal parcellation of human cerebral cortex. *Nature* 536, 171–178. doi: 10.1038/nature18933
- Glasser, M. F., Sotiropoulos, S. N., Wilson, J. A., Coalson, T. S., Fischl, B., Andersson, J. L., et al. (2013). The minimal preprocessing pipelines for the human connectome project. *NeuroImage* 80, 105–124. doi: 10.1016/j.neuroimage.2013.04.127
- Glassman, R. B. (1999). Hypothesized neural dynamics of working memory: several chunks might be marked simultaneously by harmonic frequencies within an octave band of brain waves. *Brain Res. Bull.* 50, 77–93. doi: 10.1016/s0361-9230(99)00090-8
- Glassman, R. B. (2000). A “Theory of relativity” for cognitive elasticity of time and modality dimensions supporting constant working memory capacity: involvement of harmonics among ultradian clocks? *Prog. Neuropsychopharmacol. Biol. Psychiatry* 24, 163–182. doi: 10.1016/s0278-5846(99)00096-2
- Glerean, E., Salmi, J., Lahnakoski, J. M., Jääskeläinen, I. P., and Sams, M. (2012). Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity. *Brain Connect.* 2, 91–101. doi: 10.1089/brain.2011.0068
- González-Espinoza, A., Larralde, H., Martínez-Mekler, G., and Müller, M. (2017). Multiple scaling behaviour and nonlinear traits in music scores. *R. Soc. Open Sci.* 4:171282. doi: 10.1098/rsos.171282
- Gravel, N., Harvey, B. M., Renken, R. J., Dumoulin, S. O., and Cornelissen, F. W. (2018). Phase-synchronization-based parcellation of resting state fMRI signals reveals topographically organized clusters in early visual cortex. *Neuroimage* 170, 424–433. doi: 10.1016/j.neuroimage.2017.08.063
- Gregoriou, G. G., Paneri, S., and Sapountzis, P. (2015). Oscillatory synchrony as a mechanism of attentional processing. *Brain Res.* 1626, 165–182. doi: 10.1016/j.brainres.2015.02.004
- Gupta, D. S., and Chen, L. (2016). Brain oscillations in perception, timing and action. *Curr. Opin. Behav. Sci.* 8, 161–166. doi: 10.1016/j.cobeha.2016.02.021
- Hagmann, P., Kurant, M., Gigandet, X., Thiran, P., Wedeen, V. J., Meuli, R., et al. (2007). Mapping human whole-brain structural networks with diffusion MRI. *PLoS One* 2:e597. doi: 10.1371/journal.pone.0000597
- Hakim, N., and Vogel, E. K. (2018). Phase-coding memories in mind. *PLoS Biol.* 16:e3000012. doi: 10.1371/journal.pbio.3000012
- Hanslick, E. (1891). *The Beautiful in Music; a Contribution to the Revisal of Musical Aesthetics*. London: Novello and Co.
- Hardstone, R., Poil, S.-S., Schiavone, G., Jansen, R., Nikulin, V. V., Mansvelder, H. D., et al. (2012). Detrended fluctuation analysis: a scale-free view on neuronal oscillations. *Front. Physiol.* 3:450. doi: 10.3389/fphys.2012.00450
- He, B. J., Zempel, J. M., Snyder, A. Z., and Raichle, M. E. (2010). The temporal structures and functional significance of scale-free brain activity. *Neuron* 66, 353–369. doi: 10.1016/j.neuron.2010.04.020
- Hedges, L. V. (1981). Distribution theory for Glass's Estimator of effect size and related estimators. *J. Educ. Stat.* 6, 107–128. doi: 10.3102/10769986006002107
- Higham, T., Basell, L., Jacobi, R., Wood, R., Ramsey, C. B., and Conard, N. J. (2012). Testing models for the beginnings of the Aurignacian and the advent of figurative art and music: the radiocarbon chronology of Geißenklösterle. *J. Hum. Evol.* 62, 664–676. doi: 10.1016/j.jhevol.2012.03.003
- 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
- Hodge, M. R., Horton, W., Brown, T., Herrick, R., Olsen, T., Hileman, M. E., et al. (2016). ConnectomeDB—Sharing human brain connectivity data. *NeuroImage* 124(Pt B), 1102–1107. doi: 10.1016/j.neuroimage.2015.04.046
- Hohwy, J. (2013). *The Predictive Mind*. Oxford: Oxford University Press.
- Huron, D. (2014). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge: MIT Press.
- 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

- Kaiser, M. (2011). A tutorial in connectome analysis: topological and spatial features of brain networks. *NeuroImage* 57, 892–907. doi: 10.1016/j.neuroimage.2011.05.025
- Kello, C. T., Brown, G. D. A., Ferrer-i-Cancho, R., Holden, J. G., Linkenkaer-Hansen, K., Rhodes, T., et al. (2010). Scaling laws in cognitive sciences. *Trends Cogn. Sci.* 14, 223–232. doi: 10.1016/j.tics.2010.02.005
- Kivy, P. (1990). *Music Alone: Philosophical Reflections on the Purely Musical Experience*. Ithaca: Cornell University Press.
- Kononowicz, T. W., and van Wassenhove, V. (2016). In search of oscillatory traces of the internal clock. *Front. Psychol.* 7:224. doi: 10.3389/fpsyg.2016.00224
- Kopell, N., Gritton, H. J., Whittington, M. A., and Kramer, M. A. (2014). Beyond the connectome: the dynamo. *Neuron* 83, 1319–1328. doi: 10.1016/j.neuron.2014.08.016
- Lacasa, L., Luque, B., Ballesteros, F., Luque, J., and Nuño, J. C. (2008). From time series to complex networks: the visibility graph. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4972–4975. doi: 10.1073/pnas.0709247105
- 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
- Laird, A. R., Rogers, B. P., Carew, J. D., Arfanakis, K., Moritz, C. H., and Meyerand, M. E. (2002). Characterizing instantaneous phase relationships in whole-brain fMRI activation data. *Hum. Brain Mapp.* 16, 71–80. doi: 10.1002/hbm.10027
- Levitin, D. J. (2007). *This is Your Brain on Music: the Science of a Human Obsession*. New York: Plume.
- Linkenkaer-Hansen, K., Nikouline, V. V., Palva, J. M., and Ilmoniemi, R. J. (2001). Long-range temporal correlations and scaling behavior in human brain oscillations. *J. Neurosci.* 21, 1370–1377. doi: 10.1523/jneurosci.21-04-01370.2001
- Lloyd, D. (2002). Functional MRI and the study of human consciousness. *J. Cogn. Neurosci.* 14, 818–831. doi: 10.1162/089892902760191027
- Lloyd, D. (2004). *Radiant Cool: a Novel Theory of Consciousness*. Cambridge, MA: MIT Press.
- Lloyd, D. (2011). Mind as music. *Front. Psychol.* 2:63. doi: 10.3389/fpsyg.2011.00063
- Lloyd, D. (2012). “Time after time,” in *Being in Time: Dynamical Models of Phenomenal Experience*, eds S. Edelman, T. Fekete, and N. Zach, (Amsterdam: John Benjamins), 88–81.
- Lloyd, D. (2013). The music of consciousness: can musical form harmonize phenomenology and the brain? *Constr. Found.* 8, 324–331.
- Lundqvist, M., Herman, P., Warden, M. R., Brincat, S. L., and Miller, E. K. (2018). Gamma and beta bursts during working memory readout suggest roles in its volitional control. *Nat. Commun.* 9, 1–12. doi: 10.1038/s41467-017-02791-8
- Manaris, B., Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W., et al. (2005). Zipf’s law, music classification, and aesthetics. *Comput. Music J.* 29, 55–69. doi: 10.1162/comj.2005.29.1.55
- Marcus, D., Harwell, J., Olsen, T., Hodge, M., Glasser, M., Prior, F., et al. (2011). Informatics and data mining tools and strategies for the human connectome project. *Front. Neuroinformatics* 5:4. doi: 10.3389/fninf.2011.00004
- Marcus, D. S., Harms, M. P., Snyder, A. Z., Jenkinson, M., Wilson, J. A., Glasser, M. F., et al. (2013). Human connectome project informatics: quality control, database services, and data visualization. *NeuroImage* 80, 202–219. doi: 10.1016/j.neuroimage.2013.05.077
- Maris, E., Fries, P., and van Ede, F. (2016). Diverse phase relations among neuronal rhythms and their potential function. *Trends Neurosci.* 39, 86–99. doi: 10.1016/j.tins.2015.12.004
- Marrelec, G., Horwitz, B., Kim, J., Péligrini-Issac, M., Benali, H., and Doyon, J. (2007). Using partial correlation to enhance structural equation modeling of functional MRI data. *Magn. Reson. Imaging* 25, 1181–1189. doi: 10.1016/j.mri.2007.02.012
- Maslov, S., and Sneppen, K. (2002). Specificity and stability in topology of protein networks. *Science* 296, 910–913. doi: 10.1126/science.1065103
- 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
- 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
- Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: simple building blocks of complex networks. *Science* 298, 824–827. doi: 10.1126/science.298.5594.824
- Newman, M. E. J. (2006). Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E* 74:036104.
- O’Connor, D., Potler, N. V., Kovacs, M., Xu, T., Ai, L., Pellman, J., et al. (2017). The healthy brain network serial scanning initiative: a resource for evaluating inter-individual differences and their reliabilities across scan conditions and sessions. *Gigascience* 6, 1–14. doi: 10.1093/gigascience/giw011
- Onslow, A. C. E., Bogacz, R., and Jones, M. W. (2011). Quantifying phase-amplitude coupling in neuronal network oscillations. *Prog. Biophys. Mol. Biol.* 105, 49–57. doi: 10.1016/j.pbiomolbio.2010.09.007
- Patel, A. D. (2010). *Music, Language, and the Brain*. Oxford: Oxford University Press.
- Pinker, S. (1997). *How the Mind Works*. New York, NY: Norton.
- Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., et al. (2017). Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat. Rev. Neurosci.* 18, 115–126. doi: 10.1038/nrn.2016.167
- Reichardt, J., and Bornholdt, S. (2006). Statistical mechanics of community detection. *Phys. Rev. E* 74:016110.
- Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *NeuroImage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003
- Salimi-Khorshidi, G., Douaud, G., Beckmann, C. F., Glasser, M. F., Griffanti, L., and Smith, S. M. (2014). Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *NeuroImage* 90, 449–468. doi: 10.1016/j.neuroimage.2013.11.046
- Schack, B. (2004). How to construct a microstate-based alphabet for evaluating information processing in time. *Int. J. Bifurc. Chaos* 14, 793–814. doi: 10.1142/s02181274040009478
- Sethares, W. A. (1998). *Tuning, Timbre, Spectrum, Scale*. London: Springer.
- Sethares, W. A. (2007). *Rhythm and Transforms*. London: Springer-Verlag.
- Shirazi, A. H., Jafari, G. R., Davoudi, J., Peinke, J., Tabar, M. R. R., and Sahimi, M. (2009). Mapping stochastic processes onto complex networks. *J. Stat. Mech. Theory Exp.* 2009:07046.
- Smith, S. M., Andersson, J., Auerbach, E. J., Beckmann, C. F., Bijsterbosch, J., Douaud, G., et al. (2013). Resting-state fMRI in the human connectome project. *NeuroImage* 80, 144–168. doi: 10.1016/j.neuroimage.2013.05.039
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., et al. (2011). Network modelling methods for fMRI. *NeuroImage* 54, 875–891. doi: 10.1016/j.neuroimage.2010.08.063
- Smith, S. M., and Nichols, T. E. (2018). Statistical challenges in “Big Data” human neuroimaging. *Neuron* 97, 263–268. doi: 10.1177/1740774509105380
- Sporns, O. (2011a). The human connectome: a complex network. *Ann. N. Y. Acad. Sci.* 1224, 109–125. doi: 10.1111/j.1749-6632.2010.05888.x
- Sporns, O. (2011b). The non-random brain: efficiency, economy, and complex dynamics. *Front. Comput. Neurosci.* 5:5. doi: 10.3389/fncom.2011.00005
- Sporns, O. (2012). From simple graphs to the connectome: networks in neuroimaging. *NeuroImage* 62, 881–886. doi: 10.1016/j.neuroimage.2011.08.085
- Sporns, O., Tononi, G., and Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Comput. Biol.* 1:e42. doi: 10.1371/journal.pcbi.0010042
- Stanley, M. L., Moussa, M. N., Paolini, B., Lyday, R. G., Burdette, J. H., and Laurienti, P. J. (2013). Defining nodes in complex brain networks. *Front. Comput. Neurosci.* 7:169. doi: 10.3389/fncom.2013.00169
- Termenon, M., Jaillard, A., Delon-Martin, C., and Achard, S. (2016). Reliability of graph analysis of resting state fMRI using test-retest dataset from the Human Connectome Project. *NeuroImage* 142, 172–187. doi: 10.1016/j.neuroimage.2016.05.062
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E. J., Yacoub, E., and Ugurbil, K. (2013). The WU-Minn human connectome project: an overview. *NeuroImage* 80, 62–79. doi: 10.1016/j.neuroimage.2013.05.041

- Varela, F., Lachaux, J. P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239. doi: 10.1038/35067550
- Varoquaux, G., and Craddock, R. C. (2013). Learning and comparing functional connectomes across subjects. *NeuroImage* 80, 405–415. doi: 10.1016/j.neuroimage.2013.04.007
- Ville, D. V. D., Britz, J., and Michel, C. M. (2010). EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18179–18184. doi: 10.1073/pnas.1007841107
- Voss, H. U., Timmer, J., and Kurths, J. (2004). Nonlinear dynamical system identification from uncertain and indirect measurements. *Int. J. Bifurc. Chaos* 14, 1905–1933. doi: 10.1142/s0218127404010345
- 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
- Zhang, H., Watrous, A. J., Patel, A., and Jacobs, J. (2018). Theta and alpha oscillations are traveling waves in the human neocortex. *Neuron* 98, 1269–1281. doi: 10.1016/j.neuron.2018.05.019
- Zuo, X.-N., Biswal, B. B., and Poldrack, R. A. (2019a). Editorial: reliability and reproducibility in functional connectomics. *Front. Neurosci.* 13:117. doi: 10.3389/fnins.2019.00117
- Zuo, X.-N., Xu, T., and Milham, M. P. (2019b). Harnessing reliability for neuroscience research. *Nat. Hum. Behav.* 3, 768–771. doi: 10.1038/s41562-019-0655-x
- Zuo, X.-N., Di Martino, A., Kelly, C., Shehzad, Z. E., Gee, D. G., Klein, D. F., et al. (2010). The oscillating brain: complex and reliable. *NeuroImage* 49, 1432–1445. doi: 10.1016/j.neuroimage.2009.09.037
- Zuo, X.-N., and Xing, X.-X. (2014). Test-retest reliabilities of resting-state FMRI measurements in human brain functional connectomics: a systems neuroscience perspective. *Neurosci. Biobehav. Rev.* 45, 100–118. doi: 10.1016/j.neubiorev.2014.05.009

**Conflict of Interest:** The author declares 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 © 2020 Lloyd. 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) and the copyright owner(s) 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.



# Advantages of publishing in Frontiers



## OPEN ACCESS

Articles are free to read  
for greatest visibility  
and readership



## FAST PUBLICATION

Around 90 days  
from submission  
to decision



## HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,  
and constructive  
peer-review



## TRANSPARENT PEER-REVIEW

Editors and reviewers  
acknowledged by name  
on published articles

## Frontiers

Avenue du Tribunal-Fédéral 34  
1005 Lausanne | Switzerland

**Visit us:** [www.frontiersin.org](http://www.frontiersin.org)

**Contact us:** [info@frontiersin.org](mailto:info@frontiersin.org) | +41 21 510 17 00



## REPRODUCIBILITY OF RESEARCH

Support open data  
and methods to enhance  
research reproducibility



## DIGITAL PUBLISHING

Articles designed  
for optimal readership  
across devices



## FOLLOW US

@frontiersin



## IMPACT METRICS

Advanced article metrics  
track visibility across  
digital media



## EXTENSIVE PROMOTION

Marketing  
and promotion  
of impactful research



## LOOP RESEARCH NETWORK

Our network  
increases your  
article's readership