Advances in Experimental Medicine and Biology 1455

Hugo Merchant
Victor de Lafuente *Editors*

Neurobiology of Interval Timing

Second Edition

Springer

Advances in Experimental Medicine and Biology

Volume 1455

Advances in Experimental Medicine and Biology provides a platform for scientific contributions in the main disciplines of the biomedicine and the life sciences. This series publishes thematic volumes on contemporary research in the areas of microbiology, immunology, neurosciences, biochemistry, biomedical engineering, genetics, physiology, and cancer research. Covering emerging topics and techniques in basic and clinical science, it brings together clinicians and researchers from various fields.

Advances in Experimental Medicine and Biology has been publishing exceptional works in the field for over 40 years, and is indexed in SCOPUS, Medline (PubMed), EMBASE, Reaxys, EMBiology, the Chemical Abstracts Service (CAS), and Pathway Studio.

2022 CiteScore: 6.2

Hugo Merchant • Victor de Lafuente
Editors

Neurobiology of Interval Timing

Second Edition

Springer

Editors
Hugo Merchant
Instituto de Neurobiología
UNAM, Campus Juriquilla
Querétaro, Mexico

Victor de Lafuente
Institute of Neurobiology
National Autonomous
University of Mexico
Querétaro, Mexico

ISSN 0065-2598 ISSN 2214-8019 (electronic)
Advances in Experimental Medicine and Biology
ISBN 978-3-031-60182-8 ISBN 978-3-031-60183-5 (eBook)
https://doi.org/10.1007/978-3-031-60183-5

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2014, 2024
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

If disposing of this product, please recycle the paper.

To our kids: Carla, Matías, Fernanda
and Nina. Mateo and Emilia

Preface

The ability to process temporal information in the scale of hundreds of milliseconds to seconds is critical for a wide range of behaviors, from collision avoidance and target interception—present since the invertebrates—to highly complex behaviors such as language and music. In the second edition of the book *Neurobiology of Interval Timing*, we compile the newest and most exciting research in the brain sciences of timing. We gave special emphasis to the neural underpinnings of temporal processing in behaving human and non-human primates, as well as in rodents. Thus, the new edition of *Neurobiology of Interval Timing* integrates the thrilling and revealing developments on the psychophysics of time and timing neurophysiology, as well as fascinating modelling efforts to understand the clocks of the brain across a wide variety of behaviors, including perception and production of single intervals and rhythms in music and language. It is our sincere opinion that this constitutes an excellent book for graduate programs in neuroscience.

Querétaro, Mexico
Hugo Merchant
Victor de Lafuente

Contents

Part I Neurobiology of Interval Timing

Part II Neurophysiology of Timing

Part III Timing in Humans

Contributors

M. Florencia Assaneo Instituto de Neurobiología, Universidad Autónoma de México, Santiago de Querétaro, Mexico

Ana S. Báez-Cordero Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico

Fuat Balcı Department of Biological Sciences, University of Manitoba, Winnipeg, MB, Canada

Abraham Betancourt Instituto de Neurobiología, UNAM, Campus Juriquilla, Querétaro, Mexico

Zedong Bi Lingang Laboratory, Shanghai, China

Institute for Future, Qingdao University, Qingdao, China

School of Automation, Shandong Key Laboratory of Industrial Control Technology, Qingdao University, Qingdao, China

Fleur L. Bouwer Cognitive Psychology Unit, Institute of Psychology, Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands

Department of Psychology, Brain & Cognition, University of Amsterdam, Amsterdam, The Netherlands

Dean V. Buonomano Department of Neurobiology, University of California, Los Angeles, Los Angeles, CA, USA

Jennifer T. Coull Centre for Research in Psychology and Neuroscience (UMR 7077), Aix-Marseille Université & CNRS, Marseille, France

Victor de Lafuente Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico

Marisol Espinoza-Monroy Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico

Pamela García-Saldivar Instituto de Neurobiología, UNAM, Campus Juriquilla, Querétaro, Mexico

Perla González-Pereyra Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico

Simon Grondin École de Psychologie, Université Laval, Québec, Canada

Gábor P. Háden Institute of Cognitive Neuroscience and Psychology, Budapest, Hungary

Department of Telecommunications and Media Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary

Ana E. Hidalgo-Balbuena Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico

Henkjan Honing Music Cognition group (MCG), Institute for Logic, Language and Computation (ILLC), Amsterdam Brain and Cognition (ABC), University of Amsterdam, Amsterdam, The Netherlands

Masashi Kameda Department of Physiology, Hokkaido University School of Medicine, Sapporo, Japan

Inga Korolczuk Department of Pathophysiology, Medical University of Lublin, Lublin, Poland

Sonja A. Kotz Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

Ana M. Malagón Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico

Mario G. Martinez-Montalvo Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico

Germán Mendoza Instituto de Neurobiología, UNAM, Campus Juriquilla, Querétaro, Mexico

Karla Mercado Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico

Hugo Merchant Instituto de Neurobiología, UNAM, Campus Juriquilla, Querétaro, Mexico

Benjamin Morillon Aix Marseille Université, INSERM, INS, Institut de Neurosciences des Systèmes, Marseille, France

Ken-ichi Okada Department of Physiology, Hokkaido University School of Medicine, Sapporo, Japan

Joan Orpella Department of Neuroscience, Georgetown University Medical Center, Washington, DC, USA

Oswaldo Pérez Instituto de Neurobiología, UNAM, Campus Juriquilla, Querétaro, Mexico

Luis Prado Instituto de Neurobiología, UNAM, Campus Juriquilla, Querétaro, Mexico

Pavel E. Rueda-Orozco Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico

Mildred Salgado-Menez Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico

Michael Schwartze Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

Patrick Simen Department of Neuroscience, Oberlin College, Oberlin, OH, USA

Saray Soldado-Magraner Department of Neurobiology, University of California, Los Angeles, Los Angeles, CA, USA

Masaki Tanaka Department of Physiology, Hokkaido University School of Medicine, Sapporo, Japan

Sundeep Teki University College London, London, UK

Independent AI consultant, Delhi, India

Argiro Vatakis Multisensory and Temporal Processing Laboratory (MultiTimeLab), Department of Psychology, Panteion University of Social and Political Sciences, Athens, Greece

Martin Wiener George Mason University, Fairfax, VA, USA

About the Editors

Hugo Merchant, PhD is Professor of Neurobiology and Systems Neuroscience at the Institute of Neurobiology, Universidad Nacional Autonoma de Mexico, Juriquilla, Mexico. His research focuses on the neural basis of time. He has served as Associate Editor of the *Journal of Neuroscience* and the *Journal of Neurophysiology*, and on the Editorial Boards of *eLife*, *Frontiers in System Neuroscience*, and *Frontiers in Integrative Neuroscience*.

Victor de Lafuente, PhD is Professor of Neurobiology at the Institute of Neurobiology, Universidad Nacional Autonoma de Mexico, Queretaro, Mexico. His research interests are centered on the neuronal correlates of sensory perception, motor planning, and decision making. He is on the Board of Senior Editors of *Neuroscience* and is on the Editorial Boards of *Frontiers in Molecular Neurobiology* and *Frontiers in Systems Neuroscience.*

Part I

Neurobiology of Interval Timing

A Second Introduction to the Neurobiology of Interval Timing

Hugo Merchant and Victor de Lafuente

Abstract

Time is a critical variable that organisms must be able to measure in order to survive in a constantly changing environment. Initially, this paper describes the myriad of contexts where time is estimated or predicted and suggests that timing is not a single process and probably depends on a set of different neural mechanisms. Consistent with this hypothesis, the explosion of neurophysiological and imaging studies in the last 10 years suggests that different brain circuits and neural mechanisms are involved in the ability to tell and use time to control behavior across contexts. Then, we develop a conceptual framework that defines time as a family of different phenomena and propose a taxonomy with sensory, perceptual, motor, and sensorimotor timing as the pillars of temporal processing in the range of hundreds of milliseconds.

Keywords

Time perception · Sensory timing · Motor timing · Timing models

---

H. Merchant (✉)
Instituto de Neurobiología, UNAM, Campus Juriquilla, Querétaro, Mexico
e-mail: hugomerchant@unam.mx

V. de Lafuente
Institute of Neurobiology National Autonomous University of Mexico, Querétaro, Mexico

What Is Timing?

Timing is the tracking or planning of events that are constantly changing. “Time refers to the continued sequence of happenings that occur in apparently irreversible succession from the past, through present, and into the future” (Wikipedia). Organisms have developed different mechanisms to quantify the time between successive events, which could span ten orders of magnitude. The microseconds scale (10^{-4} s) is the scenario for binaural hearing and echolocation, engaging the auditory system to determine the spatial origin of sounds and objects (Joris & van der Heijden, 2019; Schnupp & Carr, 2009; Jeffress, 1948; O’Neill & Suga, 1979). At the other end, circadian timing (10^{5} s) organizes all the fundamental body functions within 24-h oscillations. We now know that the master clock for this circadian timing is the hypothalamic suprachiasmatic nucleus, which synchronizes the internal time with the external light–dark cycle, entraining the overall rhythmicity of a wide variety of peripheral clocks in the organism (Hastings et al., 2018), as well as the wake–sleep cycle in the brain (Drucker-Colín & Merchant-Nancy, 1996). Between very fast and the very slow bordering scales, we have the scale that spans hundreds of milliseconds (10^{-2}–10^{1} s), seconds, and minutes (10^{1}–10^{2} s). Timing within the seconds-and-minutes scales relies on conscious and cognitive control and is the context in which behavior is mapped onto the

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_1

external world. This timing range is related to foraging (Henderson et al., 2006), decision making (Brody et al., 2003), sequential motor performance (Bortoletto et al., 2011), as well as multiple-step arithmetic (Sohn & Carlson, 2003), and associative conditioning (Gallistel & Gibbon, 2000). In addition, temporal processing within the seconds-to-minutes scale is highly influenced by other processes, such as attention and memory, which interact with the mechanism of a presumed internal clock (Lewis & Miall, 2003). On the other hand, timing in the hundreds of milliseconds is crucial for behaviors such as object interception and collision avoidance, which we know are present since the invertebrates (Merchant & Georgopoulos, 2006; Merchant et al., 2001, 2003; 2009; 2004a, 2004b). Complex human behaviors such as speech perception and articulation, and the execution and appreciation of music and dance also develop in the millisecond-to-second scale (Kotz & Schwartze, 2010; Poeppel & Assaneo, 2020; Merchant et al., 2015a; Lenc et al., 2021). Motion processing in the visual and tactile modalities, as well as the coordination of fine movements occurs also in this time range (Merchant et al., 1997, 2001; Romo et al., 1995, 1996; Georgopoulos et al., 2007; Narselaris et al., 2006).

In trying to define time, we often include two critical processes: (Joris & van der Heijden, 2019) tracking events that change with time, and (Schnupp & Carr, 2009) predicting the occurrence of future events. Tracking events rely on the quantification of the elapsed time that has spanned since a relevant sensory or motor event and is critical to perceive the length of an interval. It is also critical to make decisions about the length of a stimulus in tasks that demand categorization, discrimination, or identification of a duration based on a cognitive rule (Merchant et al., 2013a). Predicting the occurrence of future events relies on the quantification of the remaining time that we must wait before an expected event. We know that the ability to predict future events is intrinsically linked to time intervals that had been learned through trial and error (Merchant & Yarrow, 2016). As we review in the section on the *taxonomy of timing*, recent neurophysiological and imaging studies support the notion that tracking elapsed time (how long) and anticipating a future event (when) might rely on different neuronal mechanisms and possibly different brain areas.

The perception and production of time in the hundreds of milliseconds is crucially related to a large repertoire of behaviors, it can be triggered or guided by different sensory modalities, specially audition, vision, and touch, and it can be expressed through a variety of effectors using the skeletomotor, oculomotor, and laryngeal systems (Fig. 1) (Merchant et al., 2013a; Merchant & Yarrow, 2016; Merchant & Bartolo, 2018). It is important to note that time intervals can be produced or estimated just once (called interval based timing), and they can also be generated in sequences of nonrhythmic intervals, or in sequences of periodic musical stimuli that possess a metric structure (called beat-based timing) (Fig. 1) (Merchant et al., 2015a; Lenc et al., 2021). Furthermore, some behaviors require an explicit representation of time intervals such as when tapping to a rhythm, while in other behaviors timing is carried covertly (or implicitly) such as when we draw, where timing is an emergent property of the produced hand trajectory (Merchant et al., 2008a; Zelaznik et al., 2002).

Timing can be performed in synchrony with periodic external events, as in the case of music played by an ensemble, or internally like in the case of a soloist (Repp & Su, 2013; Merchant et al., 2008b; Wing, 2002). Another property of the sensory input that we know modulate timing performance is whether the intervals are filled or empty. For example, empty intervals can be defined by presentation of brief sensory stimuli (such as auditory clicks), whereas filled intervals are defined by the onset and offset times of a continuously present stimulus. Interestingly, it has been shown that filled intervals are perceived as being longer than empty intervals of the same length, and that the discrimination threshold is smaller for empty than for filled intervals (Fig. 1) (Grondin & Rousseau, 1991; Grondin et al., 1998).

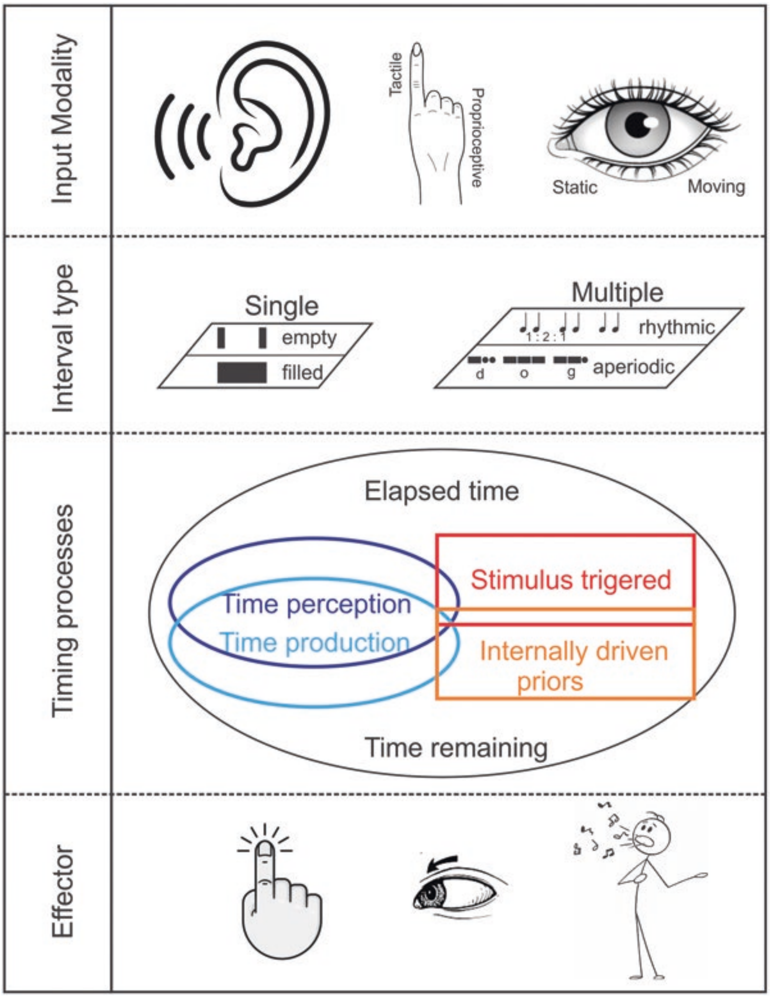


Fig. 1 Different timing processes can be flexibly engaged depending on the input modality (hearing, touch, or sight) and the type of interval (single or multiple). Timing expression can be channeled through the skeletomotor, oculomotor, and laryngeal systems

In summary, key elements of temporal processing include (Joris & van der Heijden, 2019) the time scale being quantified, (Schnupp & Carr, 2009) the modality of the stimulus that triggers a timing behavior, (Jeffress, 1948) whether the intervals are filled or empty, (O'Neill & Suga, 1979) whether the task involves single or multiple rhythmic or arrhythmic intervals, (Hastings et al., 2018) whether there is tracking of elapsed time or prediction of future events, (Drucker-Colín & Merchant-Nancy, 1996) whether time is being measured for a movement or for a perceptual decision, (Henderson et al., 2006) the effector used to express timing behavior, (Brody et al., 2003) whether timing is externally or internally generated, and (Bortoletto et al., 2011) the implicit or explicit nature of the timing task (Fig. 1). Thus, we would like to emphasize that timing most certainly is not a single process and probably depends on a set of neural mechanisms. Consistent with this idea, the last 10 years have seen an explosion of neurophysiological and imaging studies, suggesting that different brain circuits and different neural mechanisms are involved in the ability to tell and use time to control behavior. The second edition of our book strongly reflects this view.

Are Different Timing Behaviors Supported by Different Timing Mechanisms?

The present book focuses mainly on the neural basis of temporal processing in the hundreds of milliseconds range, although some of the chapters also deal with the underpinnings of timing behaviors in the seconds-to-minutes scale. Many authors defend the notion of different brain mechanisms for these two-time scales, and there is still some debate over the exact time boundary where these two scales might differ in their neuronal implementation. There is evidence that the time boundary seems to be in the order of 1300 to 2000 ms for perceptual and motor timing tasks that involve one interval or a set of isochronous intervals (see chapters of Simon Grondin, (Grondin, 2012)). However, some researchers maintain that these two-time scales are governed by the same neural clock during complex behaviors, such as the perception and execution of music with a complex hierarchical structure of tempi (see chapter of Sonja Kotz, (Rajendran et al., 2018)).

The classical model of time, often called scalar timing model, proposes a general multicontext clock composed by an internal pacemaker that monitors the passage of time once a gate signal is triggered, and it also includes an integrator that accumulates the pulses of the pacemaker (Treisman, 1963). Many psychophysical studies have successfully used this framework to explain the temporal performance of humans and animals in different perceptual and production timing tasks. At the core of this framework is the *scalar property of time*, showing that the variability of the temporal estimates increases linearly in proportion to total elapsed time (Gibbon, 1977; Gibbon et al., 1997). However, this model uses a black-box approach, in which the possible neural mechanisms behind the clock stage of the model are difficult to identify.

An alternative view emerged in the early 2000s, and it involves the notion of a ubiquitous timing mechanism, that depends on the dynamic properties of the cortical and subcortical recurrent networks of the brain (Karmarkar & Buonomano, 2007; Buonomano & Laje, 2010). The common clock hypothesis has been supported by fMRI meta-analyses reporting that the neuronal circuit composed of the medial premotor cortex (MPCs; the presupplementary motor area [preSMA], the supplementary motor area [SMA]), and the basal ganglia is engaged in many perceptual and motor timing tasks that span the hundreds of milliseconds range. The notion of a distributed timing mechanism is supported by modeling (Karmarkar & Buonomano, 2007; Zhou et al., 2022), brain slice recordings (Goel & Buonomano, 2014), and psychophysical approaches (Burr et al., 2007; Tonoyan et al., 2022).

A third possibility suggests the existence of a partially distributed timing mechanism, integrated by a main core of interconnected structures, such as the cortico-thalamic-basal ganglia circuit (CTBGc), and cortical areas that are selectively engaged depending on the specific behavioral requirement of a task (Merchant et al., 2013a, 2014a). These task-dependent areas may interact with the core timing system to produce the temporal behavior in a specific task (Merchant et al., 2015b). This recent proposal is based on psychophysical studies (Merchant et al., 2008a, 2008b), functional imaging meta-analysis (Wiener et al., 2010), and fresh neurophysiological observations (Betancourt et al., 2023). Importantly, this corpus of evidence supports neither the existence of either a common timing mechanism that functions equally every time a subject quantifies time, nor a set of timing mechanisms that are specific for each task context.

The Internal Simulation Hypothesis

A recent proposal has been put forward suggesting that the process we call timing is, in fact, an internal re-enactment of the sensory and motor actions that define the timing task to be solved in a particular behavioral context (de Lafuente et al., 2022). For example, under this internal simulation framework, if a subject is asked to time a one-second interval, they would replay the memory associated with a one-second interval.

Importantly, subjects do not make use of a general "one-second" memory, but the specific memory generated during the learning of that task. It is this long-term memory that allows the brain to conduct an internal simulation of the future states of the world.

This is closely related to the often-overlooked fact that each timing task has an almost unique experimental design, a set of very particular sensory stimuli, and, importantly, a set of specific behavioral rules that subjects must follow to solve the task. In this manner, the learning of a timing task generates a unique long-term memory or engram.

Support in favor of the internal simulation hypothesis has come recently from recordings made in nonhuman primates trained in a timing and rhythm perception task (García-Garibay et al., 2016). In this task, subjects must perceive and then internally maintain tempos defined by a visual metronome. Interestingly, when the visual metronome is extinguished and no movements are required from the subjects, the activity of single neurons and of the local field potentials (LFPs) continue to oscillate at the tempo of the metronome. This oscillatory activity was observed in V4, parietal cortices LIP and MIP, SMA, PFC, and the hippocampus (de Lafuente et al., 2022).

Under the internal simulation hypothesis, which we can also call the engram-replay hypothesis, there would be neither dedicated clocks nor specialized mechanism to measure or produce time. Instead, the brain would be using the survival-critical ability to make use of past information to predict the possible future states of the world, make plans, and act accordingly. We consider important to suggest that this new framework might encompass previous important works under a unifying point of view. This new point of view suggests that the brain does not perform "timing" as traditionally has been thought. Instead, to generate well-timed behavior, it re-enacts the sensory-motor engrams related to the task, that we as researchers instruct our subjects to perform (de Lafuente et al., 2022).

Supporting this viewpoint on how timed behavior might be achieved, recent electrophysiological evidence has shown that the neuronal activity of the motor cortices is able to oscillate with different time intervals, depending on the tempo that the subjects are holding actively on working memory (Cadena-Valencia et al., 2018). Interestingly, this oscillatory activity is also observed in parietal, V4, frontal, and hippocampal areas, and these patterns of activity are observed in the absence of any motor action or sensory stimuli (de Lafuente et al., 2022).

Taxonomy of Timing: A Second Attempt

In the first edition of our book *Neurobiology of Interval Timing,* we suggested an initial scheme for the classification of timing process. The proposal included sensory, perceptual, and motor timing as the pillars for temporal processing classification. Here, we used this scheme as a foundation, and we further suggest a new critical element: the dynamic interaction between sensory, cognitive, and motor areas to flexibly cope with the behavioral demands for time quantification and event prediction.

Sensory Timing

Organisms can extract temporal information from stimuli of all sensory modalities, even if there is no sensory organ for measuring the passage of time (Fig. 1). We still do not know how time is computed from the activation of different sensory systems or where in the sensory hierarchy is the temporal information computed for perceptual or motor purposes. To answer these fundamental questions, it is important to understand the anatomical and functional relationships that exist between the auditory, visual, and somatosensory systems, which correspond to the most important modalities for temporal information processing, particularly in the hundreds of milliseconds range. These sensory systems share a set of common operations: the sensory transduction of physical information into action potentials in the sensory receptors; the projection

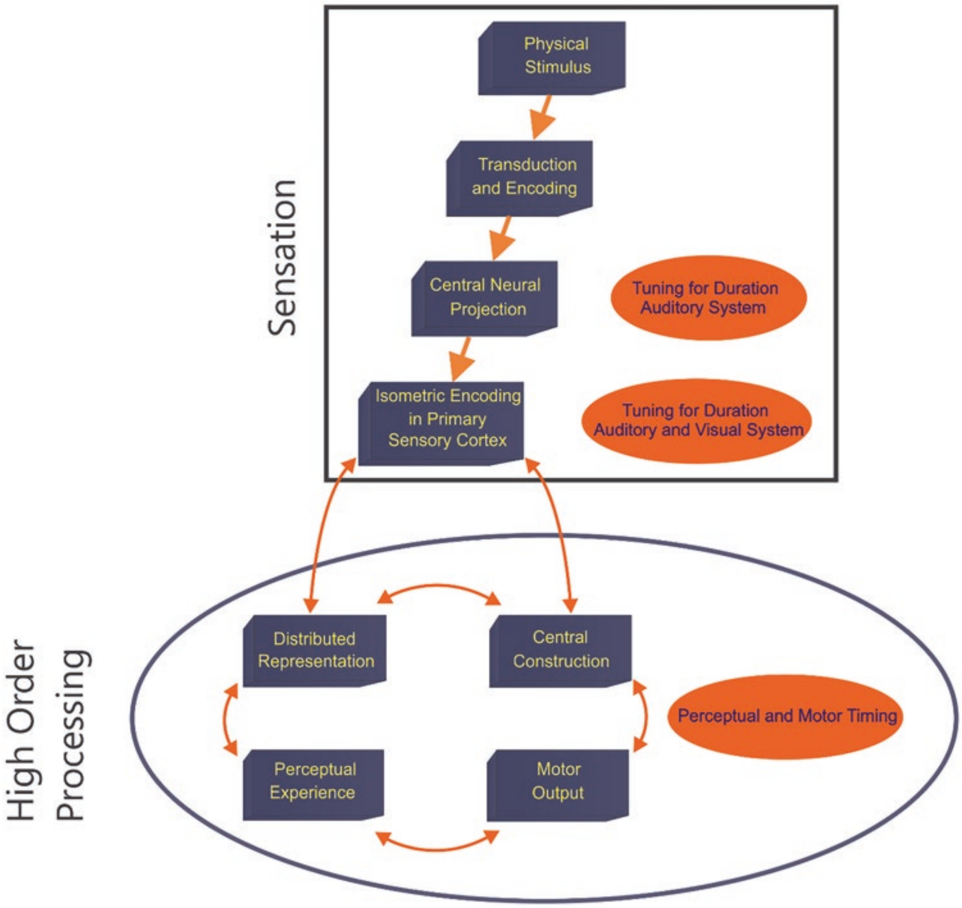


Fig. 2 A flux diagram for stimulus processing in the auditory, visual, and somatosensory modalities, ranging from sensation (square) to high order processing (ellipse). The orange ellipses highlight the levels of processing where different aspects of time quantification may occur

of this information (through thalamic nuclei) to the primary sensory areas of the cerebral cortex; the processing of the different aspects of the stimuli in the cortical and subcortical circuits; and finally, the use of high order sensory processing for perception, learning and memory, and voluntary motor action (Fig. 2). Thus, time information could be initially estimated from the transduction of the stimulus and the encoding of its physical properties in the first relays of the sensory systems. Few studies have focused on temporal processing during the transduction and conduction stages of sensation. In this regard, the auditory midbrain of many vertebrates contains cells that are tuned to the duration of stimuli in the range of tens of milliseconds (10–100 ms) that are also tuned to the stimulus frequency (Alluri et al., 2016). Studies across vertebrates have identified cells with preferred durations and auditory bandwidths for single intervals that mirror the range of species-specific vocalizations (57 for a review). Therefore, the auditory system can efficiently extract temporal information early on within the stimuli processing hierarchy, producing a switch from neural responses driven by the stimulus temporal profile to cells with average rate-tuning to durations.

The auditory cortex also shows duration selectivity for single intervals. In both cats and monkeys, cells show selectivity for the duration of auditory stimuli (He et al., 1997; Brosch et al., 1999), with preferred durations that are distributed over a wider range and for longer durations (50–350 ms) as compared to the inferior colliculus (Aubie et al., 2012).

The migration of time interval selectivity toward longer durations in the auditory cortex can be explained by the integration of duration selectivity from inferior colliculus inputs or from the stimulus temporal integration, where both temporal summation and suppression of responses within the circuits of the auditory cortex are taking place (He et al., 1997; Brosch et al., 1999). A key aspect of duration tuning implies the existence of different populations of cells encoding different intervals, filtering the continuous passage of time in subpopulations for short, medium, and long intervals. Thus, the fast identification of sounds would allow animal communication, echolocation, and even human language perception though funneling elapsed time into subpopulations of duration-tuned cells that act as feature detectors. Importantly, these feature detectors are hardwired in the inferior colliculus and auditory cortex as labeled lines for fast temporal processing (Zarco & Merchant, 2009). Overall, these studies suggest that the auditory modality has an outstanding ability to extract temporal information for single intervals in the range of early hundreds of milliseconds across the first relays of sensory processing. This indicates that the auditory modality is shaped for temporal processing.

For rhythmic auditory stimuli, recent neurophysiological studies have shown that sensory adaptation in the inferior colliculus and the auditory cortex of the rat is a critical mechanism for beat extraction. Importantly, the duration tuning of auditory cortex defines a response asymmetry that can detect stimuli that are on- or off-beat for real music excerpts, and this beat contrast selectivity aligns well with the beat tapping preference that human subjects produce in response to the same musical excerpts (Rajendran et al., 2017, 2018). Therefore, these findings suggest that the precision of the temporal coding in the inferior colliculus and auditory cortex predisposes the reliability of the estimated beat, even in the case of real music (Rajendran et al., 2020).

For vision, the first node in the visual pathway that shows duration-tuned cells for single intervals is the primary visual cortex (V1; Fig. 2). These cells show an orderly change in response magnitude to a visual stimulus that is presented in their receptive field (Duysens et al., 1996). The range of durations represented in V1 spans 50–400 ms. Interestingly, no such tuned cells have been found in the lateral geniculate nucleus of the thalamus (Duysens et al., 1996), suggesting that time selectivity is a property arising from local processing in V1. Psychophysical studies have investigated the sensory adaptation for the temporal properties of stimuli, an effect that probably depends on the primary sensory cortical areas. For instance, the apparent duration of a visual stimulus can be modified in a local region of the visual field by adaptation to oscillatory motion or flicker, suggesting that there is a spatially localized temporal mechanism for the sensation of time of visual events in the first nodes of the cortical hierarchy of visual processing (Tonoyan et al., 2022). As far as we know, there are no studies documenting neuronal responses to periodic or nonperiodic stimuli in V1.

The tactile system possesses three receptors to transduce mechanical stimuli: the rapidly adapting Pacinian, the rapidly adapting Meissner, and the slowly adapting Merkel (Romo et al., 1998). The primary somatosensory cortex contains a columnar organization that faithfully encodes the properties of these three mechanoreceptors (Romo et al., 1995; Mountcastle et al., 1969). Interestingly, rapidly adapting cells in the somatosensory cortex increase their discharge rate as a function of the duration of a moving probe in their finger receptive field (Romo et al., 1995, 1996). Thus, the primary somatosensory cortex also has a finely tuned machinery to extract temporal information (Luna et al., 2005).

Neurophysiological studies of time processing have provided evidence in favor of the idea that the auditory modality has a privileged capability for time quantification. Indeed, the precision of temporal estimations, measured in psychophysical tasks on humans, is more accurate and less variable when the intervals are defined by auditory as compared to visual or tactile stimuli. This increased precision has been observed during the perception of single and multiple intervals, as well as in the production time intervals. Interestingly, the time intervals marked by auditory signals are judged to be longer than those defined by visual stimuli (Wearden et al., 1998; Grondin & Rousseau, 1991; Grondin et al., 1998).

Thus, the perception of time passage seems to depend on specialized groups of cells in early nodes of the sensory processing which contain neurons that that are tuned to the duration of auditory and visual stimuli (Fig. 3). Consequently, it is evident that how we perceive the passage of time in the tens to hundreds of milliseconds, is highly influenced by the sensory modality, and depends on the anatomofunctional properties of each sensory system, where feature detectors are hardwired and where hearing has a clear advantage in temporal processing.

The Perception of Time

Once the sensory information is encoded as brief events or as a pattern of populations' neural activity tuned to the duration of the input stimuli, the neural system needs to translate sensation into a subjective sense of time, within the context of a specific behavior or task (Fig. 2). The most used tasks to measure time perception are the categorization and discrimination of time intervals (Merchant et al., 2008a, 2008b, 2008c). The former comes with two flavors: time bisection where subjects keep in memory the short and long prototypes to categorize intermediate probes (Wearden, 1991; Ng et al., 2011; Mendez et al., 2011), and time generalization where subjects keep in memory the boundary between the short and long categories and use it to decide on each trial which intervals are short or long (Wearden, 1992). The latter implies the comparison between the working memory trace of the first interval with the duration of a second interval to determine which one is the longer and produces a response that expresses that decision. Therefore, time perception tasks demand not only the encoding of an interval defined by an empty or filled stimulus duration, but also maintaining in memory this quantity and to categorize it or discriminate the memory trace with a second interval using specific decision rules and learned movements to express the decision (Merchant et al., 2008a, 2008b). Consequently, the representation of elapsed time needs to be dissociated from the memory trace of an interval (Chiba et al., 2015), the decision making (Merchant et al., 2011a, 2014b), and the voluntary motor signals to communicate the decision (Hernández et al., 2010; Méndez et al., 2014). This is not a trivial analytical problem, since the brain areas engaged in timing, such as the medial premotor areas, the putamen and motor thalamus, and the prefrontal and parietal areas, are also deeply involved in executive functions, working memory, and voluntary motor control (Mendoza & Merchant, 2014; Romo & Rossi-Pool, 2020; Caminiti et al., 2010; Miller & Cohen, 2001).

A recent neurophysiological study from our lab showed that the primate presupplementary cortex (preSMA) shows a boundary signal for interval categorization. These neurons showed an up-down profile of activation with a time peak that corresponded to the subjective limit between the short and long category. Notably, the time at which this peak is reached changes according to the categorical boundary of the current block, predicting the monkeys' categorical decision on a trial-by-trial basis (Mendoza et al., 2018). In addition, preSMA shows strong neural signals for the categorical choice made by the monkey and for the outcome of the categorization (Romo et al., 1993, 1997). Paradoxically, this internal prediction signal was not preceded by neurons encoding the elapsed time between the beginning and end of the intervals to be categorized. This could be due to the overtraining of the monkeys in this task (Mendez et al., 2011), which could favor the neural representation of boundary prediction than of elapsed time, while promoting the migration of time encoding to the basal ganglia (Merchant et al., 1997), where elapsed time is represented during a categorization task (Gouvêa

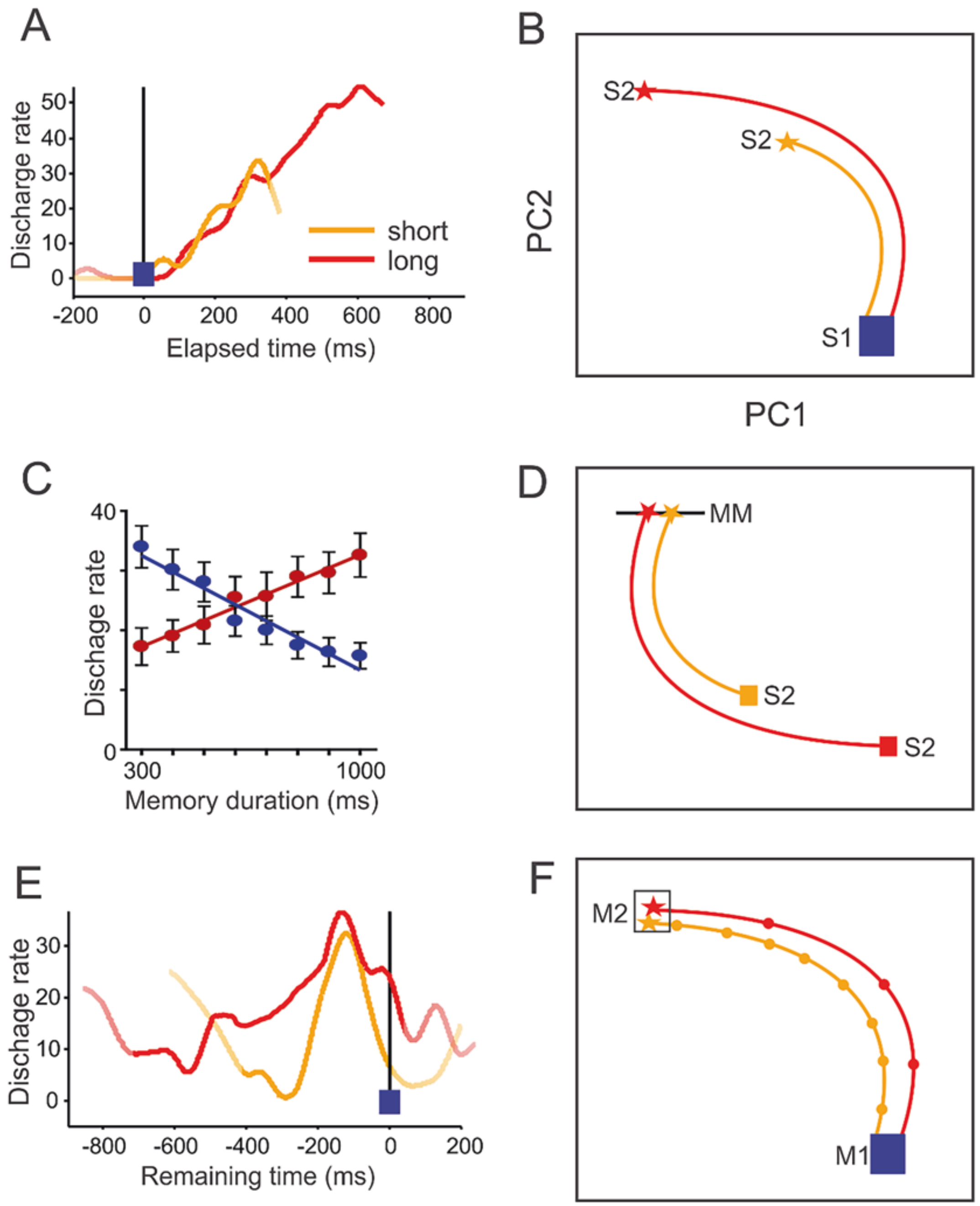


Fig. 3 (**a**) Ramping activity of a cell whose discharge rate increases with elapsed time. (**b**) Neural trajectories of a cell population starting at the same state space position at the first stimulus (S1) and encoding elapsed time in the final position of the trajectory at the time of the second stimulus (S2). (**c**) Two neurons show an increase (red) or decrease (blue) in discharge rate as a function of the memorized interval. (**d**) Neural trajectories reaching an attractor (MM during the memory delay). (**e**) Cell with a ramping activity that reaches a peak few milliseconds before the predicted event, with a larger activation slope for the shorter interval. (**f**) Neural trajectories that follow the same path but with larger speed for the shorter interval, generating temporal scaling

et al., 2015). Accordingly, recordings in monkeys performing a duration discrimination task showed single cell activity related with the encoding of elapsed time for the first interval in the striatum (Chiba et al., 2015) but not in dorsolateral prefrontal cortex, where the neurons might encode either the memory trace of the first duration, the decision rule, or the decision choice (Oshio et al., 2006; Genovesio et al., 2009).

Elapsed time can be encoded in the ramping activity of cells whose firing rates increase as a function of time and peak at preferred times (sand clock neurons); or can also be encoded by increments or decrements in firing activity that covers the length of the times interval (Merchant et al., 2011b; Knudsen et al., 2014; Henke et al., 2021) (Fig. 3a). Importantly, it is likely that thousands of ramping neurons might constitute a population clock (Merchant & Averbeck, 2017). Thus, the elapsed time between two stimuli can be represented by the dynamic interactions across the neural populations that display neural sequences (Merchant et al., 2015b; Crowe et al., 2014; Zhou et al., 2022).

This recent proposal of encoding time through dynamic neuronal activity uses dimensional reduction techniques to project the high-dimensional individual neural activity of a network into a low-dimensional state space to study the underlying computational principles of timing. This approach has been used to show that the elapsed time between events could be encoded by the final position of the neural population trajectory. Given that the trajectories triggered by the first event follow a stereotypic path for all durations, a simple linear decoder can determine time as the position of the trajectory at the second event (Fig. 3b) (Gouvêa et al., 2015; Kim et al., 2013; Sohn et al., 2019; Merchant & Pérez, 2020). Therefore, this neural population clock, observed in SMA, the striatum, and prefrontal cortex, can be started and stopped flexibly. However, this coding scheme is less efficient than the feature detectors in the midbrain or the sensory cortex that are tuned to durations. In fact, it is not yet clear whether primary sensory areas use the population mechanism for encoding elapsed time.

An important question in the encoding–decoding of time by neuronal activity is the nature of the clock reader. The final position of the neural trajectories could be used to generate a working memory representation of time for discrimination or for action triggering. Recurrent network modeling and monkey neurophysiology have shown that interval working memory might be represented as a manifold in the space generated by the neural trajectories (Fig. 3d), while at the single cell level, neurons show an increase or decrease in discharge rate as a function of the memorized duration (Chiba et al., 2015; Bi & Zhou, 2020) (Fig. 3c).

On the other hand, the study of perceptual interval learning and the generalization properties of such learning has provided important insights into the neural underpinnings of multimodal temporal information processing. For example, using interval discrimination, it has been shown that intensive learning can generalize across untrained auditory frequencies (Wright et al., 1997; Karmarkar & Buonomano, 2003), sensory modalities (Nagarajan et al., 1998; Westheimer, 1999), stimulus locations (Nagarajan et al., 1998), and even from sensory to motor timing tasks (Meegan et al., 2000; Sánchez-Moncada et al., 2020). However, none of these studies found generalization toward untrained interval durations. In addition, it has been suggested that the learning transfer depends on the improvement of temporal processing and not on more efficient memory or decision processes, at least for auditory interval discrimination (Wright et al., 1997). Therefore, these findings not only support the notion of a centralized or a partially overlapping distributed timing mechanism, but they also introduce the concept of duration-specific circuits. Regarding the first point, we can speculate that the timing signals sent from the primary sensory cortical areas to the large and distributed core timing network during the learning period may increase the global efficiency of the temporal information processing. Thus, an efficient core timing network will transfer its improved timing abilities across senses and between perceptual and motor contexts. A recent investigation found that only the subjects that can learn to efficiently discriminate visual intervals, and that show a generalization gain in tapping performance after learning, show a concomitant increase in activity in SMA, putamen, and the cerebellum (Sánchez-Moncada et al., 2020).

Another crucial aspect is that subjective time is prompt to distortions and generally does not have a one-to-one relation with physical time. As mentioned above, many factors can affect timing, and it is evident that temporal distortions can happen at different processing levels. For example, adaptation to fast visual motion strongly reduces the duration estimation of a subsequent stimulus, using a nonretinotopic reference frame. In addition, the duration of larger, brighter, or more numerous stimuli (Xuan et al., 2007; Togoli et al., 2021) is perceived to be longer than stimuli with the same duration but smaller magnitudes. Subjective time is dilated by the temporal frequency of moving flickering displays (Kanai et al., 2006) and affected by the contexts in which the stimuli are presented (Fornaciai et al., 2018; Karmarkar & Buonomano, 2007). Notably, time perception is also affected by movement (Merchant & Yarrow 2016). For example, when making a saccade to a target, a temporal expansion is produced (Yarrow et al., 2004).

Overall, the study of the neural mechanisms behind time perception would be greatly advanced by making use of experimental designs where time distortions are induced in a controlled fashion to determine their origins in the dynamics of neural population clocks across the timing circuits (De Kock et al., 2021). As a complement, studies using different modalities or contexts that share temporal resources can shed light into the properties of a core timing circuit and its interaction with context-dependent areas.

Motor Timing

Instead of reacting to sensory stimuli, the motor system can anticipate the appearance of future events. As we mentioned before, interval timing within the milliseconds range is a prerequisite for many complex behaviors, such as perception and production of speech (Kotz & Schwartze, 2010; Poeppel & Assaneo, 2020), the execution and appreciation of music and dance (Merchant et al., 2015a), and the performance of sports (Merchant & Georgopoulos, 2006). During single interval production tasks, the activity of neural populations evolves with similar trajectories across durations, reaching a common terminal state when the movement is triggered. Crucially, the trajectories are temporally scaled stretching for short and compressing for long intervals (Bi & Zhou, 2020; Wang et al., 2018; Merchant et al., 2011b).

At the single cell level, neurons encode the time to an event as ramping activity that reaches a peak shortly before the estimated time of the interval (Fig. 3c). This mechanism has been described in SMA, prefrontal cortex, and the striatum (Merchant et al., 2011b; Henke et al., 2021; Merchant & Bartolo, 2018; Kunimatsu et al., 2018). On the other hand, when monkeys produce rhythmic taps in synchrony with a metronome, neural trajectories show three main properties. First, they have circular dynamics that form a regenerating loop for every produced interval. Second, they converge to a similar state space right at the tapping time, resetting the beat-based clock at this point. Finally, the periodic trajectories increase in amplitude as a function of the length of the isochronous beat (Balasubramaniam et al., 2021; Gámez et al., 2019; Lenc et al., 2021) but also are temporal scaled (Betancourt et al., 2022). Hence, single interval and beat-based timing seem to have partially shared neural mechanisms (see chapter "Cognition of Time and Thinking Beyond").

Music and dance are behaviors that depend on intricate loops of perception and action, where temporal processing can be involved during the synchronization of movements with sensory information or during the internal generation of movement sequences (Rajendran et al., 2018). Many functional imaging studies have demonstrated that the circuits engaged in the perception of time are the same that are activated during motor timing (Coull et al., 2008; Wiener et al., 2010; Schubotz et al., 2000; Merchant et al., 2013a). The cortico-basal ganglia-thalamo-cortical circuit (CBGT), that includes the medial

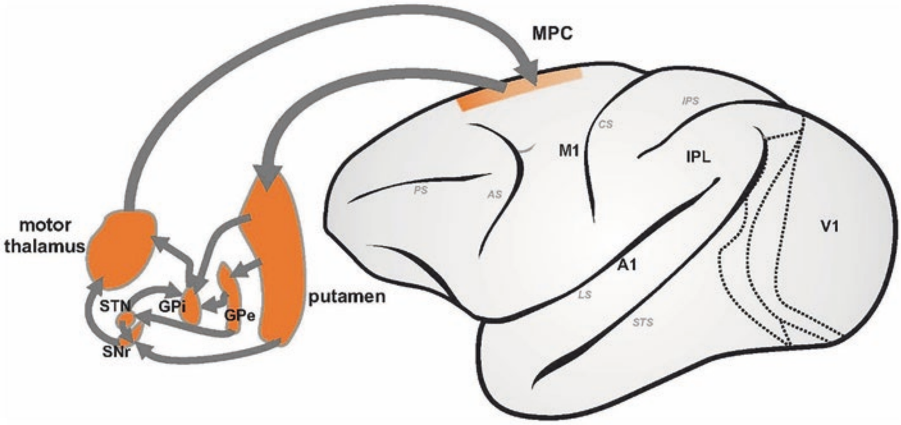


Fig. 4 The core timing network in the monkey brain is highlighted as orange structures. This core timing network is constituted by the skeleton-motor cortico-basal ganglia-thalamo-cortical circuit, which includes the medial premotor areas (MPC: SMA and preSMA), as well as the putamen, the globus pallidus, and the motor thalamus. A1 primary auditory area, AS arcuate sulcus, CS central sulcus, GPe globus pallidus externus, GPi globus pallidus internus, IPL inferior parietal lobe, IPS intraparietal sulcus, LS lateral sulcus, MPC medial premotor cortex, PS principal sulcus, SNr substantia nigra pars reticulata, STN subthalamic nucleus, STS superior temporal sulcus

premotor areas (Supplementary (SMA) and Presupplementary motor areas (preSMA), as well as the neostriatum, the globus pallidus, and the motor thalamus), is a network that is strongly engaged by interval perception and by time-constrained movements. These studies support the notion that the CBGT circuit is a key element of the core timing network, and that it is activated during the categorization and discrimination of time intervals, as well as during the perception and production of rhythms (Fig. 4). Accordingly, the neural population trajectories have now been characterized within the core timing network during perceptual and motor timing tasks for single intervals, as well as for rhythmic tapping tasks. Nevertheless, further studies are needed to determine whether the rules of time processing change in the core timing circuit when the same subjects perform perceptual and motor timing tasks, or when they perform single interval and rhythmic timing tasks. Another critical unanswered question relates to the role of the output of the basal ganglia in shaping the geometry and dynamics of the medial premotor neural trajectories during all these tasks.

In addition to the neural dynamics in population trajectories, different laboratories have shown that different core timing areas contain neurons that show duration tuning during motor timing tasks. Interval tuning during single interval and beat-based timing has been reported in medial premotor areas (Merchant et al., 2013b; Mita et al., 2009), prefrontal cortex (Henke et al., 2021), the putamen (Bartolo et al., 2014; Bartolo & Merchant, 2015), the caudate (Kameda et al., 2019; Kunimatsu et al., 2018), and the cerebellum (Ohmae et al., 2017; Okada et al., 2022). In addition, a chronomap in the medial premotor cortex has been described in humans using functional imaging, where interval-specific circuits show a topography with short preferred intervals in the anterior and long preferred intervals in the posterior portion of SMA/preSMA (Protopapa et al., 2019). Therefore, motor timing depends on distinct timing circuits composed by duration specific neurons. Each of these circuits quantify the time remaining for an event by contracting or expanding their activity patterns using temporal scaling (Merchant & Bartolo, 2018; Bartolo & Merchant, 2009). In fact, tuning and modularity are mechanisms for division of labor that are widely used in cortical and subcortical circuits to represent sensory, cognitive, and motor information (Mountcastle, 1998; Goldman-Rakic, 1984; Georgopoulos et al., 2007; Merchant et al., 2008d; Naselaris et al., 2006). Interval tuning can provide large flexibility to encode the passage of time and to predict events across many behaviors that require temporal processing by integrating timing with other task-dependent parameters with different mapping frameworks (Merchant & Bartolo, 2018; Zhou et al., 2022).

A set of functional imaging studies have revealed the neural and functional overlap between perceptual and motor timing, and the conclusion is that the motor network of the CBGT is activated across a wide range of timing contexts. A critical question, then, is what is the meaning of this anatomofunctional overlap? One possibility is that the increase in the BOLD signal in the motor areas across timing tasks reflects the presence of confounding cognitive processes, such as effector selection and motor preparation, or working memory and decision processes (Sánchez-Moncada et al., 2020). This is unlikely, however, since SMA, the CBGT circuit, and the prefrontal cortex are selectively activated even when duration estimates are registered with a perceptual discrimination (Wiener et al., 2010), or after motor preparation and/or execution processes have been rigorously controlled for (Coull et al., 2008; Schubotz et al., 2000) (see Chapter 11 of Coull and Morillon). Another possibility is that core timing network shares the neural circuitry with motor function because our general sense of time has been developed through action since childhood (Fraisse, 1982; Levin, 1992). This proposal is similar in principle to other embodied theories of time perception (Wittmann, 2013). Developmental studies have demonstrated that young children appear to represent time in motor terms (Droit-Volet & Rattat, 1999). Their duration estimates are more accurate when the duration is filled with an action than when it is empty (Fraisse, 1982), and they find it difficult to dissociate an estimate of duration from the motor act itself (Droit-Volet et al., 2006). Hence, it is possible that the motor circuits are engaged early in development to build up and acquire representations of time, forming a core timing network inside the motor system.

As has been proposed before, different cognitive functions may share the same neural representations and circuits for action and perception (Merchant et al., 2015a; Mendoza & Merchant, 2014). In the case of temporal processing, it is possible that the learned associations between particular actions and their durations have been engrained in the dynamics of the cortical and subcortical motor networks (Mendoza et al., 2016, 2018; Merchant & Averbeck, 2017; Méndez et al., 2014). Thus, the dynamic representation of time in the activity of cell populations could become a generalized temporal representation, which is independent of the motor output, and can be used for motor and perceptual acts that require a strict temporal control (Fig. 4). Longitudinal experiments recording the activity of the core timing network and context-dependent structures during learning a timing skill are required to test these ideas.

Sensory-Motor Loops

The strong correlation between fluctuations in the speed of the trajectories with trial-by-trial changes in movement times supports the notion of a robust predictive signal that triggers behavior within the core timing network. This motor top-down prediction should dynamically interact with the sensory bottom-up input to generate calibration loops of the timing system for both single interval and beat-based timing. In fact, it has been shown that predictive signals associated with a rhythmic motor behavior are fed back to the sensory areas through a corollary discharge that enhances processing of incoming auditory signals at a particular cyclic phase (Morillon et al., 2014). During beat-based timing, the motor system routinely produces dynamic signals in order to internally represent time, predict movements, enhance sensory events, and coordinate all these parameters (Merchant et al., 2015a; Merchant & Yarrow, 2016). Regarding the coordination process, two error signals most be computed: the time difference between the motor corollary discharge, and the proprioceptive reafference and the time difference between the internal beat signal and the sensory input (Repp, 2005). Both error signals could be used to calibrate the prediction of interval length (Betancourt et al., 2023). There is a large knowledge gap about how the motor and sensory timing areas accomplish these processes.

Bayesian Timing

As a final comment, the study of the neural basis of timing has been recently enriched using Bayesian approaches that can successfully account for behavioral performance across different timing tasks. This approach has also been used to dissociate the neural signals involved in tracking time, from those related to the acquired knowledge of the task parameters to optimize behavior. Bayesian estimators compute the posterior probability as the product of the likelihood function and the prior probability distribution. It has been shown that, across timing tasks, subjects tend to overestimate short intervals and overestimate long intervals, an effect described by the Viederord's Law, and now have been recently called regression toward the mean or bias effect (Jones & Mcauley, 2005; Pérez & Merchant, 2018). The prior distribution explains the bias effect by mapping the history of sample intervals that the observer has encountered during a block of trials with a mean that is close to the intermediate interval in the input distribution. Larger bias effects are captured as narrower prior distributions, reflecting more weight for previous knowledge of task conditions than the actual trial information. In addition, the likelihood function captures the scalar property of timing (Jazayeri & Shadlen, 2010; Merchant & Pérez, 2020; Perez et al., 2023). Notably, the curvature of the neural trajectories in SMA-preSMA reflects the effect of the prior on the bias effect (Sohn et al., 2019), while the variability of the neural trajectories is linked to the scalar properly of timing (Betancourt et al., 2023; Gámez et al., 2019).

Book Overview

It is with great pride and excitement that we present the second edition of the book *Neurobiology of Interval Timing*. A great deal of new and encouraging findings have been uncovered since the first edition 10 years ago. In this updated edition, we had the fortune to count among the contributions one of the finest researchers in the field of timing.

First, authors Vatakis and Teki provide us with a detailed recapitulation of the meetings and events that have now coalesced into a mature timing research field. From the first TIMELY meeting back in 2007, up until last year, these meetings brought together a highly interdisciplinary group of researchers, often with conflicting views about the future of field, which makes it more interesting for students and scientists alike.

In his excellent chapter, Prof. Grondin clearly exposes how the precision and accuracy of time estimates depend on a number of critical factors, such as the sensory stimuli utilized to define time intervals and whether the intervals are filled or empty, and importantly, the chapter provides a balanced view of the conflicting evidence pointing to the existence of a single universal Weber Fraction for time estimation, or, instead, to the fairly common observation that different time scales might result in different estimations of the value of Weber's Fraction.

Prof. Buonomano's chapter provides a very thorough and precise account of previous and current mechanistic models that have been put forward to account for the timing abilities of the brain. A remarkable observation that the authors point out is that across species and behavioral tasks, a pattern of sequential activation of neurons arises when the subjects are timing intervals. From this pattern, time can be decoded by looking at the different times at which neurons peak in their activity. Importantly, the authors make the argument that with such population clocks, the activity of each individual neuron is highly independent of the rest of the population. Hence, timing is in the neural sequences, oscillatory or ramping activity might not be sufficient to produce the flexible chronometers that are needed to account for our complex timing behavior.

Regarding electrophysiological findings in nonhuman primates, Prof. Tanaka and colleagues present and discuss an outstanding body of experimental results that strongly suggest that the basal ganglia are better suited to encode and execute timing behaviors within the range of hundreds of milliseconds, which contrasts with the activity that they recoded from the cerebellum, showing that this structure operates at shorter time scales. It is remarkable that these authors not only estimated correlation measurements between behavior and neuronal activity, but, importantly, they performed causal manipulations that provide firm support for their conclusions.

Also, in the field of primate neurophysiology, the group led by Prof. Merchant provides in their chapter a precise account of the numerous seminal experiments and results that they have produced in relation to the neuronal underpinnings of our ability to keep track of time and rhythms. This important work includes recordings from single neuron activity that could form the basis to estimate elapsed time from a past event, to estimate the time remaining to a future event, and to keep track of total elapsed time. Not only do these authors show single neuron activity, but, crucially, they demonstrate that the combined activity of large neuronal populations can account for the rhythmic tapping behavior of primates.

In their chapter, the research group led by Prof. Rueda-Orozco provides a compelling body of experimental evidence that convincingly indicates the basal ganglia do not only contain information about the muscles and articulations recruited for movements; moreover, the authors demonstrate that timing signals, such as elapsed time, constitute an integral part of basal ganglia activity. Such timing signals observed in the basal ganglia, the authors argue, might be strongly dependent on the rhythmic signals originating in sensory organs.

In the chapter by de Lafuente and colleagues, they review a body of recent literature from our group that provides support to the hypothesis that the brain estimates time and rhythm by predicting the occurrence of future stimuli and that also predicts the possible occurrence of required motor actions. They provide evidence indicating that monkeys and humans share the ability to internally maintain rhythms of different tempos. In terms of the neuronal correlates of this internal metronome, we showed that the neuronal activity within the supplementary motor area encodes the temporal and spatial characteristics of that metronome.

In a provocative and interesting review, Prof. Bi brings our attention to the large gap that stills exist between the knowledge gathered by basic science and the potential applications that this knowledge could have to understand and treat diseases like Parkinson's. Prof. Bi makes a strong case indicating that basic neuroscience and neurotechnology have been following parallel and largely independent pathways that should be brought close together. This chapter constitutes a much-needed critical view on the validity and capacity for generalization of our current results and our current approximations to the study of time and, more generally, to the study of cognition.

The chapter by Profs. Coull and Morillon makes an excellent presentation of the close interlink between auditory perception and our innate behavior to move in rhythm with periodic auditory stimuli. Moreover, the authors make a strong case arguing that it is the violations of our auditory expectations that drive us to search for and extract the beat in musical pieces. This important observation highlights the crucial function that prediction of future events plays in many aspects of brain function.

In a remarkable effort to review the large body of fMRI evidence, the chapter of Prof. Wiener provides us with an up-to-date view of the techniques that are being developed to make full use of many public databases that could contribute to a more general and unified understanding of the fMRI data currently available. The new meta-analyses presented by Prof. Wiener have provided further support for the existence of a "timing network." And within this network, the supplementary motor area is the one structure that consistently is found to be involved for solving timing tasks.

The chapter by Fleur and colleagues provides us with a well-documented well-discussed utility of the technique called Event-Related Potentials, and its successful use in probing the neuronal underpinning of rhythm processing in the brain. Such technique has the great advantage of not requiring conscious perception by the participant. The sole presentation of a predictable musical beat invariably induces brain responses that give raise to a "mismatch negativity" signal indicating that the brain has detected a sensory stimulus that is not expected to occur. The group led by Prof. Honing has produced numerous important insights into the possible brain mechanisms that could be used to detect such deviant or otherwise unexpected stimuli. It is worth noticing that this phenomenon is tightly coupled to the ability of the nervous system to predict the future states of the world.

In the chapter by Prof. Assaneo and colleagues, they describe the importance of rhythm in speech, and they make a strong case arguing that a speaker and a listener must share a—probably innate—common rhythmic structure that makes communication possible even in the most noisy and unpredictable environments.

Finally, the chapter by Schwartze and Kotz provides a solid argument in favor of the view that the basal ganglia constitute an integral node sustaining our capacity to estimate time and execute timely actions. As the authors describe, not only the basal ganglia are an integral part of the motor system, but they are also involved in modulating diverse brain structures such as the cerebellum and integrative areas of the cortex like the prefrontal cortex. Thus, the basal ganglia participate in timing not only by modulating movement execution, but also by modulating functions such as cognitive processes like attention, which we know is closely related to timing abilities.

We are certain that you will gain access to multiple insights developed by renowned colleagues, and you will enjoy reading it as much as we enjoyed the editorial process.

Acknowledgments We thank Luis Prado, Maria Antonieta Carbajo, Juan Ortíz, and Raul Paulín for their technical assistance supported by CONACYT: A1-S-8430, 151223, 167429, PAPIIT: IG200424, IA202024, IB200512, BG200521, and UNAM-DGAPA-PASPA.

References

Alluri, R. K., Rose, G. J., Hanson, J. L., Leary, C. J., Vasquez-Opazo, G. A., Graham, J. A., & Wilkerson, J. (2016). Phasic, suprathreshold excitation and sustained inhibition underlie neuronal selectivity for short-duration sounds. *Proceedings of the National Academy of Sciences, 113*(13), E1927–E1935.

Aubie, B., Sayegh, R., & Faure, P. A. (2012). Duration tuning across vertebrates. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 32*(18), 6373–6390.

Balasubramaniam, R., Haegens, S., Jazayeri, M., Merchant, H., Sternad, D., & Song, J. H. (2021). Neural encoding and representation of time for sensorimotor control and learning. *Journal of Neuroscience, 41*(5), 866–872.

Bartolo, R., & Merchant, H. (2009). Learning and generalization of time production in humans: Rules of transfer across modalities and interval durations. *Experimental Brain Research, 197*, 91–100. https://doi.org/10.1007/s00221-009-1895-1

Bartolo, R., & Merchant, H. (2015). β oscillations are linked to the initiation of sensory-cued movement sequences and the internal guidance of regular tapping in the monkey. *The Journal of Neuroscience, 35*, 4635–4640. https://doi.org/10.1523/JNEUROSCI.4570-14.2015

Bartolo, R., Prado, L., & Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *Journal of Neuroscience, 34*(11), 3910–3923.

Betancourt, A., Pérez, O., Gámez, J., Mendoza, G., & Merchant, H. (2022). Premotor population dynamics as neural substrate for auditory and visual rhythmic entrainment. bioRxiv, 2022-08.

Betancourt, A., Perez, O., Gamez, J., Mendoza, G., & Merchant, H. (2023). Amodal population clock in the primate medial premotor system for rhythmic tapping. *Cell Reports*. https://doi.org/10.1016/j.celrep.2023.113234

Bi, Z., & Zhou, C. (2020). Understanding the computation of time using neural network models. *Proceedings of the National Academy of Sciences, 117*(19), 10530–10540.

Bortoletto, M., Cook, A., & Cunnington, R. (2011). Motor timing and the preparation for sequential actions. *Brain and Cognition, 75*(2), 196–204.

Brody, C. D., Hernández, A., Zainos, A., & Romo, R. (2003). Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cerebral cortex (New York, NY: 1991), 13*(11), 1196–1207.

Brosch, M., Schulz, A., & Scheich, H. (1999). Processing of sound sequences in macaque auditory cortex: Response enhancement. *Journal of Neurophysiology, 82*(3), 1542–1559.

Buonomano, D. V., & Laje, R. (2010). Population clocks: Motor timing with neural dynamics. *Trends in Cognitive Sciences, 14*(12), 520–527.

Burr, D., Tozzi, A., & Morrone, M. C. (2007). Neural mechanisms for timing visual events are spatially selective in real-world coordinates. *Nature Neuroscience, 10*(4), 423–425.

Cadena-Valencia, J., García-Garibay, O., Merchant, H., Jazayeri, M., & de Lafuente, V. (2018). Entrainment and maintenance of an internal metronome in premotor cortex. *eLife, 7*, e38983. https://doi.org/10.7554/eLife.38983

Caminiti, R., Chafee, M. V., Battaglia-Mayer, A., Averbeck, B. B., Crowe, D. A., & Georgopoulos, A. P. (2010). Understanding the parietal lobe syndrome from a neurophysiological and evolutionary perspective. *European Journal of Neuroscience, 31*(12), 2320–2340.

Chiba, A., Oshio, K. I., & Inase, M. (2015). Neuronal representation of duration discrimination in the monkey striatum. *Physiological Reports, 3*(2), e12283.

Coull, J. T., Nazarian, B., & Vidal, F. (2008). Timing, storage, and comparison of stimulus duration engage discrete anatomical components of a perceptual timing network. *Journal of Cognitive Neuroscience, 20*(12), 2185–2197.

Crowe, D. A., Zarco, W., Bartolo, R., & Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *Journal of Neuroscience, 34*(36), 11972–11983.

De Kock, R., Gladhill, K. A., Ali, M. N., Joiner, W. M., & Wiener, M. (2021). How movements shape the perception of time. *Trends in Cognitive Sciences, 25*(11), 950–963.

de Lafuente, V., Jazayeri, M., Merchant, H., Gracía-Garibay, O., Cadena-Valencia, J., & Malagón, A. M. (2022). Keeping time and rhythm by replaying a sensory-motor engram. bioRxiv, 2022-01.

Droit-Volet, S., & Rattat, A.-C. (1999). Are time and action dissociated in young children's time estimation? *Cognitive Development, 14*(4), 573–595.

Droit-Volet, S., Delgado, M., & Rattat, A. C. (2006). The development of the ability to judge time in children. In *Focus on child psychology research* (pp. 81–104). Nova Science Publishers.

Drucker-Colín, R., & Merchant-Nancy, H. (1996). Evolution of concepts of mechanisms of sleep. In A. Kales (Ed.), *Handbook of experimental pharmacology, vol 116. Pharmacology of sleep* (pp. 1–28). Springer.

Duysens, J., Schaafsma, S. J., & Orban, G. A. (1996). Cortical off response tuning for stimulus duration. *Vision Research, 36*(20), 3243–3251.

Fornaciai, M., Togoli, I., & Arrighi, R. (2018). Motion-induced compression of perceived numerosity. *Scientific Reports, 8*(1), 6966.

Fraisse, P. (1982). The adaptation of the child to time. In *The developmental psychology of time* (pp. 113–140). Academic.

Gallistel, C. R., & Gibbon, J. (2000 Apr). Time, rate, and conditioning. *Psychological Review, 107*(2), 289–344.

Gámez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biology, 17*(4), e3000054.

García-Garibay, O., Cadena, J., Merchant, H., & de Lafuente, V. (2016). Monkeys share the human ability to internally maintain a temporal rhythm. *Frontiers in Psychology, 7*, 1971. https://doi.org/10.3389/fpsyg.2016.01971

Genovesio, A., Tsujimoto, S., & Wise, S. P. (2009). Feature-and order-based timing representations in the frontal cortex. *Neuron, 63*(2), 254–266.

Georgopoulos, A. P., Merchant, H., Narselaris, T., & Amirikian, B. (2007). Mapping of the preferred direction in the motor cortex. *Proceedings of the National Academy of Sciences, 104*(26), 11068–11072.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*(3), 279.

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology, 7*(2), 170–184.

Goel, A., & Buonomano, D. V. (2014). Timing as an intrinsic property of neural networks: Evidence from in vivo and in vitro experiments. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1637), 20120460.

Goldman-Rakic, P. S. (1984). Modular organization of prefrontal cortex. *Trends in Neurosciences, 7*(11), 421–424.

Gouvêa, T. S., Monteiro, T., Motiwala, A., Soares, S., Machens, C., & Paton, J. J. (2015). Striatal dynamics explain duration judgments. *eLife, 4*, e11386.

Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: Evidence from interval discrimination, reproduction, and categorization. *Journal of Experimental Psychology. Human Perception and Performance, 38*(4), 880–890.

Grondin, S., & Rousseau, R. (1991). Judging the relative duration of multimodal short empty time intervals. *Perception & Psychophysics, 49*(3), 245–256.

Grondin, S., Meilleur-Wells, G., Ouellette, C., & Macar, F. (1998). Sensory effects on judgments of short time-intervals. *Psychological Research, 61*(4), 261–268.

Hastings, M. H., Maywood, E. S., & Brancaccio, M. (2018). Generation of circadian rhythms in the suprachiasmatic nucleus. *Nature Reviews Neuroscience, 19*(8), 453–469.

He, J., Hashikawa, T., Ojima, H., & Kinouchi, Y. (1997). Temporal integration and duration tuning in the dorsal zone of cat auditory cortex. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 17*(7), 2615–2625.

Henderson, J., Hurly, T. A., Bateson, M., & Healy, S. D. (2006). Timing in free-living rufous hummingbirds, Selasphorus rufus. *Current Biology: CB, 16*(5), 512–515.

Henke, J., Bunk, D., von Werder, D., Häusler, S., Flanagin, V. L., & Thurley, K. (2021). Distributed coding of duration in rodent prefrontal cortex during time reproduction. *eLife, 10*, e71612.

Hernández, A., Nácher, V., Luna, R., Zainos, A., Lemus, L., Alvarez, M., et al. (2010). Decoding a perceptual decision process across cortex. *Neuron, 66*(2), 300–314.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience, 13*(8), 1020–1026.

Jeffress, L. A. (1948). A place theory of sound localization. *Journal of Comparative and Physiological Psychology, 41*(1), 35–39.

Jones, M. R., & Mcauley, J. D. (2005). Time judgments in global temporal contexts. *Perception & Psychophysics, 67*, 398–417. https://doi.org/10.3758/BF03193320

Joris, P. X., & van der Heijden, M. (2019). Early binaural hearing: The comparison of temporal differences at the two ears. *Annual Review of Neuroscience, 42*, 433–457.

Kameda, M., Ohmae, S., & Tanaka, M. (2019). Entrained neuronal activity to periodic visual stimuli in the primate striatum compared with the cerebellum. *eLife, 8*. https://doi.org/10.7554/eLife.48702

Kanai, R., Paffen, C. L., Hogendoorn, H., & Verstraten, F. A. (2006). Time dilation in dynamic visual display. *Journal of Vision, 6*(12), 1421–1430.

Karmarkar, U. R., & Buonomano, D. V. (2003). Temporal specificity of perceptual learning in an auditory discrimination task. *Learning & memory (Cold Spring Harbor, N.Y.), 10*(2), 141–147.

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: Encoding time in neural network states. *Neuron, 53*(3), 427–438.

Kim, J., Ghim, J. W., Lee, J. H., & Jung, M. W. (2013). Neural correlates of interval timing in rodent prefrontal cortex. *Journal of Neuroscience, 33*(34), 13834–13847.

Knudsen, E. B., Powers, M. E., & Moxon, K. A. (2014). Dissociating movement from movement timing in the rat primary motor cortex. *Journal of Neuroscience, 34*(47), 15576–15586.

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences, 14*(9), 392–399.

Kunimatsu, J., Suzuki, T. W., Ohmae, S., & Tanaka, M. (2018). Different contributions of preparatory activity in the basal ganglia and cerebellum for self-timing. *eLife, 7*. https://doi.org/10.7554/eLife.35676

Lenc, T., Merchant, H., Keller, P. E., Honing, H., Varlet, M., & Nozaradan, S. (2021). Mapping between sound, brain and behaviour: Four-level framework for understanding rhythm processing in humans and non-human primates. *Philosophical Transactions of the Royal Society B, 376*(1835), 20200325.

Levin, I. (1992). The development of the concept of time in children: An integrative model. In *Time, action and cognition* (pp. 13–32). Springer.

Lewis, P. A., & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: Evidence from neuroimaging. *Current Opinion in Neurobiology, 13*(2), 250–255.

Luna, R., Hernández, A., Brody, C. D., & Romo, R. (2005). Neural codes for perceptual discrimination in primary somatosensory cortex. *Nature Neuroscience, 8*(9), 1210–1219.

Meegan, D. V., Aslin, R. N., & Jacobs, R. A. (2000). Motor timing learned without motor training. *Nature Neuroscience, 3*(9), 860–862.

Mendez, J. C., Prado, L., Mendoza, G., & Merchant, H. (2011). Temporal and spatial categorization in human and non-human primates. *Frontiers in Integrative Neuroscience, 5*, 50.

Méndez, J. C., Pérez, O., Prado, L., & Merchant, H. (2014). Linking perception, cognition, and action: Psychophysical observations and neural network modelling. *PLoS One, 9*(7), e102553.

Mendoza, G., & Merchant, H. (2014). Motor system evolution and the emergence of high cognitive functions. *Progress in Neurobiology, 122*, 73–93.

Mendoza, G., Peyrache, A., Gámez, J., Prado, L., Buzsáki, G., & Merchant, H. (2016). Recording extracellular neural activity in the behaving monkey using a semichronic and high-density electrode system. *Journal of Neurophysiology, 116*, 563–574. http://www.ncbi.nlm.nih.gov/pubmed/27169505

Mendoza, G., Méndez, J. C., Pérez, O., Prado, L., & Merchant, H. (2018). Neural basis for categorical boundaries in the primate pre-SMA during relative categorization of time intervals. *Nature Communications, 9*(1), 1098.

Merchant, H., & Averbeck, B. B. (2017). The computational and neural basis of rhythmic timing in medial premotor cortex. *Journal of Neuroscience, 37*(17), 4552–4564.

Merchant, H., & Bartolo, R. (2018). Primate beta oscillations and rhythmic behaviors. *Journal of Neural Transmission, 125*, 461–470.

Merchant, H., & Georgopoulos, A. P. (2006). Neurophysiology of perceptual and motor aspects of interception. *Journal of Neurophysiology, 95*(1), 1–13.

Merchant, H., & Pérez, O. (2020). Estimating time with neural networks. *Nature Machine Intelligence, 2*(9), 492–493.

Merchant, H., & Yarrow, K. (2016). How the motor system both encodes and influences our sense of time. *Current Opinion in Behavioral Sciences, 8*, 22–27.

Merchant, H., Zainos, A., Hernández, A., Salinas, E., & Romo, R. (1997). Functional properties of primate putamen neurons during the categorization of tactile stimuli. *Journal of Neurophysiology, 77*(3), 1132–1154.

Merchant, H., Battaglia-Mayer, A., & Georgopoulos, A. P. (2001). Effects of optic flow in motor cortex and area 7a. *Journal of Neurophysiology, 86*(4), 1937–1954.

Merchant, H., Battaglia-Mayer, A., & Georgopoulos, A. P. (2003). Functional organization of parietal neuronal responses to optic-flow stimuli. *Journal of Neurophysiology, 90*(2), 675–682.

Merchant, H., Battaglia-Mayer, A., & Georgopoulos, A. P. (2004a). Neural responses during interception of real and apparent circularly moving stimuli in motor cortex and area 7a. *Cerebral Cortex, 14*(3), 314–331.

Merchant, H., Battaglia-Mayer, A., & Georgopoulos, A. P. (2004b). Neural responses in motor cortex and area 7a to real and apparent motion. *Experimental Brain Research, 154*, 291–307.

Merchant, H., Zarco, W., Bartolo, R., & Prado, L. (2008a). The context of temporal processing is represented in the multidimensional relationships between timing tasks. *PLoS One, 3*(9), e3169.

Merchant, H., Zarco, W., & Prado, L. (2008b). Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. *Journal of Neurophysiology, 99*(2), 939–949.

Merchant, H., Luciana, M., Hooper, C., Majestic, S., & Tuite, P. (2008c). Interval timing and Parkinson's disease: Heterogeneity in temporal performance. *Experimental Brain Research, 184*, 233–248.

Merchant, H., Naselaris, T., & Georgopoulos, A. P. (2008d). Dynamic sculpting of directional tuning in the primate motor cortex during three-dimensional reaching. *Journal of Neuroscience, 28*(37), 9164–9172.

Merchant, H., Zarco, W., Prado, L., & Pérez, O. (2009). Behavioral and neurophysiological aspects of target interception. *Advances in Experimental Medicine and Biology, 629*(201–20), 13.

Merchant, H., Crowe, D. A., Robertson, M. S., Fortes, A. F., & Georgopoulos, A. P. (2011a). Top-down spatial categorization signal from prefrontal to posterior parietal cortex in the primate. *Frontiers in Systems Neuroscience, 5*, 69.

Merchant, H., Zarco, W., Pérez, O., Prado, L., & Bartolo, R. (2011b). Measuring time with different neural chronometers during a synchronization-continuation task. *Proceedings of the National Academy of Sciences, 108*(49), 19784–19789.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013a). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience, 36*, 313–336.

Merchant, H., Pérez, O., Zarco, W., & Gámez, J. (2013b). Interval tuning in the primate medial premotor cortex as a general timing mechanism. *The Journal of Neuroscience: Tthe Official Journal of the Society for Neuroscience, 33*, 9082–9096. https://doi.org/10.1523/JNEUROSCI.5513-12.2013. http://www.ncbi.nlm.nih.gov/pubmed/23699519

Merchant, H., Bartolo, R., Perez, O., Mendez, J. C., Mendoza, G., Gamez, J., Yc, K., & Prado, L. (2014a). Neurophysiology of timing in the hundreds of milliseconds: Multiple layers of neuronal clocks in the medial premotor areas. *Advances in Experimental Medicine and Biology, 829*(1), 143–154. https://doi.org/10.1007/978-1-4939-1782-2_8

Merchant, H., Crowe, D. A., Fortes, A. F., & Georgopoulos, A. P. (2014b). Cognitive modulation of local and callosal neural interactions in decision making. *Frontiers in Neuroscience, 8*, 245.

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015a). Finding the beat: A neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*(1664), 20140093.

Merchant, H., Perez, O., Bartolo, R., Mendez, J. C., Mendoza, G., Gamez, J., Yc, K., & Prado, L. (2015b). Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. *The European Journal of Neuroscience, 41*(5), 586–602. https://doi.org/10.1111/ejn.12811

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*(1), 167–202.

Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., & Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nature Neuroscience, 12*, 502–507. https://doi.org/10.1038/nn.2272

Morillon, B., Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications, 5*(1), 5255.

Mountcastle, V. B. (1998). *Perceptual neuroscience: The cerebral cortex*. Harvard University Press.

Mountcastle, V. B., Talbot, W. H., Sakata, H., & Hyvärinen, J. (1969). Cortical neuronal mechanisms in flutter-vibration studied in unanesthetized monkeys. Neuronal periodicity and frequency discrimination. *Journal of Neurophysiology, 32*(3), 452–484.

Nagarajan, S. S., Blake, D. T., Wright, B. A., Byl, N., & Merzenich, M. M. (1998). Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 18*(4), 1559–1570.

Narselaris, T., Merchant, H., Amirikian, B., & Georgopoulos, A. P. (2006). Large-scale organization of preferred directions in the motor cortex II: Analysis of local distributions. *Journal of Neurophysiology, 96*(6), 3237–3247.

Naselaris, T., Merchant, H., Amirikian, B., & Georgopoulos, A. P. (2006). Large-scale organization of preferred directions in the motor cortex. I. Motor cortical hyperacuity for forward reaching. *Journal of Neurophysiology, 96*(6), 3231–3236.

Ng, K. K., Tobin, S., & Penney, T. B. (2011). Temporal accumulation and decision processes in the duration bisection task revealed by contingent negative variation. *Frontiers in Integrative Neuroscience, 5*, 77.

O'Neill, W. E., & Suga, N. (1979). Target range-sensitive neurons in the auditory cortex of the mustache bat. *Science, 203*(4375), 69–73.

Ohmae, S., Kunimatsu, J., & Tanaka, M. (2017). Cerebellar roles in self-timing for sub- and supra-second intervals. *The Journal of Neuroscience, 37*, 3511–3522. https://doi.org/10.1523/JNEUROSCI.2221-16.2017

Okada, K.-i., Ryuji, T., & Masaki, T. (2022). Neural signals regulating motor synchronization in the primate deep cerebellar nuclei. *Nature Communications, 13*, 2504. https://doi.org/10.1038/s41467-022-30246-2

Oshio, K. I., Chiba, A., & Inase, M. (2006). Delay period activity of monkey prefrontal neurones during duration-discrimination task. *European Journal of Neuroscience, 23*(10), 2779–2790.

Pérez, O., & Merchant, H. (2018). The synaptic properties of cells define the hallmarks of interval timing in a recurrent neural network. *Journal of Neuroscience, 38*(17), 4186–4199.

Perez, O., Delle Monache, S., Lacquaniti, F., Bosco, G., & Merchant, H. (2023). Rhythmic tapping to a moving beat: Motion kinematics overrules motion naturalness. bioRxiv, 2023-03.

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience, 21*(6), 322–334.

Protopapa, F., Hayashi, M. J., Kulashekhar, S., van der Zwaag, W., Battistella, G., Murray, M. M., Kanai, R., & Bueti, D. (2019). Chronotopic maps in human supplementary motor area. *PLoS Biology, 17*, e3000026. https://doi.org/10.1371/journal

Rajendran, V. G., Harper, N. S., Garcia-Lazaro, J. A., Lesica, N. A., & Schnupp, J. W. (2017). Midbrain adaptation may set the stage for the perception of musical beat. *Proceedings of the Royal Society B: Biological Sciences, 284*(1866), 20171455.

Rajendran, V. G., Teki, S., & Schnupp, J. W. (2018). Temporal processing in audition: Insights from music. *Neuroscience, 389*, 4–18.

Rajendran, V. G., Harper, N. S., & Schnupp, J. W. (2020). Auditory cortical representation of music favours the perceived beat. *Royal Society Open Science, 7*(3), 191194.

Repp, B. H. (2005). Sensorimotor synchronisation; a review of the tapping literature. *Psychonomic Bulletin & Review, 12*, 969–992.

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review, 20*, 403–452.

Romo, R., & Rossi-Pool, R. (2020). Turning touch into perception. *Neuron, 105*(1), 16–33.

Romo, R., Ruiz, S., Crespo, P., Zainos, A., & Merchant, H. (1993). Representation of tactile signals in primate supplementary motor area. *Journal of Neurophysiology, 70*(6), 2690–2694.

Romo, R., Merchant, H., Ruiz, S., Crespo, P., & Zainos, A. (1995). Neuronal activity of primate putamen during categorical perception of somaesthetic stimuli. *Neuroreport, 6*(7), 1013–1017.

Romo, R., Merchant, H., Zainos, A., & Hernández, A. (1996). Categorization of somaestetic stimuli: Sensorimotor performance and neuronal activity in primary somatic sensory cortex of awake monkeys. *Neuroreport, 7*(7), 1273–1279.

Romo, R., Merchant, H., Zainos, A., & Hernández, A. (1997). Categorical perception of somesthetic stimuli: Psychophysical measurements correlated with neuronal events in primate medial premotor cortex. *Cerebral cortex (New York, NY: 1991), 7*(4), 317–326.

Romo, R., Zainos, A., Merchant, H., Hernández, A., & Garcia, G. (1998. Cap 7). Processing of somesthetic stimuli in primate sensory-motor cortex. In J. W. Morley (Ed.), *Neural aspects of tactile perception* (pp. 239–273). Elsevier Science.

Sánchez-Moncada, I., Concha, L., & Merchant, H. (2020). The timing network is engaged in the practice of internally driven tapping independently of the learning transfer from perceptual to motor timing. *bioRxiv*, 2020-12.

Sánchez-Moncada, I., Concha, L., and Merchant, H. (2024). Pre-supplementary motor cortex mediates learning transfer from perceptual to motor timing. Journal of Neuroscience. https://doi.org/10.1523/JNEUROSCI.3191-20.2023.

Schnupp, J. W. H., & Carr, C. E. (2009). On hearing with more than one ear: Lessons from evolution. *Nature Neuroscience, 12*(6), 692–697.

Schubotz, R. I., Friederici, A. D., & von Cramon, D. Y. (2000). Time perception and motor timing: A common cortical and subcortical basis revealed by fMRI. *NeuroImage, 11*(1), 1.

Sohn, M.-H., & Carlson, R. A. (2003). Implicit temporal tuning of working memory strategy during cognitive skill acquisition. *The American Journal of Psychology, 116*(2), 239–256.

Sohn, H., Narain, D., Meirhaeghe, N., & Jazayeri, M. (2019). Bayesian computation through cortical latent dynamics. *Neuron, 103*(5), 934–947.

Togoli, I., Fedele, M., Fornaciai, M., & Bueti, D. (2021). Serial dependence in time and numerosity perception is dimension-specific. *Journal of Vision, 21*(5), 6–6.

Tonoyan, Y., Fornaciai, M., Parsons, B., & Bueti, D. (2022). Subjective time is predicted by local and early visual processing. *NeuroImage, 264*, 119707.

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". *Psychological Monographs, 77*, whole number 576.

Wang, J., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible timing by temporal scaling of cortical responses. *Nature Neuroscience, 21*(1), 102–110.

Wearden, J. H. (1991). Human performance on an analogue of an interval bisection task. *The Quarterly Journal of Experimental Psychology Section B, 43*, 59–81. https://doi.org/10.1080/14640749108401259

Wearden, J. H. (1992). Temporal generalization in humans. *Journal of Experimental Psychology. Animal Behavior Processes, 18*, 134–144. https://doi.org/10.1037/0097-7403.18.2.134

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *The Quarterly Journal of Experimental Psychology B, 51*(2), 97–120.

Westheimer, G. (1999). Discrimination of short time intervals by the human observer. *Exp Brain Res Exp Hirnforsch Expérimentation Cérébrale, 129*(1), 121–126.

Wiener, M., Turkeltaub, P., & Coslett, H. B. (2010). The image of time: A voxel-wise meta-analysis. *NeuroImage, 49*(2), 1728–1740.

Wing, A. M. (2002). Voluntary timing and brain function: An information processing approach. *Brain and Cognition, 48*, 7–30. https://doi.org/10.1006/BRCG.2001.1301

Wittmann, M. (2013). The inner sense of time: How the brain creates a representation of duration. *Nature Reviews Neuroscience, 14*(3), 217–223.

Wright, B. A., Buonomano, D. V., Mahncke, H. W., & Merzenich, M. M. (1997). Learning and generalization of auditory temporal-interval discrimination in humans. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 17*(10), 3956–3963.

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of Vision, 7*(10), 2.

Yarrow, K., Johnson, H., Haggard, P., & Rothwell, J. C. (2004). Consistent chronostasis effects across saccade categories imply a subcortical efferent trigger. *Journal of Cognitive Neuroscience, 16*(5), 839–847.

Zarco, W., & Merchant, H. (2009). Neural temporal codes for representation of information in the nervous system. *Cognitive Critique, 1*(1), 1–30.

Zelaznik, H. N., Spencer, R. M. C., & Ivry, R. B. (2002). Dissociation of explicit and implicit timing in repetitive tapping and drawing movements. *Journal of Experimental Psychology. Human Perception and Performance, 28*(3), 575–588.

Zhou, S., Masmanidis, S. C., & Buonomano, D. V. (2022). Encoding time in neural dynamic regimes with distinct computational tradeoffs. *PLoS Computational Biology, 18*(3), e1009271.

Creating a Home for Timing Researchers: Then, Now, and the Future

Argiro Vatakis and Sundeep Teki

Abstract

Our ability to perceive event duration and order is critical in every aspect of our lives, from everyday tasks like coordinating our limbs to walk safely, to uniquely human activities like planning our children's future. Many theoretical accounts of timing have been proposed to explain the mechanisms underlying our ability to estimate time and unify events in time. Continuous progress is being met in further refining and extending current theories, with the aim not only to advance our understanding of timing and time perception, but also to make timing more accessible and applicable to daily life. For this to be possible, cross-disciplinary thinking is required, which is something one cannot easily attain in a scientific conference, rather it requires a community. Having a community with an interest and/or expertise in timing can allow for cross-fertilization of ideas. This chapter introduced the story of the Timing Research Forum or else TRF.

A. Vatakis (✉)
Multisensory and Temporal Processing Laboratory (MultiTimeLab), Department of Psychology, Panteion University of Social and Political Sciences, Athens, Greece

S. Teki
University College London, London, UK

Independent AI consultant, Delhi, India

Keywords

Time perception · Timing · Timing Research Forum · Scientific community

Introduction

Our ability to perceive event duration and order is critical in every aspect of our lives, from everyday tasks like coordinating our limbs to walk safely, to uniquely human activities like planning our children's future. And yet, unlike other perceptual dimensions, we do not possess a sense organ specifically dedicated to time. Rather, our temporal percepts seem to result from a conglomerate of brain processes that allow us to time our behavior to act in a highly complex world (e.g., Allman et al., 2014; Macar & Vidal, 2009; Matthews & Meck, 2016; Tsao et al., 2022). Many theoretical accounts of timing have been proposed to explain the mechanisms underlying our ability to estimate time and unify events in time (e.g., Buhusi & Meck, 2005; Matthews & Meck, 2016; Eagleman & Pariyadath, 2009; New & Scholl, 2009; Tse et al., 2004; Ulrich et al., 2006). Continuous progress is being met in further refining and extending current theories, with the aim not only to advance our understanding of timing and time perception, but also to make timing more accessible and applicable to daily life. For this to be possible, cross-disciplinary think-

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_2

ing is required, which is something one cannot easily attain in a scientific conference, rather it requires a community. Having a community with an interest and/or expertise in timing can allow for cross-fertilization of ideas. Such crosstalk allows not only for multidisciplinary work to be possible but also interdisciplinary excellence.

In this chapter, we will describe the story of the making of a "home" for the timing researchers around the world to allow for the above-mentioned cross/interdisciplinary thinking. We will describe the first steps made initiating a more general timing community, the restart of creating a more focused timing community, and how this community has supported and encouraged the growth and expansion of timing in different disciplines and different settings, but also fostering more basic research and more accessibility in research materials and knowledge.

TIMELY: The First Step in Creating a Timing Community

The idea of creating a multidisciplinary network of scientists working on timing and time perception was entertained in 2007 (during AV's last year of her PhD studies), was formulated as a funding proposal in 2008, and was realized in 2009, when funded by the European COST (European Cooperation in Science and Technology) agency under the title "TD0904: Time In MEntaL activitY: theoretical, behavioral, bioimaging, and clinical perspectives (TIMELY)" (see Fig. 1). TIMELY mainly focused on the following issues.

A. The conceptual analysis and measurement of time given that there was no common code of communication as to the different aspects of time perception. A debatable definition also leads to problematic measurement methodologies. Reaching an agreement regarding the conceptual analysis of time, we expected to also lead to more efficient and accurate measures of timing in humans and other animals.
B. The exploration of factors associated with timing variability given that high timing variability had been reported within and between individuals. Such variability may represent a barrier in understanding time perception. Thus, a close examination of various cognitive/biological processes in relation to timing was considered essential.
C. The expansion of timing research to ecologically valid stimuli and real-world applications. Most timing research had focused on simple stimuli, necessitating the use of more informationally rich stimuli (e.g., music, action) for advancing time perception research and, thus, extending the results to real-world applications.
D. The uncovering of the neural correlates of time perception. Advances in neuroimaging allow observation of the brain in action. It was considered necessary to identify the techniques appropriate for studying timing in humans and other animals and for examining time distortions in specific neurological/psychiatric conditions and other impairments.

These ambitious aims lead to a vigorous community and the development of an active hub for: (a) European and international cooperation: 17 short-term research visits across Europe and internationally; (b) Timing events: 3 large conferences and 20 workshops and symposiums (see Fig. 2); (c) Training schools: 6 training schools with amazing participation from senior and junior scientists (see Fig. 2); and (d) Publications: 14 collaborative works (e.g., Vatakis & Allman, 2015; Vatakis et al., 2011, 2018). A significant result that was also initiated due to TIMELY was the first journal dedicated on timing, *Timing & Time Perception*, a name suggested by Warren Meck and adopted by the editorial team. This journal continues to serve the aim to host all basic research studies, including interdisciplinary and multidisciplinary works on timing and time perception, and serve as a forum for discussion and extension of current knowledge on the topic.

TIMELY lasted for four years and served as a unique "meeting place" of researchers working on time from different disciplines (given their different perspectives and approaches on time as a concept and percept), different laboratories, and

Fig. 1 The logo of the TIMELY community and the attempt to recreate it in our very first training school in Groningen (NL) on "Psychophysical, Computational, and Neuroscience Models of Time Perception" on April 4–8, 2011

different countries. At its conclusion, TIMELY brought together 318 individuals—including 112 females and 122 early-stage researchers—working on timing and laid the foundation for the next stage of this community. It was a community that would be focused on timing and that would continue to be present independent of funding.

Timing Research Forum (TRF): The Current and Future Home of Timing

After the conclusion of TIMELY in 2014, there was a lot of crosstalk among the community of how to move forward. One thing was certain that this next step had to be a group effort. Warren Meck at some point in 2015 had forwarded to AV an article under review that was written by ST and which urged for a community in timing. AV followed up on this, ST was immediately positive about the notion of a community, and it was in 2016 that the Timing Research Forum (TRF) was first introduced (Teki, 2016; see Fig. 3).

TRF's aim is also to bring together researchers that have been studying timing and time perception from many different perspectives and through different methodologies and techniques. We believe that true advancement comes from collaborations and the sharing of the knowledge and experience gained by each person, each laboratory, and each department and institution (Teki, 2016). TRF, therefore, started as a "free existing and evolving" entity that will continuously change depending on the activities organized and the people interested with a focus on the maintenance, continuation, and advancement of the timing and time perception community through open

Fig. 2 Group pictures of almost all attendees of the (**a**) first TIMELY training school in Groningen (NL) on the "Psychophysical, Computational, and Neuroscience Models of Time Perception" held on April 4–8, 2011, (**b**) TIMELY workshop on "Temporal processing in clinical populations" in Thessaloniki (GR) during March 26–29, 2012, and (**c**) final conference of the TIMELY project held on Corfu (GR) on March 31st till April 3rd, 2014

Fig. 3 The logo of the TRF community and the very first group picture on the very first TRF conference held in Strasbourg (FR) on October 23–25, 2017

science, discussions, and collaborations, and the advancement and dissemination of the new knowledge accumulated.

Since 2016, TRF has been serving this vision with several activities initiated by the community. TRF's regular activities include: (a) monthly newsletters that cover job advertisements, new papers published, calls for special journal issues, and, in general, all that pertains to timing, (b) the organization of conferences every 1.5 years (excluding the COVID-19 break; Giersch & Coull, 2018; Merchant, 2019) that include initiatives such as the Young Investigator Keynote award and the MoonShot event (where scientists discuss the most important questions to be solved in the next 5 years for better understanding time in the brain), (c) the monthly Journal Club talks from both junior and senior scientists (see https://www.youtube.com/@timingresearchforum6392/playlists), (d) the TRF Carte Blanche series, where leading researchers in timing present and openly discuss their past and future ideas; this project, which was conceived by Virginie van Wassenhove and realized with the help of Fuat Balci, kicked off in 2022 by hosting the amazing (Charles) Randy Gallistel, and (e) the online presence through TRF's website, social media (Facebook and Twitter), and YouTube channel storing all the online activities of the community.

TRF was also present, with Virginie van Wassenhove leading the effort, in bringing together researchers during the pandemic of COVID-19 in a massive effort to study the "Effects of physical and social distancing on time perception and temporal cognition," which lead to a large collection of data (Chaumon et al., 2022), publications (e.g., Rioux et al., 2022), and a large body of data and code open to the research community (https://osf.io/359qm/). TRF is also the host of the "Timing Database" (Aydoğan et al., 2023), an initiative of compiling a shared database of published and unpublished work on interval timing. Such a database will allow researchers to conduct secondary analysis of large datasets to test basic assumptions (e.g., large scale tests of the scalar property and its limits, order effects, differences between sub- and suprasecond timing), investigate novel research questions, and/or make new discoveries based on much larger datasets than those collected by a single laboratory. Importantly, this database will also contain unpublished datasets and, thus, serve at least as a partial correction factor for publication bias.

The sixth birthday of TRF in 2022 was met with a membership of 1207 members, plans for the post-COVID TRF3 conference, and new ideas to be implemented (e.g., annual training schools). This would not be possible without the support from the community and the continuous work of the Early Career Researchers (ECR) committee that devote their time and effort to keep TRF current, fresh, and alive.

Then, Now, and the Future of Timing and TRF

At the closing of TIMELY, Matthews and Meck (2014) published a paper reviewing the current challenges in the field of timing and their proposed directions for the future of timing. They correctly argued that there were three general challenges (and opportunities, we would say) in timing research: (a) the malleability of timing to context and other factors, (b) the individual differences in both the mechanisms and substrates of interval timing, and (c) the limited transfer of laboratory timing research to the real world (Matthews & Meck, 2014). The timing community has been vigorously working on these challenges with an ever-increasing rate of publications in the area (see Fig. 4). Furthermore, the maturity of the timing research and the presence of the timing community in multiple areas of science and society lead to the initiation, in 2017, of the discussions of the importance and relevance of timing in technology at the European Commission level. Specifically, in 2018, we had the launch of the very first Future Emerging Technologies (FET) Proactive call for proposals seeking "new technological possibilities inspired by notions of time, not seen as a given and singular background

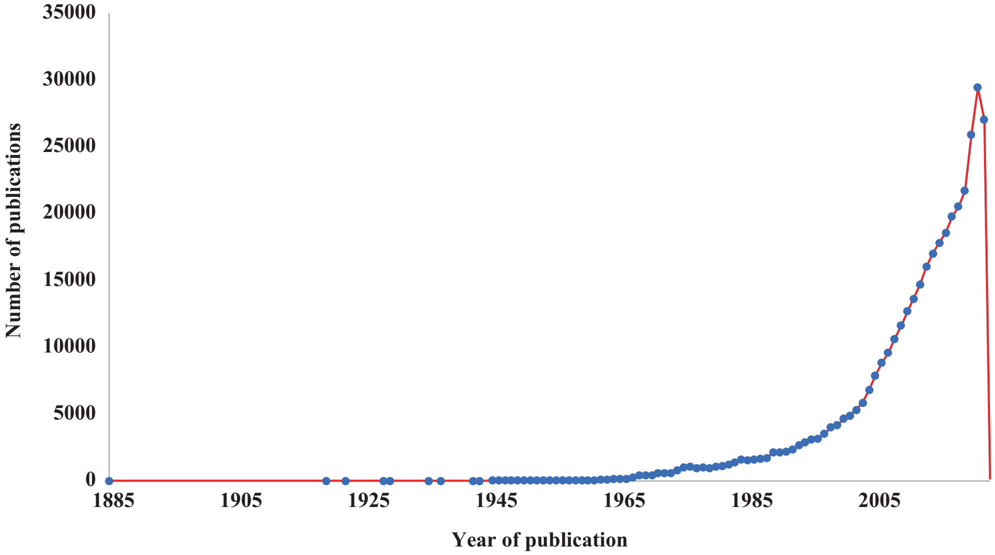


Fig. 4 The frequency of publications—according to PubMed—with the keywords "Time perception" OR "Timing" (as captured on 6/12/2022)

against which things unfold, but rather as a resource that can be experienced and used in different ways" (Digital4Science, 2017).

The efforts of the timing community to make timing more relevant to technology and other more applied fields, as well as the increasing needs of other fields to interact with psychology and neuroscience have flourished in recent years (see Fig. 5a). Timing research is now included or central to funding sources that aim to develop technologies for the future (see Fig. 5b). It must be noted here that one of the first European-funded projects on time was a project developed by members of the TIMELY community. The project was a FET project called TIMESTORM (Mind and Time: Investigation of the Temporal Traits of Human-Machine Convergence) that aimed to equip artificial agents with temporal cognition to establish a new framework for the investigation and integration of knowing, doing, and being in artificial systems. Subsequently, the VIRTUALTIMES FET project aimed to create tools to speed up or slow down events in immersive artificial worlds, resulting in the development of the MetaChron platform that is a time-sensitive and "intelligent" VR platform that enables experiences of virtually modified time and tools for diagnosing and treating neurological pathologies. More recently, another FET project called ChronoPilot—that AV coordinates—aims to capitalize on the knowledge produced in timing research to develop mixed reality tech-

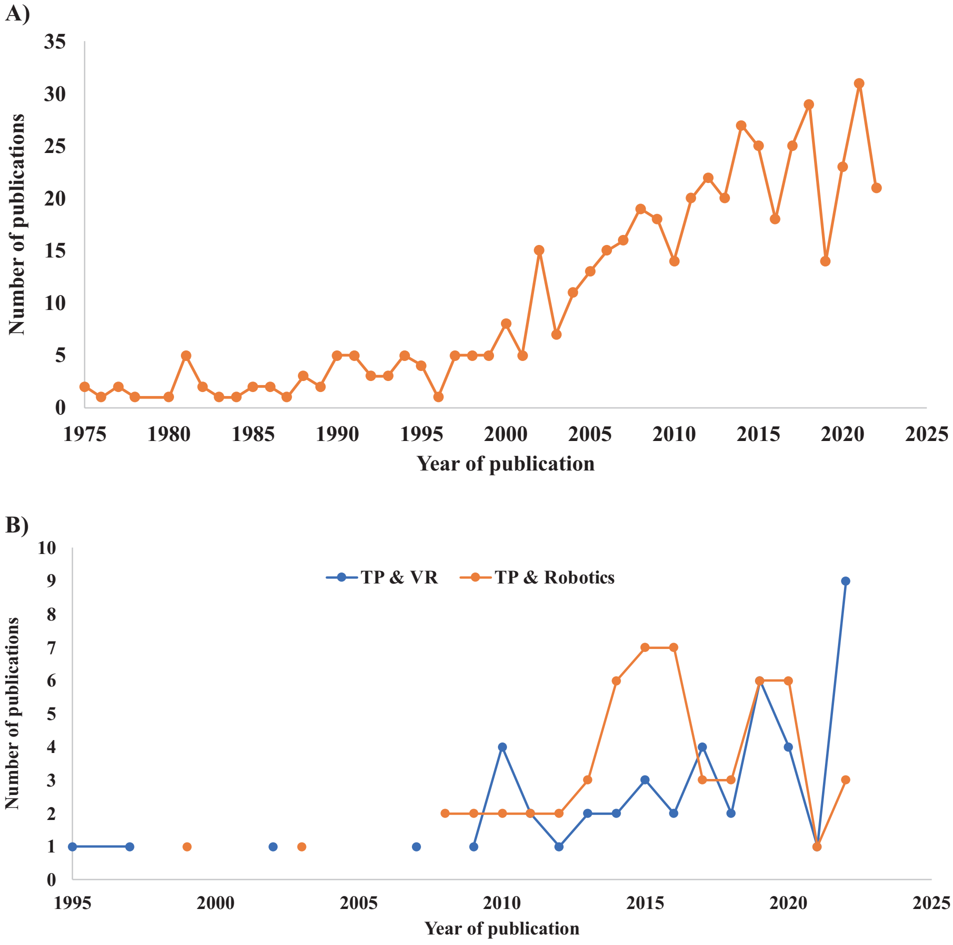


Fig. 5 The frequency of publications—according to PubMed—on (**a**) "time perception" AND "applied" and (**b**) "time perception" AND "Virtual Reality" (VR) or "Robotics" (as captured on 6/12/2022)

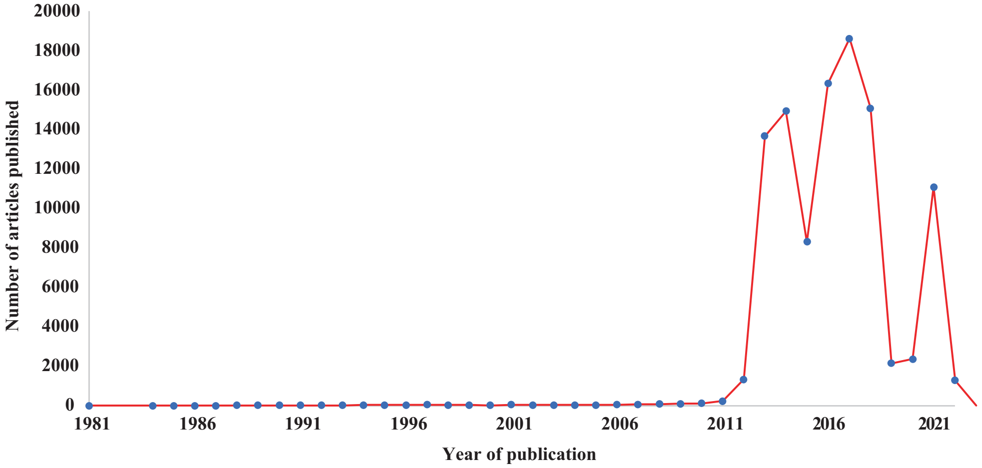


Fig. 6 The frequency of publications with associated datasets—according to PubMed—on "time perception" OR "timing" (as captured on 6/12/2022)

nologies to adapt dynamically one's time perception, as well as to find ways to coordinate time plasticity in collaborative settings with human and/or artificial agents. The aim is to modulate timing for better decision making and well-being, thus aiming for a profound impact on both technology and society.

The above-mentioned activity demonstrates that the timing community is very active in addressing all challenges posed in the field but also open to new ones. Of note here are efforts devoted to openness of the data collected (see Fig. 6) to attempt to answer more general questions on timing that can only be answered by comparing large data. Examples of these efforts were seen in the Chaumon et al.'s (2022) and Aydoğan et al.'s (2023) work, where comparison of larger datasets may allow us to answer questions related to the central tendency in timing, differences in various timing intervals and tasks, individual differences, the role of cognitive process in timing, and how isolation and wellbeing relate to timing, to name just a few. TRF will certainly be here, now, and in the future, to disseminate and support all that is *timing*.

Acknowledgments AV was supported from the European Union's Horizon 2020 research and innovation program under grant agreement No 964464.

References

Allman, M. J., Yin, B., & Meck, W. H. (2014). Time in the psychopathological mind. In D. Lloyd & V. Arstila (Eds.), *Subjective time: The philosophy, psychology, and neuroscience of temporality* (pp. 637–654). MIT Press.

Aydoğan, T., Karşılar, H., Duyan, Y. A., Akdoğan, B., Baccarani, A., De Corte, B., Crystal, J., Çavdaroğlu, B., Gallistel, C. R., Grondin, S., Gür, E., Hallez, Q., de Jong, J., Matell, M., Narayanan, N., Özoğlu, E., van Maanen, L., Öztel, T., Vatakis, A., Freestone, D., & Balcı, F. (2023). The Timing Database: An open-access, live repository for interval timing studies. *Behavioral Research Methods*. https://doi.org/10.3758/s13428-022-02050-9

Buhusi, C., & Meck, W. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience, 6*, 755–765. https://doi.org/10.1038/nrn1764

Chaumon, M., Rioux, P.-A., Herbst, S., Spiousas, I., Kübel, S., Gallego Hiroyasu, E., Leman Runyun, Ş., Micillo, L., Thanopoulos, V., Mendoza-Durán, E., Wagelmans, A., Mudumba, R., Tachmatzidou, R., Cellini, N., D'Argembeau, A., Giersch, A., Grondin, S., Gronfier, C., Alvarez Igarzábal, F., Klarsfeld, A., Jovanovic, L., Laje, R., Lannelongue, E., Mioni, G., Nicolaï, C., Srinivasan, N., Sugiyama, S., Wittmann, M., Yotsumoto, Y., Vatakis, A., Balcı, F., & van Wassenhove, V. (2022). The Blursday database as a resource to study subjective temporalities during COVID-19. *Nature Human Behavior, 6*, 1587–1599.

Digital4Science. (2017). *Time*. Retrieved from https://ec.europa.eu/futurium/en/digital4science/time.html (9/12/2022).

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

Giersch, A., & Coull, J. T. (2018). TRF1: It was the best of time(s)…. *Timing & Time Perception, 6*(3-4), 231–414. https://doi.org/10.1163/22134468-00603001

Macar, F., & Vidal, F. (2009). Timing processes: An outline of behavioural and neural indices not systematically considered in timing models. *Canadian Journal of Experimental Psychology, 63*, 227–239. https://doi.org/10.1037/a0014457

Matthews, W. J., & Meck, W. (2014). Time perception: The bad news and the good. *WIREs Cognitive Science, 5*, 429–446. https://doi.org/10.1002/wcs.1298

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

Merchant, H. (2019). TRF2: Having the time of our lives. *Timing & Time Perception, 7*(4), 291–452. https://doi.org/10.1163/22134468-00704001

New, J. J., & Scholl, B. J. (2009). Subjective time dilation: Spatially local, object-based, or a global visual experience? *Journal of Vision, 4*, 1–11.

Rioux, P., Chaumon, M., Demers, A., Fitzback-Fortin, H., Kübel, S. L., Lebrun, C., Mendoza-Duran, E., Micillo, L., Racine, C., Thibault, N., van Wassenhove, V., & Grondin, S. (2022). Psychological time during the COVID-19 lockdown: Canadian data. *Timing & Time Perception, 10*(4), 326–343. https://doi.org/10.1163/22134468-bja10063

Teki, S. (2016). A citation-based analysis and review of significant papers on timing and time perception. *Frontiers in Neuroscience, 10*, 330. https://doi.org/10.3389/fnins.2016.00330

Tsao, A., Yousefzadeh, S. A., Meck, W. H., Moser, M. B., & Moser, E. I. (2022). The neural bases for timing of durations. *Nature Reviews Neuroscience, 23*, 646–665. https://doi.org/10.1038/s41583-022-00623-3

Tse, P., Intriligator, J., Rivest, J., & Cavanagh, P. (2004). Attention and the subjective expansion of time. *Perception & Psychophysics, 66*, 1171–1189.

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

Vatakis, A., & Allman, M. (2015). *Time distortions in mind: Temporal processing in clinical populations*. Brill.

Vatakis, A., Esposito, A., Giagkou, M., Cummins, F., & Papadelis, G. (2011). *Multidisciplinary aspects of time and time perception*. Springer LNCS/LNAI Proceedings Volume.

Vatakis, A., Balci, F., Di Luca, M., & Correa, A. (2018). *Timing and time perception: Procedures, measures, and applications*. Brill.

The Processing of Short Time Intervals: Some Critical Issues

Simon Grondin

Abstract

Humans have the capability to make judgments about the relative duration of time intervals with accuracy (correct perceived duration) and precision (low variability). However, this capability has limitations, some of which are discussed in the present chapter. These limitations, either in terms of accuracy or precision, are obvious when there are changes in the physical characteristics of the stimuli used to mark the intervals to be judged. The characteristics are the structure (filled vs. empty) of the intervals and the sensory origin of the stimuli used to mark them. The variability of time estimates also depends on the use of single intervals by opposition to the use of sequences of intervals, and on the duration range under investigation. In addition to the effect caused by the physical characteristics of the stimuli, the perceived duration also relies on the way of presenting successive stimuli and on whether the intervals are marked by a single source or by different sources with distance (spatial effect) between them.

Keywords

Time psychophysics · Duration discrimination · Variability · Weber's law · Perceived duration

Introduction

In order to understand how a mechanism works, it is important to understand its properties and the extent of the influence that some factors may exert on it. When it comes to the judgment of time intervals, we do not know exactly which mechanism is responsible, although there are many classical interpretations (e.g., Gibbon et al., 1984; Jones, 2019; Mauk & Buonomano, 2004; Zakay & Block, 1997), as well as more recent ones (e.g., Gilden & Mezaraups, 2022; Killeen & Grondin, 2022). This chapter aims to identify the impact of some factors on the processing of temporal information.

To what extent is it reasonable to posit that a unique timekeeping mechanism is responsible for processing temporal intervals (1) from different duration ranges, (2) marked by different sensory signals, (3) presented or not in a continuous sequence, and (4) when adopting or not an explicit counting strategy? These avenues are of interest in this chapter as they have an impact on either the variability of time judgments or on the perceived duration.

S. Grondin (✉)
École de psychologie, Université Laval,
Québec, QC, Canada
e-mail: simon.grondin@psy.ulaval.ca

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_3

Methods for Studying Time Perception

The reading of this chapter will be easier if some critical distinctions are kept in mind. It is only a small portion of the field of psychological time that is addressed here. What is addressed is neither the judgments about the order of arrival of different sensory signals, nor the judgments about the durations of past events or activities. What is of interest here are judgments about time in conditions where participants know in advance that a judgment about time will have to be made (what is called prospective timing [see Grondin & Laflamme, 2015; Tobin et al., 2010]), and the intervals to be processed are usually relatively short.

Also critical is the following distinction. If you have a device to keep track of time, you want this device to be accurate on average (that the mean judgments reflect the exact physical/chronometric time), and you want this average to be obtained with precision, i.e., with low variability (Grondin, 1998; Killeen & Grondin, 2022). This chapter is about the factors affecting the variability of time judgments and the mean perceived duration. These two dependent variables can take different forms, depending on the method adopted for investigating the processing of short time intervals.

There are several classical methods for testing timing and time perception capabilities (Grondin, 2008).[1] One is the verbal estimates of a given interval, using chronometric units (seconds, or minutes for longer intervals). In another method, participants could be asked to produce an interval of a given length, with this production typically being made with two successive finger taps, or by pressing a button for the duration of the targeted length.[2] Interval reproduction is also a classical method. Instead of using chronometric units to define a targeted interval, the experimenter marks the interval with a continuous sensory signal, or with two brief signals marking an empty duration. After this presentation, the participant reproduces this interval. This can be done with two successive finger taps or by pressing a button for the duration of the targeted length, as described for interval production, or by ending with a single finger tap the reproduced interval whose onset was initiated with a sensory signal provided by the experimenter. Note that the exact method adopted for reproducing an interval will have an incidence on the results (Mioni et al., 2014). In the case of interval production and reproduction, it is possible to have an index of perceived duration and of variability. When a series of intervals are reproduced, the mean is kept as the estimate of perceived duration, and the standard deviation is the estimate of variability.

Another type of classical method is often referred to as the comparison method. In this case, two (or more) intervals are presented and must be discriminated ("Which one is longer?" or "Was the second interval shorter or longer than first one?"). The method of constant stimuli involves such a comparison, and most often, on each trial, a standard is compared to one of a series of comparison intervals. This method requires building a psychometric function where the probability of responding that the comparison interval is longer than the standard (on the y axis) is plotted as a function of the value of the comparison intervals (on the x axis). The point on the x axis corresponding to 50% of long responses is the point of subjective equality (PSE) and could

---

[1] In addition to the classical psychophysical methods, it is possible to use indirect methods such as drawing or segmenting a line presented visually on a screen (Bisson et al., 2009). For instance, instead of using a binary-forced-choice system as is the case with a bisection task, a participant could be asked, after being presented a short and a long standard interval, to indicate on a line, between Point S (short) and Point L (Long) on this line, a point representing the length of the comparison interval which has just been presented (Gamache & Grondin, 2010—Experiment 2). With such a line-segmentation task, if the duration of the presented interval appears to be, for example, exactly between the short and the long standard, the participant would have to click on the middle of the line.

[2] Interval production may also mean that, after the presentation of an interval or a series of intervals of equal duration, a series of equal intervals are produced by a participant (usually with a series of successive finger taps), and these productions may or may not involve a synchronization phase (Bartolo et al., 2014; Grondin, 1992; Wing & Kristofferson, 1973).

be interpreted as an index of perceived duration: a PSE value shorter that the standard interval means that the comparison interval is more often perceived as longer than the standard. The slope of the function provides an index of variability. One can use one standard deviation to account for the variability of judgments (Killeen & Weiss, 1987). It is also possible to take the distance between the points on the *x* axis corresponding to 75% and 25% of long responses. This distance, divided by 2, corresponds to what is known in psychophysics as the difference threshold (Grondin, 2016).

There are variants of this constant method, based on a categorization of intervals. With the many-to-few method, only a distribution of intervals is used; over the trials, participants develop an implicit standard and after each presentation of an interval, they determine whether this interval is shorter or longer than the implicit standard (Allan, 1979). Another variation of the constant method—a very popular one indeed—is called, following the terminology used in animal timing, the bisection method. With this method, two standard intervals (i.e., the shortest and longest of a distribution of intervals) are presented several times at the beginning of the experimentation. After these standard intervals are presented, other intervals of varying durations are introduced, and participants must determine, after each presentation of an interval, whether it is closer in duration to the short or to the long standard interval. With the variations of the constant method, it is also possible to build a psychometric function and extract information about perceived duration (PSE) and variability.

There are several other ways to estimate difference threshold when comparing time intervals. For example, one is the use of an adaptive procedure, where the level of difficulty is adjusted after each trial. Though not often used in the study of time perception, the method of adjustment can also be used: after the successive presentation of a standard and a comparison interval, participants have to adjust the duration of the comparison until the length of both intervals are perceived as equal (Kuroda & Hasuo, 2014). The standard deviation of the values of comparison intervals kept after each trial provides the estimate of variability.

Finally, discrimination levels, or variability, are sometimes expressed with other performance indexes like the percentage of correct responses, signal detection theory measures such as A′ or d′, or the Weber fraction, i.e., the difference threshold divided by the standard (variability to time ratio) (Grondin, 2010). Perceived duration could be accounted for by a percentage of long responses, the PSE (when psychometric functions are used), the constant error (CE: PSE—Standard), or the relative CE (PSE—Standard in absolute values; or (PSE—Standard/Standard)).

Variability of Time Estimates

There are several factors that may influence the variability of time estimates. Among these factors, three are described below: the structure of the interval and the sensory modality used for marking it, the mode of interval presentation (single intervals vs. sequences), and the range of duration under investigation.

Structure and Sensory Modality

The way of indicating the beginning and end of an interval to be discriminated has a major impact on performance. Essentially, an interval could be filled, i.e., marked by a single, continuous sensory signal, or empty, i.e., marked by two brief sensory signals with an empty duration between the signal. For the discrimination of 50-ms intervals, there is evidence that discrimination is easier in the filled than in the empty condition (Rammsayer & Lima, 1991; Rammsayer & Skrandies, 1998). For longer intervals (but still <1 s), there is evidence that discrimination is easier with empty than with filled intervals (Grondin, 1993; Grondin et al., 1998),[3] but this finding holds only if there is no confusion in the

---

[3] In these studies, this superiority of empty vs. filled intervals was observed in auditory and visual conditions.

presentation of successive empty intervals (if a single-stimulus method is adopted, instead of a forced-choice procedure). Also, Rammsayer and Leutner (1996) report clear evidence that the discrimination of empty intervals is impaired when marking signals are longer than 200 ms (for a summary of the filled vs. empty comparisons, see Grondin, 2003).

In the case of empty intervals, the length of the signals marking the beginning and end of the timing period has an impact, but the sensory modality used as a marker is also critical. There is abundant literature on the processing of temporal information revealing the superiority of the auditory modality over the visual (Bratzke & Ulrich, 2019; Bueti et al., 2008; Espinoza-Monroy & de Lafuente, 2021; Gontier et al., 2013; Kuroda et al., 2014; Rousseau et al., 1983; Ulrich et al., 2006) or tactile modality (Azari et al., 2020; Azari et al., 2023; Espinoza-Monroy & de Lafuente, 2021; Grondin & Rousseau, 1991; Jones et al., 2009; Mayer et al., 2014; Villalonga et al., 2021). The differences between performances when the visual and the tactile conditions are compared are usually nonsignificant, and this conclusion holds with a staircase method (Jones et al., 2009), and a bisection method (Azari et al., 2020, 2023).

In most investigations on duration discrimination involving empty intervals, signals marking the beginning and end of the intervals are identical. If these signals differ, the performance could be impaired. If sounds marking intervals are of different frequencies, discrimination is lower (Grondin & Rousseau, 1991), and this deterioration increases with more differences between signals (van Noorden, 1975). Even more damaging are conditions where the signals marking an empty interval are delivered from different sensory modalities (intermodal intervals) instead of from the same modality (intramodal intervals). This finding holds for the different combinations of auditory and visual markers (Azari et al., 2020, 2023; Rousseau et al., 1983; Rousseau & Kristofferson, 1973) and for combinations involving a tactile stimulus (Azari et al., 2020, 2023; Grondin & Rousseau, 1991; Mayer et al., 2014).

This distribution of performance levels when different interval structures or different sensory signals are used to mark intervals is very challenging for researchers in the field of time perception. A classical account of the capability to process temporal information is based on the idea that there is a unique, central, internal clock. This clock is reported to be a pacemaker-counter device, where the pacemaker emits pulses which are then accumulated in the counter (Gibbon et al., 1984; Zakay & Block, 1997). This accumulation process thus forms the basis of time judgments: the more pulses accumulated, the longer the perceived duration. The properties of the pulses' emission by the pacemaker that could account for the errors observed in duration discrimination tasks. These errors are also partly accounted for by the involvement of attention in the process, either at the beginning or end of the accumulation process (switch process Meck, 1984), or during the accumulation (gate process Zakay & Block, 1997). In the global information processing account of time perception and estimation, errors are also attributed to the memory and decisional processes involved in a specific method/task (Gibbon et al., 1984). Within this perspective, the superiority of the auditory modality could be explained by an increased firing rate of the pacemaker when sounds are used to mark intervals, with an increased rate leading to a finer temporal resolution. Such an explanation based on the firing rate would also be useful to explain why filled intervals are easier to discriminate than empty intervals, assuming that the firing rate is increased with filled intervals (Rammsayer & Lima, 1991). Note that some authors rather propose the contribution of a two-level process, with one level being controlled by a supraordinate, modality-independent processing system, and the other level relying on a modality-specific (auditory or visual) processing of temporal information (Stauffer et al., 2012). When faced with the difficulty to account for the problems caused by intermodal conditions, Grondin and Rousseau (1991) proposed that if the marking of an interval involved only one sensory modality, the processing would be taken on by a modality-specific system, but when an interval is intermodal, the processing could be based on the functioning of the central (amodal) clock process.

Neuroscience offers a different viewpoint for explaining modality effects. Based on investigations involving the use of transcranial magnetic stimulations (TMS), Bueti et al. (2008) came to the conclusion that temporal information may be independently represented within a given sensory modality. Also based on the effect of TMS, Kanai et al. (2011) rather emphasize the critical role of the auditory cortex. Applying TMS on the auditory cortex impairs the processing of both auditory and visual temporal intervals. However, applying TMS on the visual cortex affects the processing of visually-marked intervals but not that of intervals marked with sounds. This finding led Kanai et al. to the conclusion that the auditory cortex is crucial for processing time intervals. Within such a perspective, weaker duration discrimination performances with intervals marked by visual, tactile, or intermodal sequences could be accounted for by the need to transfer any non-auditory signals (visual or tactile signals) into an auditory code.

Single Intervals vs. Sequences of Intervals

Another factor to take into account when assessing duration discrimination levels, or variability of time estimates, is whether the assessment is based on single intervals or sequences of intervals. Sequences of intervals are particularly relevant for researchers interested in the mechanisms involved in rhythm perception.

Typically, in duration discrimination tasks, the comparison interval(s) is(are) presented after the standard interval(s) and an interstimulus interval. A classical finding is that multiple presentations of the standard intervals result in an increased discrimination level (Grondin, 2012; Ivry & Hazeltine, 1995; McAuley & Kidd, 1998). This improvement of performance with multiple standard intervals was accounted for by a multiple-look model (Drake & Botte, 1993). In this model, a good part of the variance in the discrimination process is located at the memory level. The representation of the standard interval, with which is compared the comparison intervals, is improved with the repetition of standard intervals. However, at some point, after four presentations of an interval, there is no improvement to expect from additional presentations (ten Hoopen et al., 2011).

In several studies, the number of standard and comparison intervals covaries. When the number of repetitions of standard and comparison intervals is controlled, performance can be improved by the sole repetition of the comparison intervals, presented after the standard and without the repetition of the standard (Miller & McAuley, 2005). This finding of an increased discrimination level following a second sequence of comparison intervals is inconsistent with the multiple-look model; multiple presentations do not reduce the variability only by reducing the memory representation of the standard. Indeed, more presentations of both the standard and the comparison intervals contribute to improve discrimination capabilities (McAuley & Miller, 2007).

The influence of the number of comparison intervals is applicable to conditions involving visual stimuli, and even when comparing an auditory sequence to a visual sequence. In Grondin and McAuley (2009), using four comparison intervals instead of only one improved performance when both standard and comparison sequences involved only visual stimuli, or when one sequence was visual and the other auditory (cross-modal conditions). Interestingly, performances were better in the cross-modal condition than in the only visual condition (but lower than in the only auditory condition).[4] These results hold in conditions where the standard intervals are presented first (fixed value in a block of trials), and comparison intervals second (varied values within a block of trials), and when using the opposite order (varied first and fixed second).

---

[4] With single intervals, Ulrich et al. (2006) report better discrimination when both the standard and comparison intervals are of the same modality (A or V) than if one is A and the other is V, and the V-A than in an A-V sequence, and intervals are not marked by the same modality, discrimination is better if an A rather than a visual signal is the comparison interval.

Also, this study shows that presenting the fixed sequence first leads to better performances than presenting the varied sequence first. This finding is consistent with what is observed with the presentation of single intervals (Lapid et al., 2008; Ulrich et al., 2006), a phenomenon referred to as a Type B effect (Dyjas et al., 2014; Ellinghaus et al., 2021).

There is another type of methodological control that may influence discrimination levels. The comparison interval(s) could be presented in continuity after the standard interval(s) or after an interstimulus interval (discontinuous condition). This distinction is especially critical when sequences of flashes are used to mark very brief intervals. With a continuous sequence, the discrimination threshold will be as high in a 150-ms condition as in a 300-ms condition, while with a discontinuous sequence, the Weber fraction holds for intervals lasting 150–900 ms (Grondin, 2001a). This finding indicates the importance of having sufficient time for the resolution of the signal marking intervals in a duration discrimination task.

Globally, the superiority of the use of multiple presentations of intervals for discriminating time could reflect a limitation of the pacemaker-counter interpretation, which is more suitable for accounting for the presentation of single intervals. An interpretation such as the *dynamic attending theory* of time perception (Jones, 2019; Jones & Boltz, 1989; Large & Jones, 1999), where time judgments are based on the attending of a given event after the regular occurrences of such events, seems more appropriate. As soon as a series of signals marking intervals are repeated with some regularity, it is possible to develop expectations about the moment of occurrence of the forthcoming signals.

The neuroscience literature dedicated to timing reveals that this single intervals vs. sequences of intervals distinction is crucial. Based on the use of fMRI, Teki et al. (2011) identified the potential contribution of two cerebral mechanisms, where the availability of one mechanism seems to depend on the temporal structure of the events in the environment. One mechanism is dedicated to the processing of absolute duration; it is based on a representation of an interval. This mechanism is an olivo-cerebellar network; it includes the contribution of the inferior olivary nucleus, the vermis, and the dentate nucleus. The other mechanism is based on the expectation of the moment of arrival of an event occurring in the continuity of regular events. This mechanism is a striato-thalamo-cortical network; it involves the contribution of several brain structures: the putamen, caudate nucleus, premotor cortex, supplementary motor area (SMA), dorsolateral prefrontal cortex, and thalamus (see also Cadena-Valencia et al., 2018).

Duration Range[5]

The capacity to discriminate intervals is not the same for all durations on the chronometric time continuum. One way of addressing this issue is to rely on Weber's law. According to this law, the minimum difference (difference threshold) between the stimuli magnitudes necessary for a discrimination of these stimuli should increase proportionally with the increase in magnitude (Grondin, 2001b; Grondin, 2016; Killeen & Weiss, 1987). In other words, the difference threshold to magnitude ratio should be constant. In the field of time perception, this constancy is often referred to as the scalar property.

The Weber fraction for time perception is, in fact, not constant for a large duration range; it seems to remain constant only for a brief range of durations, and this range depends on the method adopted for assessing the discrimination capabilities. There seems to be some constancy of this fraction for intervals lasting approximately 0.3–1.2 s, at least when intervals are marked by auditory signals (Friberg & Sundberg, 1995).[6] For very brief duration values, the Weber fraction

---

[5] A complete chapter was dedicated to this issue in the first volume edited by Merchant and de Lafuente (Grondin, 2014).

[6] Some studies have investigations on a limited range but show the constancy of the Weber fraction from 400 to 800 ms with sounds (Schultze, 1978), and between 500 and 740 ms with sounds or flashes marking intervals (Grondin et al., 2001).

is much higher (Friberg & Sundberg, 1995; Getty, 1975; Grondin, 1993). Such an increase of the Weber fraction for low magnitudes of chronometric time is likely due to the properties of the signals marking time, not to the properties of the underlying mechanism responsible for processing time. In fact, this increase can be accounted for mathematically quite easily, using a generalized form of Weber's law (Killeen & Weiss, 1987), which might indicate that a common timing system could be used for processing intervals shorter than 1.2 s. Nevertheless, for some researchers, this increased Weber fraction for very brief intervals (200–300 ms) is an important breakpoint and is interpreted as an indication that two different mechanisms are at play (Hibi, 1983; Michon, 1964).

There is another form of violation of Weber's law that indicates that the duration range has a critical impact on the capability to discriminate intervals. While the timing literature dedicated to the discrimination of brief intervals (Getty, 1975; Grondin, 1993; Grondin et al., 2001), and even to the production of time intervals (Madison, 2001) shows some constancy of the Weber fraction for intervals lasting 0.3–1.2 s, there are several indications that, beyond 1.2 s, the Weber fraction is increased. Clearly, there is an increase of the Weber fraction from 1.3 to 1.6 s with interval discrimination and categorization tasks with auditory markers, and with an interval reproduction method (Grondin, 2012).[7] This finding is consistent with the composite figure reported by Gibbon et al. (1997)[8] showing a likely transition point for intervals briefer or longer than 1.5 s. Even with visual markers, the Weber fraction is reported to be higher at 1.2 s than at 0.9 s (Grondin, 2001a).

The breakpoint around 200–300 ms remains disputable given the generalized version of Weber's law; however, a breakpoint at approximately 1.2–1.5 s is more difficult to interpret. It may well reflect a fundamental limitation in the flow of psychological time. This limitation in the capability to process a temporal extent probably speaks to a temporal limitation of human cognitive processes. Beyond what might be representative of some "psychological present" (Gruber et al., 2022), efficiency is lost.

While the results obtained with several classical timing tasks point in the direction of this temporal limitation, it can also be observed using a different method. With this method (Grondin et al., 2015a), at the beginning of each trial, a brief sound is presented four times in order to mark three equal time intervals lasting 800, 1200, or 1600 ms. Participants have to count from 1 up to a target number, using the pace dictated by the sounds, and to press the spacebar (of a computer keyboard) twice, that is, when saying 1 and when reaching the target number. The target number, *x*, was 16, 21, and 31 for trials involving, respectively, 1600-, 1200-, and 800-ms intervals, *y*. In other words, the theoretical duration to be reproduced was 24 s, i.e., *y* (*x*-1): 1600 (16-1), 1200 (21-1), and 800 (31-1). After 30 trials in each condition, the variability of the reproduced intervals, divided by the target (24 s), was the dependent variable of interest (the Weber fraction). The results show that there is not much difference between the 0.8- and 1.2-s conditions, but the Weber fraction is much higher at 1.6 s. This finding indicates once again, with a very basic task (just counting at a given pace), a temporal limitation, that is, a loss of efficiency when intervals are too long. This result was also obtained with stuttering participants in a study showing that there are benefits to expect from counting aloud instead of silently (Plamondon & Grondin, 2020).

When it comes to counting, adopting such a strategy has an impact on the range effect. The exact value of the Weber fraction in classical human timing tasks heavily depends on the adoption or not of a strategy consisting of the segmentation (with explicit counting for instance) of the time intervals to be processed (Grondin et al.,

---

[7] See also Halpern and Darwin (1982) for an increased Weber fraction between 1150 and 1300 ms in an auditory rhythm discrimination task.

[8] This figure includes data from animal and human timing. For a clear violation or the Weber fraction with long intervals (categorization and production tasks), see Bizo et al. (2006).

2004).[9] When participants have to discriminate intervals lasting 0.7–1.9 s and marked with sounds, there are benefits to expect from a counting strategy when intervals are longer than approximately 1.2 s (Grondin et al., 1999); and interestingly, with explicit counting, variability can remain quite constant from 1.3 to 1.9 s. Other data also indicate that singing or counting can help to maintain the Weber fraction's constancy in the reproduction of 6- to 24-s intervals and, if participants are musicians, the fraction may even be lower with longer intervals (Grondin & Killeen, 2009).

Recent findings in neurosciences also support this distinction between intervals below and above 1.2 s, showing the presence of distinct brain dynamics and networks to support the perception of time intervals above and below 1.2 s (Thibault et al., 2023). In this study involving a totally passive task, participants were exposed to a series of sounds occurring in some trials every 0.8 s or, in other trials, every 1.6 s. In addition to these standard intervals, some deviant interstimulus intervals were presented. This is known as an oddball paradigm, and EEG recordings (auditory event-related potentials) were taken during the presentation of standards and deviants. Whole-brain cluster-based permutation statistics performed around the peak of N1 and P2 components, contrasting deviants' event-related potentiels (ERPs) against their standard, revealed that the deviance detection in the 0.8- and 1.6-s conditions differed. Temporal-related deviance detection occurred during the N1 time period for the condition using intervals above 1.2 s, and during the P2 time period for the condition using intervals below 1.2 s.

Perceived Duration of Intervals

Several factors exert influence on the perceived duration of a given temporal extent. Three of these factors are described here: the structure of the interval and sensory modality used for marking it, the mode of interval presentation, and whether or not there is space between the signals marking the intervals.

Structure and Sensory Modality

One important factor that influences perceived duration is the structure of a time interval. As mentioned earlier, an interval could be filled (marked by a continuous signal) or empty (the period between two brief signals). It is most often reported that filled intervals are perceived longer as compared with empty intervals of the same length, and in several studies, the difference is very large (Craig, 1973; Goldfarb & Goldstone, 1963; Wearden et al., 2007). Nevertheless, some reports indicate that this filled vs. empty difference is prone to methodological effects and individual differences, these limitations being reported for intervals in the range of 20–180 ms (Hasuo et al., 2011a) and for intervals up to 520 ms (Hasuo et al., 2014a).

This filled vs. empty structure issue, sometimes referred to as the sustained sound illusion (Repp & Marcus, 2010), could be seen as a special case of what is more generally referred to as the filled duration illusion where an empty time interval (marked by two brief signals) is filled with one or more brief stimuli inserted within this interval (ten Hoopen et al., 2008; Thomas & Brown, 1974).[10]

Different hypotheses can be developed to account for the difference in perceived duration between filled (continuous signal) and empty intervals. Somewhat along the line described earlier about the pacemaker-counter device, the difference may depend, according to Wearden et al. (2007), on the pace of pulse emissions by the pacemaker being faster with the presence of a continuous signal than with no signal. Another explanation rather emphasizes the possibility that the end of the timekeeping period arrives earlier

---

[9] In Grondin (1993) (Experiment 5), participants were not refrained from counting, and the Weber fraction remained roughly constant up to 4 s (the longest interval tested), a finding applying to both auditory and visual conditions (and to filled and empty intervals).

[10] This is reminiscent of the Oppel-Kundt illusion in visual perception, where a line segment divided into several parts is perceived as longer than an undivided line segment of the same length (Grondin, 2016).

when a signal must be detected, as is the case for an empty interval, than it takes for perceiving the offset of a signal, as is required with a filled interval (Craig, 1973) (see the internal-marker hypothesis (Grondin, 2003; Grondin et al., 2018)).

It is important to note that several physical characteristics can influence the perceived duration of an interval (Berglund et al., 1969). The intensity of signals seems to influence the perceived duration of a filled interval (louder is longer), but, in the case of audition, the key factor is apparently not the absolute intensity of signals but rather the relative intensity of signal and background sounds (Matthews et al., 2011). For empty auditory intervals, lengthening the duration of either the first or the second marker (intervals of 500 ms), when the task is to judge from the offset of the first signal to the onset of the second, increases the perceived duration of these intervals (Woodrow, 1928). In an experiment involving empty intervals from 120 to 360 ms that had to be estimated from the onset of the first signal to the onset of the second, Hasuo et al. (2012) found that perceived duration is longer when the second marker lasts 100 instead of 20 ms. Hasuo et al. (2011b) also found that in a sequence of three successive sounds marking two intervals, increasing the length of the second and third sounds increases the perceived duration of the first and second intervals, respectively. Finally, increasing the length of markers also increases the perceived duration of brief empty intermodal intervals (Grondin et al., 1996; Kuroda et al., 2014).

The sensory modality used to mark the intervals also has an impact on their perceived duration (for a review of older work, see Grondin, 2003). In a series of experiments using temporal generalization and verbal estimation methods, Wearden et al. (1998) showed that intervals marked by sounds are perceived as longer than intervals marked by a visual stimulus (investigation with intervals lasting 77–1183 ms). Jones et al. (2009) also showed with a verbal estimation method and the same duration range, that intervals marked by sounds are perceived as longer than intervals marked by a visual stimulus, but they also showed that there was no difference between intervals marked by visual and tactile stimuli. In both studies by Wearden et al. and Jones et al., the difference between the sensory conditions increased as the durations of the intervals to be estimated increased. The modality effect was attributed to the properties of the pacemaker, positing that it runs faster when sounds are used to mark intervals.

For empty time intervals marked with two signals delivered from different modalities (intermodal intervals), Grondin and Rousseau (1991) report that intervals (around 250 ms) tend to seem longer when the first marker is tactile, and shorter when the first marker is visual. In a study by Azari et al. (2020), in experiments involving 300- and 900-ms standard durations, an interval marked by an auditory-tactile sequence is perceived as longer than an interval marked by a visual-tactile sequence. This finding could be attributed to the alerting properties of the auditory and visual signals. The detection of a signal is faster in the auditory than in the visual modality. Therefore, the timekeeping activity begins earlier when the first marker is auditory instead of visual, and consequently, the interval is perceived as longer if the first marker is a sound. In a study by Azari et al. (2023), an interval marked by a tactile-visual sequence is perceived as longer than an interval marked by a tactile-auditory sequence. This finding could also be accounted for by the faster detection with an auditory than with a visual stimulus; in the present case, it is the rapidity of the offset process, i.e., the moment that determines the end of the timekeeping period, that is at play.

This section emphasized the physical characteristics of intervals that can impact perceived duration. It is relevant to note that the psychological characteristics of the signals used to mark an interval also influence perceived duration. For instance, if an interval is marked with a visual image, the perceived duration of this image will be influenced, for instance, by the symbolic meaning of the image (Gagnon et al., 2018; Mioni et al., 2015; Ouellet et al., 2023), by the emotion expressed if a visage is presented (Grondin et al., 2015b), or by the emotion generated by the image (Grondin et al., 2014). The psychological characteristics of sounds also influence perceived duration (Mioni et al., 2018; Noulhiane et al., 2007; Voyer & Reuangrith, 2015).

Mode of Interval Presentation

In the field of visual perception, there are several spatial and temporal contexts that will determine what exactly is perceived. There are different contrast effects leading to the evidence that visual perception goes beyond simply receiving sensory information on the retinae. The same is true for the processing of the temporal information.

The way of presenting successive stimuli influences perceived duration. In a comparison task involving a forced-choice procedure, depending on the length of intervals under investigation, either the first or the second presented interval may tend to be perceived as longer (Hellström, 1977). This tendency is referred to as the Time Order Error (TOE) (Allan, 1979). When, with a given interval, there is no TOE, this interval is referred to as the indifference interval (Eisler et al., 2008). Several factors, like practice or using feedback, attenuate TOE (Allan, 1979; Jamieson & Petrusic, 1976; Jamieson & Petrusic, 1978).

When a sequence of intervals marked by consecutive tones is used, instead of two distinct intervals with an interstimulus interval, the perceived duration of the intervals in the sequence can be severely distorted. In the auditory modality, there is a classical illusion called time-shrinking (Nakajima et al., 1992). This illusion occurs when three successive sounds mark two neighboring time intervals, T1 and T2. When T1 $\leq$ 200 ms and is shorter than T2, T2 will appear much briefer than when it is presented in isolation. The maximum shrinking effect occurs when T1 − T2 = 80 ms (ten Hoopen et al., 2008). Time-shrinking is also observed when visual (Arao et al., 2000) and tactile (Hasuo et al., 2014b; Van Erp & Spapé, 2008) stimuli are used.

Other phenomena related to the presentation of successive intervals may occur. For instance, when a sequence of four identical flashes lasting 600 or 667 ms are presented, the duration of the first flash is overestimated by about 50% (Rose & Summers, 1995). For the processing of filled auditory intervals, when a sine tone is preceded by a noise, it is perceived as much longer than if it is presented in isolation, a phenomenon called the time-stretching illusion (Sasaki et al., 1992). Note that there is no such illusion when a tone was followed, instead of being preceded, by a noise (Kuroda & Grondin, 2013a).

Spatial Effects

There is a huge literature on the effect of space on temporal judgments, but many pieces of this literature belongs to research avenues that are different from the main concerns of the present chapter. These research avenues touch on, for example, the field of time-to-collision or time-to-contact, or the Piagetian relationship between space, time, and speed.

In most studies on time perception related to duration discrimination, signals marking intervals are delivered from a single sensory source. Nevertheless, there is a series of studies where signals marking time intervals are delivered from different locations in space. These studies indicate that having space between signals influences the perceived duration of a time interval. Indeed, there are two distinct questions: how is perceived duration changed if two signals marking an interval are delivered from different locations instead of from the same location? And when signals are delivered from different locations, does more space result in longer perceived duration?

With Visual Signals

The most classical demonstrations of the effect of space on perceived duration were made with visual signals. A typical case involves the succession of three consecutive flashes and is referred to as the kappa effect (Abe, 1935; Cohen et al., 1953).[11] Let us call these flashes A, B, and C. If

---

[11] There are also classical works on the effect of time on perceived distance, the tau effect, where longer duration between equidistant signals results in longer perceived distance.

the duration between these flashes is the same, but the distance between the source of flashes B and C is longer than the distance between the source of flashes A and B, the duration of the time interval marked by flashes B and C will be perceived as longer than the duration of the time interval marked by flashes A and B.

One explanation of this effect is based on the hypothesis that what is perceived by an observer is not three distinct objects, but one object flashing three times (imputed-velocity model: ten Hoopen et al., 2008; Jones & Huang, 1982), and that this object moves at a constant speed. But this constancy is not maintained if the time intervals are equal and the spatial distance is not. To maintain speed constancy, the longer distance is rather interpreted as a longer time interval.

The influence of space on time judgments has also been shown when only two flashes were used (imputed velocity is less likely in such a condition). In an experiment using the reproduction of 7- to 11-s intervals marked by flashes from a same source, either A or B, or by A followed by B (32 inches between A and B), Price-Williams (1954) reports that having space between signals, instead of having flashes from a same source, results in longer reproductions. In an experiment on the discrimination of short empty time intervals (around 250 ms) marked by brief signals delivered from sources A and B located in the left vs. right visual field, intervals marked by AA or BB sequences were perceived as shorter than intervals delivered by bilateral sequences, AB or BA (Grondin, 1998). In this experiment, intervals presented in a BA sequence (right to left) were perceived as shorter than intervals presented in an AB sequence (left to right). Also, when the visual sources are located on a vertical plane, the difference between perceived durations in the upper vs. lower visual field depends on the fixation point (Roussel et al., 2009).

With Auditory Signals

In the auditory modality, the idea of a kappa effect has taken different forms. In some studies, it is not the spatial distance between sound sources that is manipulated, but the different frequencies of successive sounds. For example, Shigeno (1986) reported that with three consecutive sounds (total duration of 1000 ms), if the frequency of the middle sound is closer to that of the last sound than to that of the first sound, the time interval between the first two sounds is likely to be perceived as longer than the time interval between the last two sounds. The reader will also find a demonstration of a kind of kappa effect with successive auditory signals of different frequencies in Henry and McAuley (2009) and with three successive sounds of different intensities in Alards-Tomalin et al. (2013).

There are not many demonstrations of the effect of spatial sound sources on perceived duration (Sarrazin et al., 2007). In Grondin and Plourde (2007), a sequence of four consecutive brief sounds was presented, delimiting three time intervals. The first three sounds were delivered from the same spatial location, marking the first and second time intervals. These standard time intervals lasted 75, 150, or 225 ms. The moment of occurrence of a fourth sound was systematically varied and was delivered from a different location on a vertical plane. Participants had to compare the duration of the third interval (between the third and fourth sound) with the preceding intervals. In the 150- and 225-ms standard conditions, there was an overestimation of the duration of the last time interval (an occurrence of the kappa effect). This effect occurred only when participants were uncertain of the provenance of the sounds.

The effect of space on perceived duration can be completely different when single intervals, instead of consecutive intervals, are presented. In a study by Roy et al. (2011), two sounds marked intervals (standard equals 125 or 250 ms) and the sources of these sounds were 1.1 or 3.3 m apart. Time intervals were perceived as longer with less distance between sound sources: longer spatial distance resulted in shorter perceived duration.

A recent study also showed the presence of a kappa effect in the auditory modality with a reproduction task (500- to 1600-ms intervals), and with a discrimination task (500-ms standard) involving three consecutive sounds (Bausenhart, 2018). In the reproduction task (500- to 1600-ms intervals), the kappa effect occurred with sounds presented from different sources (vs. from the same source), and this effect held both when the auditory and visual signals were delivered simultaneously, and when the auditory and visual signals were presented simultaneously with only sounds being delivered from different locations.

With Tactile Signals

There is not much work on the effect of space on tactile duration discrimination. In one study, electrotactile stimuli were used to mark intervals (standard = 500 ms). Two brief consecutive signals were delivered on the same hand (left or right: no space between markers) or on different hands: space between signals). The results indicate that when there is space between signals, duration is perceived as longer than when there is no space (Grondin et al., 2011). In this experiment, there was no difference in perceived duration between intervals marked on the left vs. right hand, or marked in the left-to-right vs. the right-to-left directions. Note also that discrimination was better in the same hand than in the different hand condition. Finally, marking a time interval with a tactile signal on different fingers of the same hand (with space between signals) is also perceived as longer duration than an interval of the same length marked by two tactile signals on the same finger (no space) (Kuroda & Grondin, 2013b).

Conclusion

It is difficult to say whether judging the relative duration of time intervals with accuracy (correct perceived duration) and precision (low variability) requires a central, or unique, mechanism able to adapt to all environmental circumstances or all methodological contexts, or whether a multitude of timing mechanisms are available, with the physical constraints imposed by the environment (or methodological choices in experiments) deciding the best-suited mechanism the brain has to offer in a given circumstance. The series of limitations described in the present chapter should be seen as challenges for any model or theory developed to account for the capability to perceive and estimate time intervals.

A theory like the *dynamic attending theory* of time perception (Jones, 2019; Large & Jones, 1999; Teki et al., 2011) seems most suitable for environments offering sequences with regularities, while the pacemaker-counter is more appropriate for processing single intervals, but requires the contribution of several additional sources of errors (attentional, memory, and decisional processes (Gibbon et al., 1984; Zakay & Block, 1997)) to account for the various distortions and discrimination levels described in the chapter. Another theory, referred to as the trace theory of time perception (Killeen & Grondin, 2022), proposes that duration estimation relies on the memorial strength of the signal used to initiate the interval at the moment this interval ends, with the signal that terminates the interval providing a reminder of the initial intensity. This theory was developed to account for some of the limiting factors reported above, but offers no speculations about the brain mechanisms underlying this type of processing.

References

Abe, S. (1935). Experimental study on the co-relation between time and space. *Tohoku Psychologica Folia, 3*, 53–68.

Alards-Tomalin, D., Leboe-McGowan, L. C., & Mondor, T. A. (2013). Examining auditory kappa effects through manipulating intensity differences between sequential tones. *Psychological Research, 77*, 480–491.

Allan, L. G. (1979). The perception of time. *Perception & Psychophysics, 26*(5), 340–354.

Arao, H., Suetomi, D., & Nakajima, Y. (2000). Does time-shrinking take place in visual temporal patterns? *Perception, 29*, 819–830.

Azari, L., Mioni, G., Rousseau, R., & Grondin, S. (2020). An analysis of the processing of intramodal and intermodal time intervals. *Attention, Perception & Psychophysics, 82*, 1473–1487.

Azari, L., Drouin, J., Plante, G., & Grondin, S. (2023). Discrimination of brief empty time intervals when the first marker is tactile. *Timing and Time Perception, 11*, 343–361.

Bartolo, R., Prado, L., & Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *The Journal of Neuroscience, 34*, 3910–3923.

Bausenhart, K. M. (2018). Quinn KR on the interplay of visuospatial and audiotemporal dominance: Evidence from a multimodal kappa effect. *Attention, Perception & Psychophysics, 80*, 535–552.

Berglund, B., Berglund, U., Ekman, G., & Frankehaeuser, M. (1969). The influence of auditory stimulus intensity on apparent duration. *Scandinavian Journal of Psychology, 10*, 21–26.

Bisson, N., Tobin, S., & Grondin, S. (2009). Remembering the duration of joyful and sad musical excerpts. *NeuroQuantology, 7*, 46–57.

Bizo, L. A., Chu, J. Y. M., Sanabria, F., & Killeen, P. R. (2006). The failure of Weber's law in time perception and production. *Behavioural Processes, 71*(2), 201–210.

Bratzke, D., & Ulrich, R. (2019). Temporal sequence discrimination within and across senses: Do we really hear what we see? *Experimental Brain Research, 237*, 3089–3098.

Bueti, D., Bahrami, B., & Walsh, V. (2008). Sensory and associative cortex in time perception. *Journal of Cognitive Neuroscience, 2008*(20), 1054–1062.

Cadena-Valencia, J., García-Garibay, O., Merchant, H., Jazayeri, M., & de Lafuente, V. (2018). Entrainment and maintenance of an internal metronome in supplementary motor area. *eLife, 7*, e38983. https://doi.org/10.7554/eLife.38983

Cohen, J., Hansel, C. E. M., & Sylvester, J. D. (1953). A new phenomenon in time judgment. *Nature, 172*, 901.

Craig, J. C. (1973). A constant error in the perception of brief temporal intervals. *Perception & Psychophysics, 13*, 99–104.

Drake, C., & Botte, M. C. (1993). Tempo sensitivity in auditory sequences: Evidence for a multiple-look model. *Perception & Psychophysics, 54*(3), 277–286.

Dyjas, O., Bausenhart, K. M., & Ulrich, R. (2014). Effects of stimulus order on duration discrimination sensitivity are under attentional control. *Journal of Experimental Psychology: Human Perception and Performance, 40*(1), 292–307.

Eisler, H., Eisler, A., & Hellström, A. (2008). Psychophysical issues in the study of time perception. In S. Grondin (Ed.), *Psychology of time* (pp. 75–109). Emerald.

Ellinghaus, R., Giel, S., Ulrich, R., & Bausenhart, K. M. (2021). Humans integrate duration information across sensory modalities: Evidence for an amodal internal reference of time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 47*(8), 1205–1225.

Espinoza-Monroy, M., & de Lafuente, V. (2021). Discrimination of regular and irregular rhythms explained by a time difference accumulation model. *Neuroscience, 459*, 16–26.

Friberg, A., & Sundberg, J. (1995). Time discrimination in a monotonic, isochronic sequence. *The Journal of the Acoustical Society of America, 98*, 2524–2531.

Gagnon, C., Bégin, C., Laflamme, V., & Grondin, S. (2018). Temporal processing of joyful and disgusting food pictures by women with an eating disorder. *Frontiers in Human Neuroscience, 12*, 29.

Gamache, P.-L., & Grondin, S. (2010). The life span of time intervals in reference memory. *Perception, 39*, 1431–1451.

Getty, D. (1975). Discrimination of short temporal intervals: A comparison of two models. *Perception & Psychophysics, 18*(1), 1–8.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. G. Allan (Eds.), *Timing and time perception* (Annals of the New York Academy of Sciences) (Vol. 423, pp. 52–77). New York Academy of Sciences.

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology, 7*(2), 170–184.

Gilden, D., & Mezaraups, T. M. (2022). Allometric scaling laws for temporal proximity in perceptual organization. *Psychological Review, 129*, 457–483.

Goldfarb, J. L., & Goldstone, S. (1963). Time judgment: A comparison of filled and unfilled durations. *Perceptual and Motor Skills, 16*(2), 376.

Gontier, E., Hasuo, E., Mitsudo, T., & Grondin, S. (2013). EEG investigations of duration discrimination: The intermodal effect is induced by an attentional bias. *PLoS One, 8*(8), e74073.

Grondin, S. (1992). Production of time intervals from segmented and nonsegmented inputs. *Perception & Psychophysics, 52*(3), 345–350.

Grondin, S. (1993). Duration discrimination of empty and filled intervals marked by auditory and visual signals. *Perception & Psychophysics, 54*(3), 383–394.

Grondin, S. (1998). Judgments of the duration of visually marked empty time intervals: Linking perceived duration and sensitivity. *Perception & Psychophysics, 60*, 19–330.

Grondin, S. (2001a). Discriminating time intervals presented in sequences marked by visual signals. *Perception & Psychophysics, 63*, 1214–1228.

Grondin, S. (2001b). From physical time to the first and second moments of psychological time. *Psychological Bulletin, 127*, 22–44.

Grondin, S. (2003). Sensory modalities and temporal processing. In H. Helfrich (Ed.), *Time and mind II: Information processing perspectives* (pp. 75–92). Hogrefe & Huber.

Grondin, S. (2008). Methods for studying psychological time. In S. Grondin (Ed.), *Psychology of time* (pp. 51–74). Emerald Group Publishing.

Grondin, S. (2010). Timing and time perception: A review of recent behavioral and neuroscience findings and theoretical directions. *Attention, Perception, & Psychophysics, 72*, 561–582.

Grondin, S. (2012). Violation of the scalar property for time perception between 1 and 2 seconds: Evidence from interval discrimination, reproduction, and categorization. *Journal of Experimental Psychology: Human Perception and Performance, 38*, 880–890.

Grondin, S. (2014). About the (non)scalar property for time perception. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of interval timing*. Springer Editorial System.

Grondin, S. (2016). *Psychology of perception*. Springer.

Grondin, S., & Killeen, P. R. (2009). Tracking time with song and count: Different Weber functions for musicians and non-musicians. *Attention, Perception & Psychophysics, 71*, 1649–1654.

Grondin, S., & Laflamme, V. (2015). Stevens's law for time: A direct comparison of prospective and retrospective judgments. *Attention, Percept & Psychophys, 77*, 1044–1051.

Grondin, S., & McAuley, J. D. (2009). Duration discrimination in crossmodal sequences. *Perception, 38*(10), 1542–1559.

Grondin, S., & Plourde, M. (2007). Discrimination of time intervals presented in sequences: Spatial effects with multiple auditory sources. *Human Movement Science, 26*, 702–716.

Grondin, S., & Rousseau, R. (1991). Judging the relative duration of multimodal short empty time intervals. *Perception & Psychophysics, 49*(3), 245–256.

Grondin, S., Ivry, R. B., Franz, E., Perreault, L., & Metthé, L. (1996). Markers' influence on the duration discrimination of intermodal intervals. *Perception & Psychophysics, 58*, 424–433.

Grondin, S., Meilleur-Wells, G., Ouellette, C., & Macar, F. (1998). Sensory effects on judgments of short-time intervals. *Psychological Research, 61*, 261–268.

Grondin, S., Meilleur-Wells, G., & Lachance, R. (1999). When to start explicit counting in time-intervals discrimination task: A critical point in the timing process of humans. *Journal of Experimental Psychology: Human Perception and Performance, 25*(4), 993–1004.

Grondin, S., Ouellet, B., & Roussel, M.-E. (2001). About optimal timing and stability of weber fraction for duration discrimination. *Acoustical Science and Technology, 22*, 370–372.

Grondin, S., Ouellet, B., & Roussel, M.-È. (2004). Benefits and limits of explicit counting for discriminating temporal intervals. *Canadian Journal of Experimental Psychology, 58*, 1–12.

Grondin, S., Kuroda, T., & Mitsudo, T. (2011). Spatial effects on tactile duration categorization. *Canadian Journal of Experimental Psychology, 65*, 163–167.

Grondin, S., Laflamme, V., & Gontier, E. (2014). Effect on perceived duration and sensitivity to time when observing disgusted faces and disgusting mutilation pictures. *Attention, Perception & Psychophysics, 76*, 1522–1534.

Grondin, S., Laflamme, V., & Mioni, G. (2015a). Do not count too slowly: Evidence for a temporal limitation in short-term memory. *Psychonomic Bulletin & Review, 22*, 863–868.

Grondin, S., Laflamme, V., Bienvenue, P., Labonté, K., & Roy, M.-L. (2015b). Sex effect in the temporal perception of faces expressing anger and shame. *International Journal of Comparative Psychology, 28*, 27200.

Grondin, S., Hasuo, E., Kuroda, T., & Nakajima, Y. (2018). Auditory time perception. In R. Bader (Ed.), *Springer handbook of systematic musicology* (pp. 423–440). Springer.

Gruber, R. P., Block, R. A., & Montemayor, C. (2022). Physical time within human time. *Frontiers in Psychology, 13*. https://doi.org/10.3389/fpsyg.2022.718505

Halpern, A. R., & Darwin, C. J. (1982). Duration discrimination in a series of rhythmic events. *Perception & Psychophysics, 31*(1), 86–89.

Hasuo, E., Nakajima, Y., & Ueda, K. (2011a). Does filled duration illusion occur for very short time intervals? *Acoustical Science and Technology, 32*, 82–85.

Hasuo, E., Nakajima, Y., & Hirose, Y. (2011b). Effects of sound-marker durations on rhythm perception. *Perception, 40*, 220–242.

Hasuo, E., Nakajima, Y., Osawa, S., & Fujishima, H. (2012). Effects of temporal shapes of sound markers on the perception of interonset time intervals. *Attention, Perception & Psychophysics, 74*, 430–445.

Hasuo, E., Nakajima, Y., Tomimatsu, E., Grondin, S., & Ueda, K. (2014a). The occurrence of the filled duration illusion: A comparison of the method of adjustment with the method of magnitude estimation. *Acta Psychologica, 147*, 111–121.

Hasuo, E., Kuroda, T., & Grondin, S. (2014b). About the time-shrinking illusion in the tactile modality. *Acta Psychologica, 147*, 122–126.

Hellström, Å. (1977). Time errors are perceptual. *Psychological Research, 39*, 345–388.

Henry, M. J., & McAuley, J. D. (2009). Evaluation of an imputed pitch velocity model of the auditory kappa effect. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 551–564.

Hibi, S. (1983). Rhythm perception in repetitive sound sequence. *Journal of the Acoustical Society of Japan (E), 4*, 83–95.

Ivry, R. B., & Hazeltine, R. E. (1995). Perception and production of temporal intervals across a range of durations: Evidence for a common timing mechanism. *Journal of Experimental Psychology: Human Perception and Performance, 21*(1), 3–18.

Jamieson, D. G., & Petrusic, W. M. (1976). On a bias induced by the provision of feedback in psychophysical experiments. *Acta Psychologica, 4*, 199–206.

Jamieson, D. G., & Petrusic, W. M. (1978). Feedback versus an illusion in time. *Perception, 7*(1), 91–96.

Jones, M. R. (2019). *Time will tell: A theory of dynamic attending*. Oxford University Press.

Jones, M. R., & Boltz, M. (1989). Dynamic attending and responses to time. *Psychological Review, 96*, 459–491.

Jones, B., & Huang, Y. L. (1982). Space-time dependencies in psychophysical judgment of extent and duration: Algebraic models of the tau and kappa effect. *Psychological Bulletin, 91*, 128–142.

Jones, L. A., Poliakoff, E., & Wells, J. (2009). Good vibrations: Human interval timing in the vibrotactile modality. *Quarterly Journal of Experimental Psychology, 62*, 2171–2186.

Kanai, R., Lloyd, H., Bueti, D., & Walsh, V. (2011). Modality-independent role of the primary auditory cortex in time estimation. *Experimental Brain Research, 209*, 465–471.

Killeen, P. R., & Grondin, S. (2022). A trace theory of time perception. *Psychological Review, 129*, 603–639.

Killeen, P. R., & Weiss, N. A. (1987). Optimal timing and the Weber function. *Psychological Review, 94*(4), 455–468.

Kuroda, T., & Grondin, S. (2013a). No time-stretching illusion when a tone is followed by a noise. *Attention, Perception, & Psychophysics, 75*, 1811–1816.

Kuroda, T., & Grondin, S. (2013b). Discrimination is not impaired when more cortical space between two electro-tactile markers increases perceived duration. *Experimental Brain Research, 224*, 303–312.

Kuroda, T., & Hasuo, E. (2014). The very first step to start psychophysical experiments. *Acoustical Science and Technology, 35*, 1–9.

Kuroda, T., Hasuo, E., Labonté, K., Laflamme, V., & Grondin, S. (2014). Discrimination of two neighboring intra- and intermodal empty time intervals marked by three successive stimuli. *Acta Psychologica, 149*, 134–141.

Lapid, E., Ulrich, R., & Rammsayer, T. (2008). On estimating the difference limen in duration discrimination tasks: A comparison of the 2AFC and the reminder task. *Perception & Psychophysics, 70*, 291–305.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How we track time varying events. *Psychological Review, 106*, 119–159.

Madison, G. (2001). Variability in isochronous tapping: Higher order dependencies as a function of intertap interval. *Journal of Experimental Psychology: Human Perception and Performance, 27*(2), 411–421.

Matthews, W. J., Stewart, N., & Wearden, J. H. (2011). Stimulus intensity and the perception of duration. *Journal of Experimental Psychology: Human Perception and Performance, 37*, 303–313.

Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. *Annual Review of Neuroscience, 27*, 307–340.

Mayer, K. M., Di Luca, M., & Ernst, M. O. (2014). Duration perception in crossmodally-defined intervals. *Acta Psychologica, 147*, 2–9.

McAuley, J. D., & Kidd, G. R. (1998). Effect of deviations from temporal expectations on tempo discrimination of isochronous tone sequences. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1786–1800.

McAuley, J. D., & Miller, N. S. (2007). Picking up the pace: Effects of global temporal context on sensitivity to the tempo of auditory sequences. *Perception & Psychophysics, 69*, 709–718.

Meck, W. H. (1984). Attentional bias between modalities: Effects on the internal clock, memory, and decision stages used by animal time discrimination. In J. Gibbon & L. G. Allan (Eds.), *Timing and time perception* (Annals of the New York Academy of Sciences) (Vol. 423, pp. 528–541). New York Academy of Sciences.

Michon, J. A. (1964). Studies on subjective duration. I. Differential sensitivity in the perception of repeated temporal intervals. *Acta Psychologica, 22*, 441–450.

Miller, N. S., & McAuley, J. D. (2005). Tempo sensitivity in isochronous tone sequences: The multiple-look model revisited. *Perception & Psychophysics, 67*, 1150–1160.

Mioni, G., Stablum, F., McClintock, S. M., & Grondin, S. (2014). Different methods for reproducing time, different results. *Attention, Perception, & Psychophysics, 76*, 675–681.

Mioni, G., Zakay, D., & Grondin, S. (2015). Faster is briefer: The symbolic meaning of speed influences time perception. *Psychonomic Bulletin & Review, 22*, 285–1291.

Mioni, G., Laflamme, V., Grassi, M., & Grondin, S. (2018). The effect of emotional spoken words on time perception depends on the gender of the speaker. *Timing and Time Perception, 6*, 1–13.

Nakajima, Y., ten Hoopen, G., Hilkhuysen, G., & Sasaki, T. (1992). Time-shrinking: A discontinuity in the perception of auditory temporal patterns. *Perception & Psychophysics, 51*, 504–507.

Noulhiane, M., Mella, N., Samson, S., Ragot, R., & Pouthas, V. (2007). How emotional auditory stimuli modulate time perception. *Emotion, 7*(4), 697–704.

Ouellet, C., Tétreault, É., & Grondin, S. (2023). Politically biased time perception and perspective. *Timing & Time Perception, 10*, 386–402.

Plamondon, M., & Grondin, S. (2020). Counting fast or slow, aloud or silently? *Cognitive Processing, 21*, 461–467.

Price-Williams, D. (1954). The Kappa effect. *Nature, 173*, 363–364.

Rammsayer, T. H., & Leutner, D. (1996). Temporal discrimination as a function of marker duration. *Perception & Psychophysics, 58*(8), 1213–1223.

Rammsayer, T. H., & Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: Cognitive and perceptual factors. *Perception & Psychophysics, 50*(6), 565–574.

Rammsayer, T. H., & Skrandies, W. (1998). Stimulus characteristics and temporal information processing: Psychophysical and electrophysiological data. *Journal of Psychophysiology, 12*(1), 1–12.

Repp, B. H., & Marcus, R. J. (2010). No sustained sound illusion in rhythmic sequences. *Music Perception, 28*, 121–133.

Rose, D., & Summers, J. (1995). Duration illusions in a train of visual stimuli. *Perception, 24*(10), 1177–1187.

Rousseau, R., & Kristofferson, A. B. (1973). The discrimination of bimodal temporal gaps. *Bulletin of the Psychonomic Society, 1*, 115–116.

Rousseau, R., Poirier, J., & Lemyre, L. (1983). Duration discrimination of empty time intervals marked by intermodal pulses. *Perception & Psychophysics, 34*, 541–548.

Roussel, M.-E., Grondin, S., & Killeen, P. R. (2009). Spatial effects on temporal categorization. *Perception, 38*, 748–762.

Roy, M., Kuroda, T., & Grondin, S. (2011). Effect of space on auditory temporal processing with a single-stimulus method. In P. Strumillo (Ed.), *Advances in sound localization* (pp. 95–104). InTech.

Sarrazin, J. C., Giraudo, M. D., & Pittenger, J. B. (2007). Tau and Kappa effects in physical space: The case of audition. *Psychological Research, 71*, 201–218.

Sasaki, T., Suetomi, D., Nakajima, Y., & ten Hoopen, G. (1992). Time-shrinking, its propagation, and Gestalt principles. *Perception & Psychophysics, 64*, 919–931.

Schultze, H. H. (1978). The detectability of local and global displacements in regular rhythmic patterns. *Psychological Research, 40*, 173–181.

Shigeno, S. (1986). The auditory tau and kappa effects for speech and nonspeech stimuli. *Perception & Psychophysics, 40*, 9–19.

Stauffer, C. C., Haldemann, J., Troche, S. J., & Rammsayer, T. H. (2012). Auditory and visual temporal sensitivity: Evidence for a hierarchical structure of modality-specific and modality-independent levels of temporal information processing. *Psychological Research, 76*, 20–31.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience, 31*, 3805–3812.

ten Hoopen, G., Miyauchi, R., & Nakajima, Y. (2008). Time-based illusions in the auditory mode. In S. Grondin (Ed.), *Psychology of time* (pp. 139–187). Emerald Group Publishing.

ten Hoopen, G., van den Berg, S., Memelink, J., Bocanegra, B., & Boon, R. (2011). Multiple-look effects on temporal discrimination within sound sequences. *Attention, Perception & Psychophysics, 73*, 2249–2269.

Thibault, N., Albouy, P., & Grondin, S. (2023). Distinct brain dynamics and networks for processing short and long auditory time intervals. *Scientific Reports, 13*, 22018. https://doi.org/10.1038/s41598-023-49562-8

Thomas, E. A. C., & Brown, I. (1974). Time perception and the filled-duration illusion. *Perception & Psychophysics, 16*, 449–458.

Tobin, S., Bisson, N., & Grondin, S. (2010). An ecological approach to prospective and retrospective timing of long durations: A study involving gamers. *PLoS One, 5*(2), e9271.

Ulrich, R., Nitschke, J., & Rammsayer, T. (2006). Crossmodal temporal discrimination: Assessing the predictions of a general pacemaker-counter model. *Perception & Psychophysics, 68*, 1140–1152.

Van Erp, J. B. F., & Spapé, M. M. A. (2008). Time-shrinking and the design of tactons. In M. Ferre (Ed.), *Haptics: Perception, devices and scenarios* (Vol. 5024, pp. 289–294). Springer.

van Noorden, L. P. A. S. (1975, February). *Temporal coherence in the perception of tone sequences*. Doctoral Thesis, Eindhoven University of Technology.

Villalonga, M. B., Sussman, R. F., & Sekuler, R. (2021). Perceptual timing precision with vibrotactile, auditory, and multisensory stimuli. *Attention, Perception & Psychophysics, 83*, 2267–2280.

Voyer, D., & Reuangrith, E. (2015). Perceptual asymmetries in a time estimation task with emotional sounds. *Laterality: Asymmetries of Body, Brain and Cognition, 20*, 211–231.

Wearden, J. H., Edwards, H., Fakhri, M., & Percival, A. (1998). Why "sounds are judged longer than lights": Application of a model of the internal clock in humans. *The Quarterly Journal of Experimental Psychology, 51B*, 97–120.

Wearden, J. H., Norton, R., Martin, S., & Montford-Bebb, O. (2007). Internal clock processes and the filled-duration illusion. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 716–729.

Wing, A. M., & Kristofferson, A. B. (1973). Response delay and the timing of discrete motor responses. *Perception & Psychophysics, 14*, 5–12.

Woodrow, H. (1928). Behavior with respect to short temporal stimulus forms. *Journal of Experimental Psychology, 11*, 167–193.

Zakay, D., & Block, R. (1997). Temporal cognit. *Current Directions in Psychological Science, 6*, 12–16.

Neurocomputational Models of Interval Timing: Seeing the Forest for the Trees

Fuat Balcı and Patrick Simen

Abstract

Extracting temporal regularities and relations from experience/observation is critical for organisms' adaptiveness (communication, foraging, predation, prediction) in their ecological niches. Therefore, it is not surprising that the internal clock that enables the perception of seconds-to-minutes-long intervals (interval timing) is evolutionarily well-preserved across many species of animals. This comparative claim is primarily supported by the fact that the timing behavior of many vertebrates exhibits common statistical signatures (e.g., on-average accuracy, scalar variability, positive skew). These ubiquitous statistical features of timing behaviors serve as empirical benchmarks for modelers in their efforts to unravel the processing dynamics of the internal clock (namely answering how internal clock "ticks"). In this chapter, we introduce prominent (neuro)computational approaches to modeling interval timing at a level that can be understood by general audience. These models include Treisman's pacemaker accumulator model, the information processing variant of scalar expectancy theory, the striatal beat frequency model, behavioral expectancy theory, the learning to time model, the time-adaptive opponent Poisson drift-diffusion model, time cell models, and neural trajectory models. Crucially, we discuss these models within an overarching conceptual framework that categorizes different models as threshold vs. clock-adaptive models and as dedicated clock/ramping vs. emergent time/population code models.

Keywords

Computational Models · Drift-Diffusion · Interval Timing · Pacemaker Accumulator · Population Clock Models

Introduction

The simplest definition of interval timing is the perception of event times in the seconds to minutes range. Our sense of time can be encoded retrospectively in episodic memory after a single occurrence, and it can prospectively guide anticipation following the repeated experience of temporally predictable events. For instance, we automatically encode how long we have been sitting in the waiting room (retrospectively) and

---

F. Balcı (✉)
Department of Biological Sciences, University of Manitoba, Winnipeg, MB, Canada
e-mail: fuat.balci@umanitoba.ca

P. Simen
Department of Neuroscience, Oberlin College, Oberlin, OH, USA
e-mail: psimen@oberlin.edu

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_4

learn how long the red light lasts at a specific intersection en route from home to work (prospectively). It is this capacity, coupled with earlier experiences or prior knowledge, that enables us to question whether we mixed up the appointment time or whether the traffic light is broken.

Such temporal experiences often become part of our daily behaviors in the same contexts and daily conversations during which we give approximate estimates of time intervals (Balcı et al., 2023). Humans can reach such duration estimates without a sensory system dedicated to physical time, which is something we cannot feel, see, hear, taste, or smell. Other animals also exhibit the capacity to monitor, remember, and organize adaptive behaviors around biologically critical events. For example, rats can learn how much time has elapsed before they experience an electric shock in a test chamber after a single shock (Bevins & Ayres, 1995) or learn how long after a stimulus onset a reward becomes available (Roberts, 1981). This was observed by Ivan Pavlov as part of classical conditioning (Pavlov, 1927), which he referred to as “inhibition of delay,” as well as by Burrhus Frederic Skinner in operant conditioning in continuous fixed intervals schedules (i.e., the classic “scallop”-shaped cumulative response plot in the fixed interval task, as opposed to a straight line of cumulative responses in variable interval schedules—Ferster & Skinner, 1957). The pioneering work of John Gibbon and his colleagues (e.g., Russell Church, Warren H Meck, Charles Randy Gallistel, Peter Balsam, and others) has further reinforced psychophysics as a benchmark in modeling the interval timing behaviors of humans and nonhuman animals. This work has focused on the nature of mapping between objective and subjective time including its noise characteristics (as the main focus of the psychophysical approach to explaining Weber’s law - Fig. 1). Based on several seminal studies, it is now widely accepted that both humans and other animals exhibit time-tracking capacity with similar statistical signatures, which suggests common, evolutionarily conserved timing mechanisms that are ideal for study in comparative psychology and behavioural neuroscience.

For instance, these researchers have concluded that organisms, on average, reach accurate estimates of time intervals. In other words, when an organism emits anticipatory responses over many trials, the highest rate of responding happens at the time of the learned interval. Another statistical characteristic of timing behavior across different species is that the uncertainty in estimates of time intervals is proportional to the duration of events, such that the discriminability of two event times depends on the ratio and not the absolute difference between the corresponding event times (i.e., Weber’s Law, or in Gibbon et al.’s application to timing data, “time-scale invariance,” or “scalar invariance” in his terminology). In other words, according to Weber’s Law, the discriminability of 2 s vs. 4 s is the same as 10 s vs. 20 s (same logarithmic distance or ratio) and based on the same law 2 s vs. 4 s is more discriminable than 10 s vs. 12 s (same linear distance). In mathematics, scale invariance refers to the invariance of functions (e.g., probability density functions) when the scale over which they are expressed is multiplied by a common factor. In simpler terms, this feature refers to self-similarity, where a given mathematical function looks the same irrespective of the distance from which one looks at it (e.g., as in fractals). Time-scale invariance is the application of the scale invariance property to the time domain, in which the standard deviation of time estimates grows linearly with the duration being timed. As a result, when timed response curves for different target intervals are expressed in relative time (when absolute time is divided by the target interval, analogous to zooming in on the fractals), they superimpose. Figure 1 illustrates this property by expressing the intervals on absolute (left panel) and relative time (right panel).

An important statistical property is implied by timescale invariance. The *coefficient of variation* (CV)—defined as the standard deviation of the time estimates divided by the mean estimate—is constant for timescale-invariant distributions. For distributions such as the Gaussian, inverse Gaussian, and gamma distributions, specifying the CV specifies the distribution completely, so it is most convenient in these cases to work directly with the CV, as we do below.

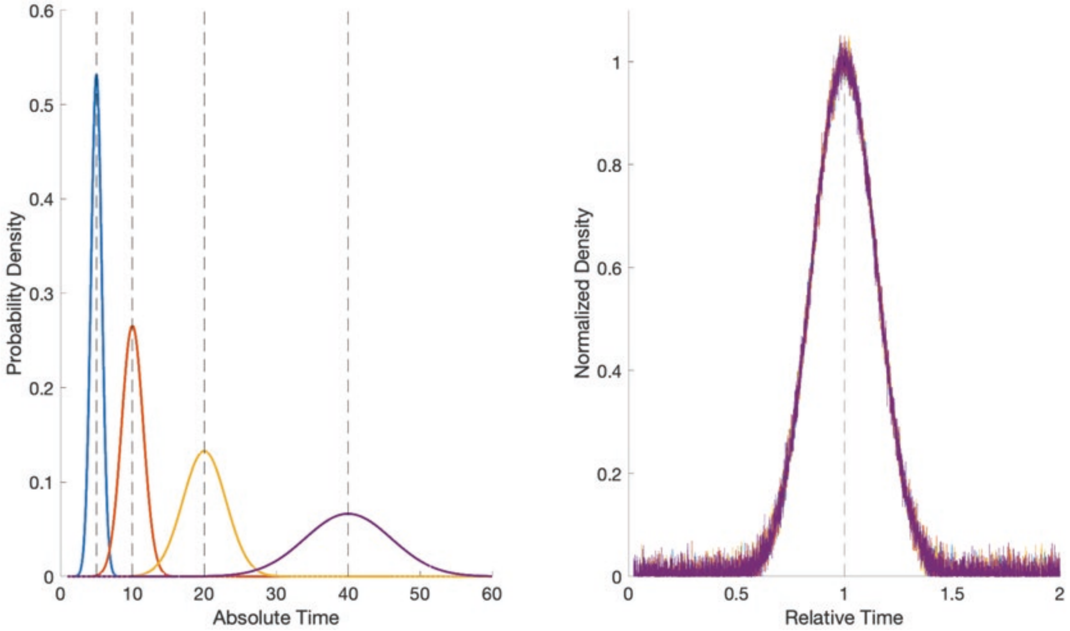


Fig. 1 *Left Panel:* Probability density functions describing anticipatory responses for four target intervals illustrated as Gaussian distributions with a constant coefficient of variation of .15 (σ/μ). The critical observation is that the probability density gets more spread out for longer intervals. *Right Panel:* Superposition of density functions when they are expressed in a relative time scale (absolute time/target duration). Note that the density values were perturbed by additive Gaussian noise since without noise these curves are completely indiscriminable

It is these widely-accepted signature statistical properties that constitute empirical/ psychophysical benchmarks for interval timing modelers and constrain the processing dynamics of the generative models (analogously, in two-alternative perceptual discrimination of the sort that yielded Weber's original law—e.g., comparing two weights to determine which is heavier—Simen et al. (2016) demonstrated that multiplying both stimulus intensities by a common factor leads to timescale invariance as well as Weber's Law behavior; but if this is not done, violations of timescale invariance emerge; see also Link, 1992 for similar findings). Thus, any successful model of interval timing should account for on-average accuracy and scalar variability of timing behaviors at least for suprasecond intervals. This is analogous to using a limited working memory span (e.g., 7 −/+ 2) or the serial position effect as empirical benchmarks for judging the success of working memory models. In this chapter, we will nonexhaustively overview modeling efforts in timing research and evaluate how they account for the above-mentioned psychophysical features of interval timing characterizing the nature of mapping between objective and subjective time (i.e., accurate timing and time-scale invariance). Our coverage of these models will target a general audience to make them understandable even to those new to the field (see Addyman et al., 2016, De Corte et al., 2022; Hass & Durstewitz, 2014 for earlier similar attempts). Our primary focus will be on the pacemaker-accumulator family of models since they represent a historically prominent and successful theoretical approach to interval timing.

Computational Models of Interval Timing

In a nutshell, the pacemaker-accumulator family of models assumes that the sense of time results from the accumulation of signals (e.g., action potentials, sub-threshold membrane potentials, hypothetical ticks, phasic dopamine signals) up to a certain level that can be treated effectively as a decision threshold. Although time has been a subject of interest since before modern science (e.g., St. Augustine, 2008), the earliest attempts to formally model interval timing were made in the early 1960s by Creelman (1962) and Treisman (1963). Treisman assumed that a timing agent integrates regularly emitted pulses while experiencing an event, and the number of pulses is encoded in the memory as a measure of the elapsed time. This would be akin to using a water clock in ancient Greece (clepsydra, which literally means water thief) or chronometric counting in daily life (one Mississippi, two Mississippi, …) with the critical difference that there is no symbolic tag for mental magnitudes in nonhuman animals and humans when chronometric counting is suppressed. Instead, it is the total number of pulses that generates the time representation, which would inherit noise elements of the generative processes that underlie the clock signals, as the cognitive atomic constituents of the resulting temporal (magnitude) representation.

Treisman made one simple but critical assumption that accounts for the psychophysics of interval timing; he assumed that the rate of the pulses is a normally distributed random variable sampled once per trial (thus staying constant within a given trial but resampled and thus varying across trials). This is similar to the assumption of trial-to-trial variability in the slope of the linear approach to the decision threshold in the Linear Ballistic Accumulator (LBA) model (Brown & Heathcote, 2008; see also Grice, 1968). Note that LBA also assumes trial-to-trial variability in the starting point of evidence accumulation. There is no reason for assuming such randomness other than the fact that neural information processing is noisy, and the normal distribution is a relatively safe assumption for capturing the underlying process as a common practice in cognitive modeling (Farrell & Lewandowsky, 2018). This simple assumption for accounting for variability in time estimates from one trial to the next also results, to a high degree of approximation, in a constant ratio between the standard deviation and the mean number of pulses counted for different target intervals (accounting for the scalar property). Figure 2 illustrates Treisman's model along with its predictions.

Although many behavioral scientists or neuroscientists would be at ease with normally distributed variation in the accumulation rate between trials, it strikes us as implausible that average interpulse durations would vary across trials but remain fixed within trials. It is more parsimonious to assume that interpulse intervals are highly variable both within and across trials.

Gibbon, Church, and Meck (e.g., Gibbon et al., 1984) came up with an alternative to Treisman's approach (the information processing implementation of Scalar Expectancy Theory; Gibbon, 1977) in which the source of the scalar variability was the memory instead of the clock. In the information processing implementation of SET (IPI-SET), interpulse intervals were explicitly formulated as exponential distributed (resulting in Poisson distributed pulses). Treisman made no specific distributional assumption regarding interpulse intervals but simply assumed that within an individual trial, they would be correlated. In contrast to Treisman (1963), in their version of Treisman's (1963) pacemaker accumulator model, the interpulse times were exponentially distributed (and thus with no memory of how long the previous interpulse time was) but without the between-trial variability in the rate of accumulation. Like Treisman and as characteristic of the pacemaker accumulator family of models, the timing agent was again assumed to encode and use the count of pulses as the measure of elapsed time.

But without between-trial variability in the rate of accumulation or threshold setting, there was still a dire need for a source of intertrial variability. This was simply because when one times intervals by counting Poisson distributed pulses and no between-trial variability in threshold setting or clock speed, the time-estimate coefficient of variation (CV) decreases proportionally to the square root of the timed interval (thus violating timescale invariance). This was the cost of fore-

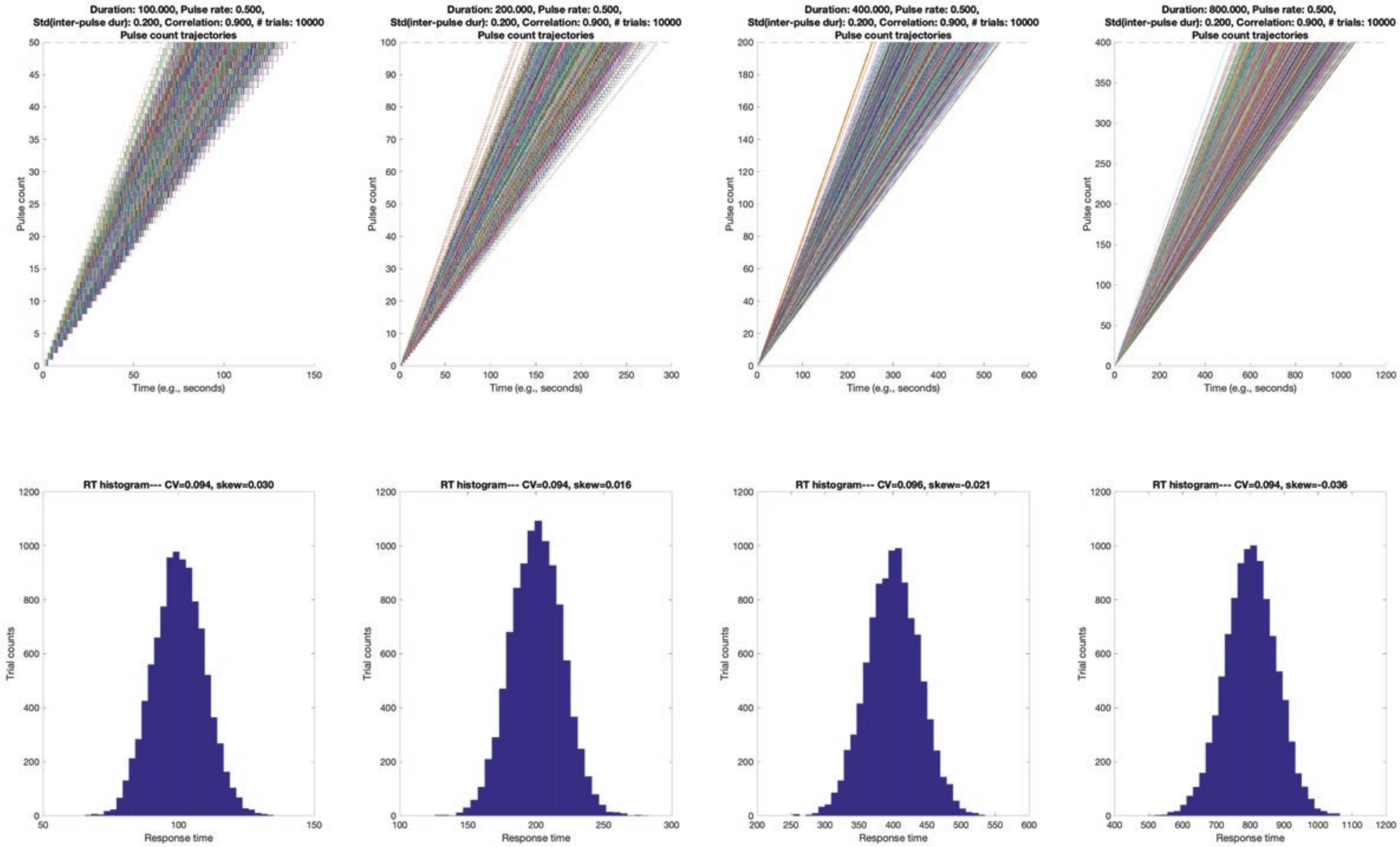


Fig. 2 Illustration of Treisman's pacemaker accumulator model of interval timing. The model is simulated for four different intervals achieved by modulating the threshold setting (higher thresholds for longer intervals). The density of threshold first crossing times is shown in the bottom panel as histograms of response times. Note that the coefficients of variation for four intervals are virtually identical. Simulations were performed 10,000 times per duration with a tick rate of 2 Hz, interpulse interval standard deviation of .2, and a correlation coefficient between any two within-trial interpulse durations of .9

going the intertrial variability assumptions of Treisman (1963), namely the variability in the "clock speed." The scalar property was rescued by assuming trial-by-trial variability in the encoding of counts (Gibbon & Church, 1990; Gibbon et al., 1984). This critical element of the model was harshly criticized by John Staddon (Staddon & Higa, 1996, 1999) as the "injection" of scalar variability rather than it being an emergent property of the inherent dynamics of temporal information processing. In formal terms, one can think of this encoding noise as a normally distributed random variable with a mean of 1, which was referred to as the *translation coefficient*. In psycho-mechanistic terms, the integrated amount of ticks in the accumulator (akin to working memory in the cognitive system or a capacitor in an electrical circuit) would be multiplied by normally distributed noise during its "consolidation" into the reference memory. This simple assumption results in more spread-out long-term memory representations of longer intervals, accounting for the scalar property at the level of memory processes/representation and therefore the time-scale invariance of the resulting behavior.

But this resulting time-scale invariance of the long-term memory representation is not sufficient for its manifestation at the behavioral level as the functional outputs of the internal clock. This brought about another critical assumption of IPI-SET, which is related to the decision process. IPI-SET assumes that the comparison of elapsed time to the long-term memory representation (more precisely, a random sample from LTM) was based on the ratio rather than the absolute difference between these two memory representations of time intervals. It is only with these assumptions that IPI-SET can account for the time-scale invariance, which puts the source of scalar variability in the memory and ratio-based decision rules rather than as an emergent outcome of the generative process that underlies the clock ticks. Note that the multiplicative memory translation noise can as easily be conceptualized as threshold variability that increases linearly with increasing threshold setting. This alternative interpretation was not exercised in IPI-SET.

In the simplest form of IPI-SET, when the organism monitors the duration of an event to match the temporal statistics of the environment, the model assumes that a Poisson clock keeps track of elapsed time and that it is constantly compared against a random sample interval from the long-term memory. The ratio of the difference between these quantities normalized by the random sample from the reference memory is used as the decision variable that is compared against a threshold value in real time. Since the decision variable is a normalized difference between elapsing time and memory for time, the lowest possible value of zero will be reached when the elapsed time matches the random memory sample. This results in a V-shaped trajectory for the decision variable as a function of elapsing time (see bottom left inset on Fig. 3 for two sample trajectories). When this decision variable takes a value that is lower than the threshold, the agent responds in anticipation of the reward and does not respond when the same decision variable is above the same threshold. Within this framework, the anticipatory response is assumed to be initiated when the decision variable first hits this threshold and terminated when the decision variable hits the same threshold a second time. In the peak interval procedure (Balcı & Freestone, 2020; Roberts, 1981), where subjects are trained to anticipate reward delivery after a constant delay following the onset of a timing stimulus, this mechanism results in the initiation of timing behavior before the reward delivery and its termination if the reward is omitted during the continued presentation of the timing stimulus (peak interval trials). The first instance is referred to as the *start time*, and the second instance is referred to as the *stop time*.

Loosely speaking, this mechanism is similar to the working of the conventional thermostat; the heater kicks in when the measured room temperature dips below a set level (lower than the desired ambient temperature) and keeps heating the room until the measured room temperature goes above another set level (higher than the desired ambient temperature). Such a mechanism

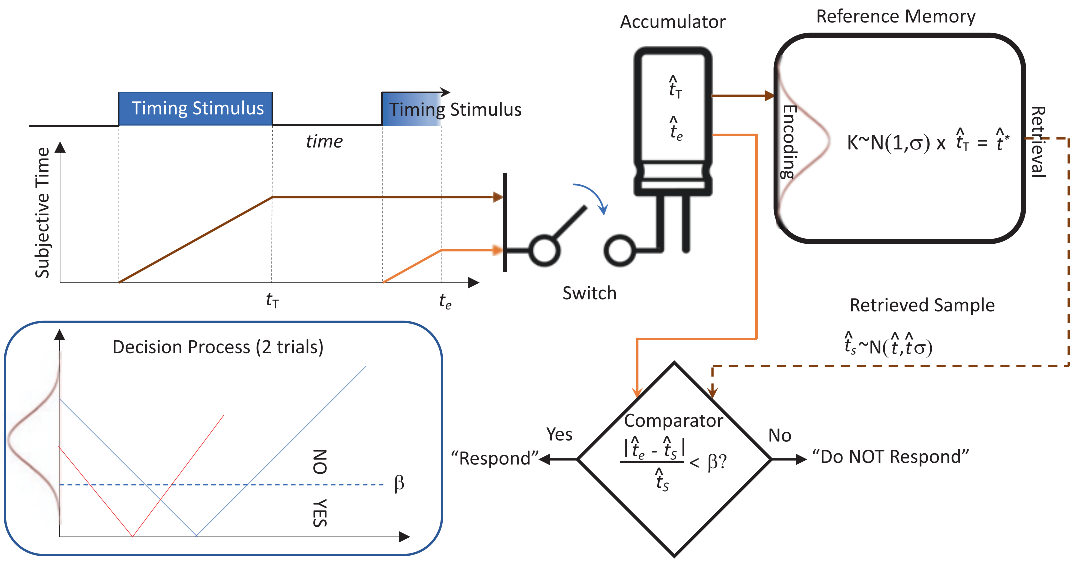


Fig. 3 Illustration of IPI-SET. The illustration shows temporal information processing in a single trial. Clock signals are integrated by the accumulator with the completion of the circuit (closure of the switch). Signals integrated in the accumulator stage can be viewed as the working memory representation of elapsed time. When the timing event is over, the information in the accumulator is transferred to the long-term memory during which it is multiplicatively perturbed by a normally distributed random variable with an unbiased mean (i.e., 1). When the timing agent compares elapsing time to previous experiences of the corresponding event, a random sample is drawn from the reference memory. The decision variable is the normalized difference between elapsing time and memory for time (normalized by the random memory sample), which reaches zero when the elapsed time matches the random memory sample. This results in a V-shaped trajectory for the decision variable as a function of elapsing time (bottom left inset). Briefly, the decision-making stage relies on ratio comparisons

keeps the room temperature around the desired level while saving energy and reducing mechanical costs by avoiding frequent fluctuations between the different states of the heater. IPI-SET has been used widely to explain behavioral phenomena in humans and other animals. For instance, the correlational patterns of when the agent starts and stops responding can provide information regarding the relative contribution of memory and decision processes to the variability in timing data as well as whether there is a single threshold or separate response thresholds for start and stop times (Gibbon & Church, 1990). Figure 3 illustrates the workings of IPI-SET and its control of anticipatory responses in the peak interval procedure.

Although the Poisson-distributed ticks of IPI-SET better capture what is known regarding neuronal activity (e.g., exponentially distributed interspike intervals) compared to assuming periodic pulses, this model lacked neurobiological/neuroanatomical implementation. Warren Meck demonstrated one of the earliest efforts to fill this conceptual gap by postulating the "Dopamine Clock Hypothesis" (Meck, 1983; 1996). He assumed that the dopaminergic activity of neurons in substantia nigra pars compacta (the nigrostriatal pathway that is primarily implicated in motor function and action selection) is a good candidate as the implementational counterpart of hypothetical clock ticks. The Dopamine Clock Hypothesis assumes that the phasic dopamine input in the nigrostriatal pathway is integrated by the dorsal striatum, which corresponds to the accumulator component of IPI-SET.

Meck, his colleagues, and others provided empirical evidence supporting this hypothesis by showing that timing behaviors were left-shifted by acutely administered dopamine agonists, whereas they are right-shifted by acutely administered dopamine antagonists (e.g., Meck, 1983, 1996; Maricq et al., 1981; Abner et al., 2001;

Çevik, 2003; Matell et al., 2006; Drew et al., 2003). These effects could be accounted for by the effect of dopaminergic modulation on the speed of the internal clock (rate of ticks fed into the accumulator). The disappearance of the acute effects of dopaminergic manipulations after chronic administration of the same compounds further supported this hypothesis by assuming calibration of the long-term memory representation to the new (biased) clock speed (Meck, 1983, 1996—but see Balcı, 2014 and Odum & Ward, 2007 for alternative accounts). If this was the case, then the termination of chronic treatment with dopaminergic agents should result in the effects in the opposite direction compared to the effects of acute administration on timing behavior. This is indeed what was observed in the seminal psychopharmacological work by Warren Meck, his colleagues, and others (e.g., Meck, 1983, 1996).

Another set of evidence in favor of the Dopamine Clock Hypothesis comes from lesion studies. Meck showed that lesioning of both substantia nigra pars compacta and dorsal striatum results in the loss of temporal control over anticipatory responses in the peak interval procedure. These detrimental effects of ablation can be rescued by the administration of the dopamine precursor levodopa for animals that have gone through lesioning of SNC but not the dorsal striatum (Meck, 2006). In other words, dopamine supplementation can rescue the loss of clock signals but not the accumulation of those signals by the dorsal striatum. Related work showed that the administration of cholinergic agents also resulted in shifts in the timing behavior, but critically, these effects emerged only gradually and as a result of chronic administration (Meck & Church, 1987). Based on these findings, cholinergic function was implicated in the neurobiological implementation of the translation coefficient.

A more biophysically detailed approach to modeling the neural basis of interval timing (in light of IPI-SET) was taken on by Warren Meck and Mathew Matell, followed by Sorinel Oprisan and Catalin Buhusi (Matell & Meck, 2000, 2004; Meck et al., 2008; Oprisan & Buhusi, 2013). Their model was computationally inspired by Miall's (1989) beat frequency model, which explains how units that spike frequently, but with different periods, could time intervals by encoding them as the time of synchronous firing of the units. These times of firing synchrony could occur over periods much longer than the interfiring period of any individual unit. However, the beat frequency model could not account for the scalar property of interval timing due to the lack of a noise element. Meck and Matell mapped a noisy version of Miall's model onto the corticostriatal pathway, giving rise to the name "striatal beat frequency" (SBF). Their basic assumption was that time intervals are encoded by cortical oscillations that can be primarily attributed to a population of oscillating neurons in the prefrontal cortex and, therefore, subject to the effects of dopaminergic signaling in the mesocortical pathway. They assumed that dopaminergic signaling in the mesocortical pathway would phase reset (synchronize) the cortical oscillations, after which each oscillating neuron would oscillate according to its endogenous frequency. The temporal pattern of moments at which these oscillators with different frequencies happened to fire synchronously, or "beat," generated by a large number of neurons, would form a unique pattern as a function of time and resultingly a reliable-enough pattern for timing intervals in different trials. But these signals would be of no use to the motor system unless synchrony was detected by a simple form of neural network similar to a perceptron (Rosenblatt, 1958).

To this end, the SBF model assumes that GABAergic medium-spiny neurons in the dorsal striatum implement such coincidence detection for the specific pattern of different oscillatory signals (beats of cortical oscillators) for coding time. The role of the nigrostriatal dopamine pathway in this model is attributed to the training of the medium spiny neurons to recognize the beats of the cortical oscillators corresponding to biologically significant events through dopamine-dependent (through nigrostriatal pathway) long-term potentiation and depression (LTP and LTD). Specifically, nigrostriatal dopamine input strengthens those corticostriatal synapses that were active at the time of the reinforcement (LTP) while weakening those that were inactive at this time point. As a result of this phasic input, contingent on feedback or reward marking the end of the interval, the efficacy of the synapses connecting a group of oscillating cortical neurons and medium spiny neurons (MSN) can be changed such that MSNs are trained to recognize the corresponding unique cortical input (perceptron role of the striatum).

This model can be likened to recognizing specific moments and composition that precedes and follows those specific moments in orchestra music (where each instrument can be treated as a different oscillatory basis function) after listening to it repeatedly. When the learning for a given time interval is established, and as predicted by the temporal difference learning algorithm, the phasic dopamine input would occur at the beginning of the timing cue to push striatal neurons to their resting membrane potentials. Crucially, the bistability of MSNs has been shown as a property that makes them suitable for being responsive (driven to their up/excitatory state ~−60 mV from down state ~−90 mV) to highly coherent cortical input from oscillating neurons (e.g., Gu et al., 2015). Figure 4 illustrates the functional architecture of the SBF model. The SBF model accounts for the scalar property by introducing trial-by-trial change in the speed of oscillators and the neuronal firing thresholds in a globally coherent manner to preserve the informative value of the oscillations. If these variations that commonly apply to different oscillation frequencies are violated, the information value of the cortical signals with respect to time would be lost.

Oprisan and Buhusi (2013) have extended the SBF model to account for the scalar property as an emergent property of the model (rather than enforcing it with globally coherent noise—see Buhusi & Oprisan, 2013) as well as to account for the effect of dopaminergic manipulations outlined earlier in the chapter by modeling the effect of these manipulations on the firing frequency of

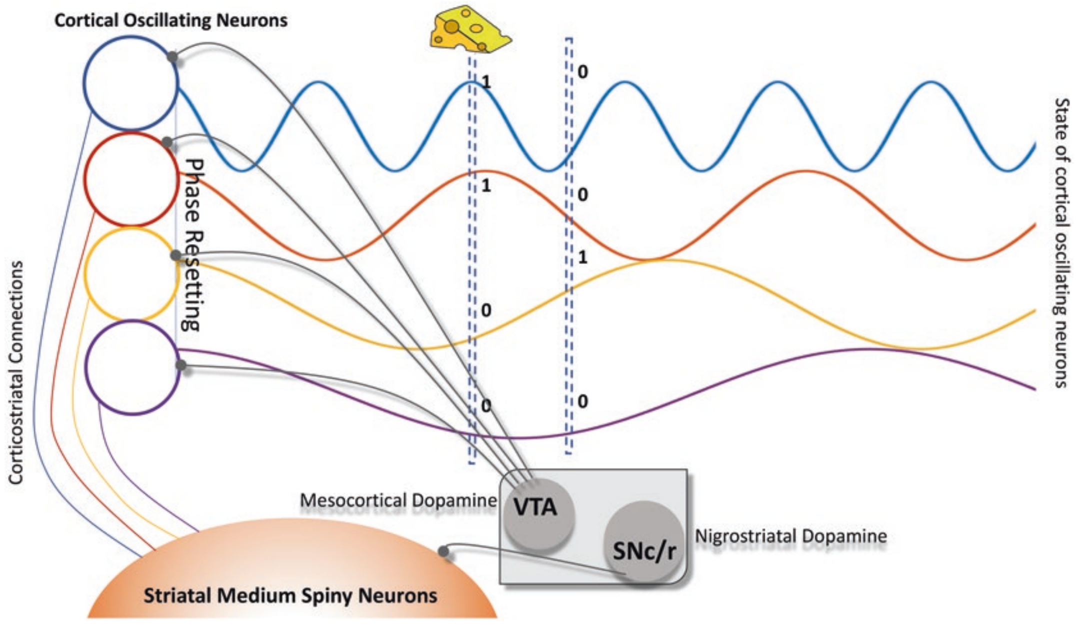


Fig. 4 Illustration of Striatal Beat Frequency model with four cortical oscillatory units that are phase reset with mesocortical dopamine input. Striatal medium spiny neurons serve as coincidence detectors for recognizing the critical pattern of cortical oscillatory states. The oscillatory states are recoded as zero and one to demonstrate the change in the population code as a function of time. In this example, the code for the target interval is 1100 while keeping the identity of their neuronal sources. Nigrostriatal dopamine input trains striatal medium spiny neurons to recognize these patterns

cortical oscillators (Oprisan & Buhusi, 2011). Unlike the original version of the SBF model, these variants utilized biophysically realistic neurons with explicitly formulated ionic dynamics. For instance, Oprisan and Buhusi (2013) showed that accurate timing and time-scale invariance emerge independent of the source of the noise implemented in the SBF model (i.e., frequency variability, memory variability). These variants of the SBF model pointed to its robust nature in terms of accounting for the psychophysical properties of interval timing along with explicit neurobiological characterization. Empirical evidence for the SBF model comes from a variety of studies conducted with animal models (Emmons et al., 2017, 2019, 2020; Matell et al., 2003).

Note that an earlier connectionist model interval timing based on the phases of an oscillating neuronal set at the time of reinforcement was also introduced by Church and Broadbent (1990). The model compared the oscillatory states to the long-term memory representation of critical oscillator phases, which guided the timing behavior and accounted for the scalar property. But this model did not receive attention arguably because of the lack of a clear neuroanatomical mapping such as that proposed by the SBF Model.

At the algorithmic level of description, IPI-SET and Treisman's model of timing can be treated as a threshold adaptive approach to interval timing, where different intervals are monitored by keeping an on-average constant clock speed and relying on the pulse-count sum as a measure of elapsed time. Critically, this is not the only way of keeping track of different intervals while ensuring that predictions abide by the psychophysical properties of interval timing. An alternative approach is to adjust the speed of the clock (i.e., adaptive clock approach) and use a constant threshold to time different intervals (or both, although this would not be preferred on the grounds of parsimony).

One of the early approaches to using adaptive clocks to account for timing accuracy and the scalar property was proposed by Killeen and Fetterman (1988). According to their Behavioral Expectancy Theory (BET), time intervals are captured based on behavioral states that are sequentially linked to each other (e.g., rodents frequently engage in sequential grooming behaviors—elliptical stroke → unilateral stroke → bilateral stroke → body licking; Kalueff et al., 2016). With the onset of the timing stimulus, the first behavioral state is activated, and this activation propagates through a chain of states, with exponentially distributed intervals between state transitions. When the propagation reaches the terminal state in the chain, the temporally predictable event is expected to occur. The result is effectively a Poisson clock (embedded in a chain architecture) because the Poisson distribution denotes the number of events that happen during a given period when the event spacings are exponentially distributed. This important statistical feature is indeed widely used in different fields (e.g., the mathematical study of the waiting lines—queuing theory—Sundarapandian, 2009). Crucially, BET assumes that the speed of the between-state transitions increases with higher reward rates (and thus delay to rewards while keeping the reward magnitude constant) associated with the timing stimulus. Analogously, an increase in the incoming rate of customers at a cafe results in meeting a constant customer count earlier during a happy hour. BET is thus effectively a clock-speed-adaptive model of interval timing that accumulates clock signals up to a constant threshold. Figure 5 (top panel) illustrates the generative dynamics and assumptions of BET based on the escape of a particle from energy wells until it completes a specific path.

The threshold first crossing times of accumulation of Poisson distributed signals are gamma distributed, whose coefficient of variation stays constant when the clock speed is inversely proportional to the target delay (= $1/t$). In this way, BET accounts both for high timing accuracy and scalar property. But note that under this critical assumption, BET also predicts that the clock speed should increase with increasing reward magnitude (when the target interval is constant) while empirical evidence suggests that such experimental manipulations affect threshold setting for willingness to initiate anticipatory responses (i.e., start times) rather than causing

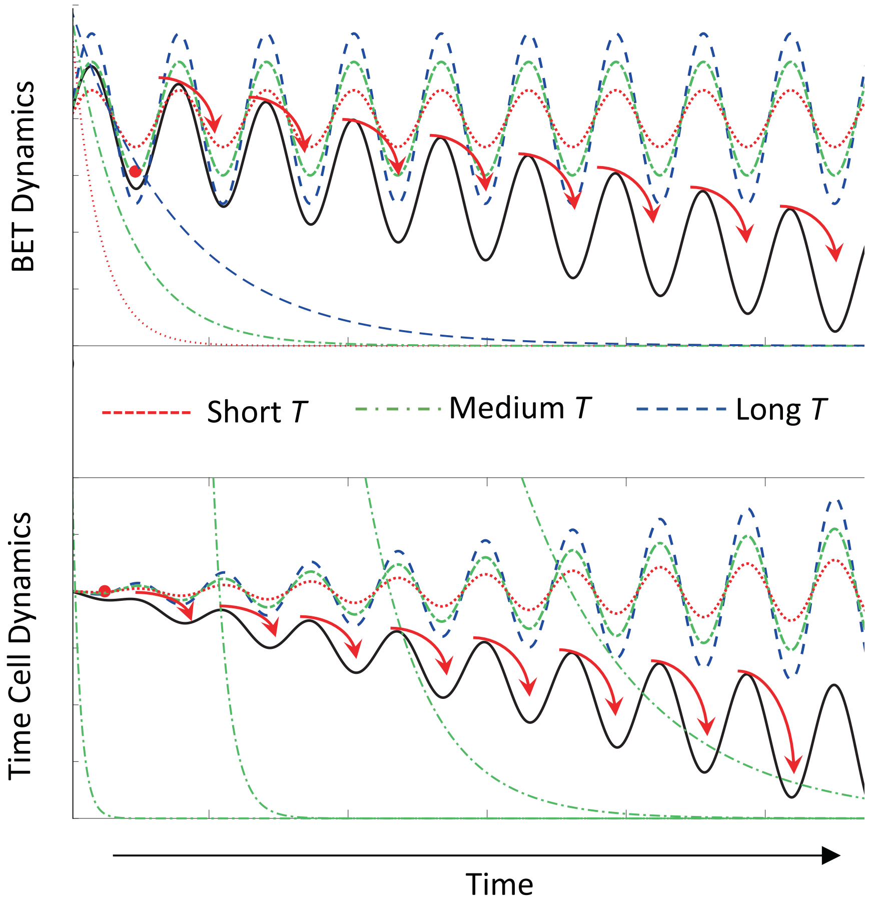


Fig. 5 Illustration of the signal propagation dynamics according to BET (top panel) and time cell functional architecture (bottom panel). *Top Panel:* The illustration is based on the unidirectional energy well-escape behavior of a particle (approximately exponential distribution), which can also be thought of as the flow of incoming liquid from one well to the next (left to right) for convenience. The rescaling characteristics of the two approaches are shown by delinearized amplitude differences in the sine waves (dashed, dotted-dashed, dotted sine waves). The exponential distribution that would account for energy well escape behavior is shown for three different amplitudes on the top panel. Note that the exponentially distributed escape times are slower for sine waves with higher amplitude. *Bottom Panel:* The same dynamic is shown with dotted-dashed green sine waves for subsequent wells on the bottom panel. Note that the exponential escape times become slower for later points in time, which translates into longer dwelling times of the particle in the subsequent wells corresponding to the longer activity of time cells that are activated later in time

the predicted full leftward shift in timed response curves (Balcı, 2014; Ludvig et al., 2007).

One of the common problems of the models outlined above is that they do not have explicitly specified learning rules (except for SBF for which synaptic plasticity in the dorsal striatum is implicated). Note that SET's encoding function can also be treated as a learning instance with a learning rate of 1. The Learning to Time Theory (LET) of Machado (1997; see also Machado et al., 2009) has overcome this gap in earlier models by assuming that each state of BET is associated with a common response form(s), and the degree of association between the time units and the behavioral unit is subject to plasticity through delta-based trial-and-error learning.

Thus, unlike BET, LET allows the learning of associations between time and behavioral outputs and thereby accounts for the acquisition of timed responses (but see Balci et al., 2009 for counter-evidence to predictions of LET regarding acquisition). Note that in essence, LET also constitutes a clock-speed-adaptive model of interval timing (inherited from BET's functional architecture) with an associative interface that links the temporal representation to timing responses and thus adds learning capacity to interval timing (e.g., learning to do different things at different points during an interval). This contribution of LET to the interval timing literature filled an important gap, particularly when considered within the framework of conventional associative learning models.

Another problem that is common to the models outlined above (including LET regarding acquisition) is that all these models effectively assume a perfect counter function in which the cumulative value of the counted clock signals increases monotonically as a function of time (activation does not back-propagate). In other words, although there is randomness manifested at the level of intersignal generation times (as an exponentially distributed random variable), the counted ticks are not subject to perturbation (akin to counting errors) or never lost (or follow a very strict unidirectional path as in the case of the propagation of action potentials through the axon).

One approach that overcomes the need for such strong assumptions and that unifies interval timing research with research on perceptual decision making and memory is to model the internal stopwatch as a process of noisy evidence integration, which is mathematically identical to the movement of a tiny particle in liquid or gas undergoing drift (from a current) along with some random movement, or diffusion (from heat). Drift-diffusion models (DDMs) have been widely used for decades to model perceptual decision making because they can explain both the accuracy and the entire distribution of response times in two-alternative forced choice tasks (e.g., Ratcliff, 1978; Ratcliff & McKoon, 2008). For instance, when participants are repeatedly asked to determine the direction of coherent motion in an array of dynamically moving dots, a subgroup of which move coherently either to the left or right while others move in random directions, the resulting choices and associated response times are numerically analyzed together to estimate the parameters of the underlying decision process (including the drift rate and level of diffusion) based on the shape and relative density of response times for correct and incorrect choices. This model extends the standard psychological approach that classically treats response times and accuracy as isolated outputs of the same underlying decision process. This standard approach does not utilize all the information in the data. This is nicely captured by the parable of the blind man and elephant retold by Rumi. This parable is about blind men who have never come across an elephant and try to learn about the elephant by touching only one part of the elephant's body (leg, trunk, ear, or tusk). As a result, each blind man would describe the elephant based on what they touched resulting in widely different descriptions of the entity (missing the forest for the trees).

The primary assumptions of the DDM in the case of two-alternative forced choice tasks are that the sensory evidence is noisy, the difference between the evidence supporting two different hypotheses constitutes the decision variable, this noisy decision variable is integrated over time toward one of the two thresholds, and when the decision variable hits one of these two thresholds the corresponding decision is made. The time it takes for the decision variable to reach that threshold is taken as the response time (composed of the decision time plus a nondecision time-related sensorimotor delay). This model also allows the integration of biases (e.g., due to prior belief states, different probability of different signals) as well as trial-to-trial variation in the core parameters (Ratcliff et al., 2016).

The DDM can be viewed as the continuum limit of the sequential probability ratio test (SPRT—Wald, 1947). In the SPRT, the agent computes the log-likelihood ratio of continuously incoming data under two different hypotheses and reaches a decision when this log-likelihood ratio reaches a given criterion (positive for one hypothesis and negative for the alternative hypothesis). Importantly, for a range of simple perceptual tasks with a stationary signal-to-noise ratio, the SPRT is the optimal procedure, which maximizes accuracy for a given response time and minimizes response time for a given accuracy (Wald & Wolfowitz, 1940). Thus, the DDM is the optimal procedure for two alternative forced-choice behaviors in which aspects of the task (signal-to-noise ratio, prior probability of signals) are stationary within and across trials (e.g., Bogacz et al., 2006).

The DDM can be formalized as an ordinary stochastic differential equation:

$$dx = A \cdot dt + c \cdot dW$$

where dx is the change in the position of the decision particle x between two thresholds, A refers to the average rate at which the decision particle is displaced in this area toward the correct threshold (slope by which it approaches the correct threshold), dt refers to an infinitely small time interval (denoting the continuous nature of the decision process) and dW refers to the Gaussian white noise (repeated samples that are fully independent). Thus, c, the diffusion coefficient, determines the amount of Gaussian noise added to the trajectory of the decision particle at each time point (dt), which is typically assumed to be constant for different signal strengths in decision making. When the decision particle reaches one of the two thresholds $-z$ or z (one below and one above the starting point of x [x_0]), a decision that corresponds to that threshold is made and the time it takes for the decision particle to reach that threshold denotes the decision time.

When the decision particle is bound to hit only one of these thresholds (e.g., due to a very high drift rate and high decision threshold with respect to the diffusion coefficient), the threshold first crossing time distributions are inverse Gaussian (aka Wald) distributions (Luce, 1986), where the mean equals z/A, and the shape parameter of the inverse Gaussian distribution equals z/c^2. Thus, there is a closed-form solution for the threshold first crossing times of the decision particle when there is a single threshold (as opposed to the need for numerical solutions for the two threshold version of DDM).

The problem with modeling interval timing as a drift-diffusion model with a constant diffusion coefficient (c) either by modulating the single threshold or drift rate to time different intervals is that it does not result in time-scale invariant threshold first crossing times for different durations. As outlined earlier, this is a terminal failure for any model of interval timing. For instance, when different intervals are timed with different drift rates and a constant threshold (clock speed adaptive approach to timing), then CV grows with the square root of T (z/A) (Simen et al., 2013). If, on the other hand, one times different intervals by keeping the drift rate constant but adjusting the threshold (threshold adaptive approach to timing), then CV decreases with the square root of T. In both cases, the time-scale invariance of interval timing is systematically violated by different changes in timing imprecision/uncertainty as a function of the target time. In the first approach, the longer intervals end up having a larger CV whereas, in the second approach, they end up having a smaller CV. In contrast, a successful model should predict a constant CV.

Then how can one model interval timing with a diffusion process? The answer lies in two critical assumptions regarding how the drift rate and diffusion coefficient scale with target times, which is assumed to be constant in the conventional DDM as it is applied in decision science. The time-scale invariance of threshold first crossing times is achieved when the drift rate is inversely proportional to the target time and the diffusion coefficient is scaled as the square root of the drift rate while keeping the threshold constant. But this naturally begs the question of why such a relationship should exist in the first place and how making this assumption is different from injecting the right kind of noise into the generative process to achieve time-scale invariance. Crucially, this is an emergent feature that results from a few assumptions outlined in Simen, Balci, de Souza, Cohen, and Holmes (2011a; see also Simen et al., 2013), such as exponentially distributed interspike times and balanced excitatory and inhibitory inputs (as a common feature of neural information processing; e.g., Okun & Lampl, 2009).

Specifically, the time adaptive opponent Poisson drift-diffusion model (TopDDM) keeps track of the cumulative sum of excitatory Poisson inputs with the rate I (by probabilistic and independent activation of bistable units from OFF to ON) and inhibitory Poisson inputs that are a fixed proportion of excitatory inputs $\gamma{\cdot}I$; the average accumulation rate is the difference between the rates of excitatory and inhibitory Poisson inputs ($I - \gamma{\cdot}I$). The variance of such a competitive process is the sum of the variances of each spiking count, which is equal to the sum of the rate parameters of each Poisson input train ($c^2 = I + \gamma{\cdot}I$). This results in a diffusion standard deviation that is proportional to the square root of drift.

$$dx = (1-\gamma)\cdot I\cdot dt + \sqrt{(1+\gamma)\cdot I}\cdot dW \qquad (1)$$

When we define $A = (1 - \gamma){\cdot}I$, and define $m = \sqrt{(1+\gamma)/(1-\gamma)}$, substitution into Eq. 1 with $c = \sqrt{(1+\gamma)\cdot I}$, gives $c = \sqrt{(1+\gamma)/(1-\gamma)\cdot A}$. For convenience, after getting $\sqrt{(1+\gamma)/(1-\gamma)}$ out of the equation as a constant of proportionality (m), c becomes $m\sqrt{A}$, thus resulting in a particular form of drift-diffusion process in which m is constant across durations and a diffusion coefficient (c) scales with the square root of the drift rate (A).

$$dx = A\cdot dt + m\cdot\sqrt{A}\cdot dW \qquad (2)$$

Briefly, the specific relationship between the drift rate and diffusion coefficient that enables a drift-diffusion process to account for time-scale invariance of timing behaviors comes for free out of the assumption of time-adaptive opponent Poisson processes and balanced excitation and inhibition. Figure 6 illustrates the resulting drift-diffusion process for three different target intervals.

At the neuronal activity level, the threshold can be treated as the firing rate of the population of neurons implementing the accumulator that is needed for the corresponding decision to be made. At around the same time as us, Francois Rivest and Yoshua Bengio developed an identical drift-diffusion model of timing that simply posited that noise was proportional to the square root of the drift (Rivest & Bengio, 2011). The two models were developed in parallel without mutual awareness of the other group's effort until a common colleague (Elliot A. Ludvig) informed FB during his visit to Princeton University.

One way to simulate the TopDDM is to generate a sequence of spikes according to one Poisson rate (exponentially distributed interspike times with low mean) and another sequence of spikes according to a lower Poisson rate (exponentially distributed interspike time with high mean). Then, one would compute the cumulative difference between the cumulative spike count of higher rate and lower rate Poisson processes. When the cumulative interspike time at which the cumulative difference between the two Poisson processes reaches a predetermined level (i.e., threshold), the timing behavior is manifested. Two built-in functions of Matlab would be sufficient to conduct these simulations (i.e., *exprnd* and *cumsum*). For instance, *exprnd(.5,1,100)* would generate 100 interpulse durations with a mean of .5. If this is the excitatory Poisson process, then one can set the first input argument to the inhibitory Poisson process to $1/\gamma \times .5$; which would result in 1 with a γ parameter of .5 (i.e., *exprnd(1,1,100)*) and 2.5 with a γ parameter of .2 (i.e., *exprnd(2.5,1,100)*). Take the cumulative sum of low rate exponential and the cumulative sum of high rate exponential and assign values of 1 to the cumulative sum values of high rate exponential in a second column and −1 to the cumulative sum values of low rate exponential in the second column. Concatenate the two arrays and then sort all rows according to the cumulative sums and then sum over the second column (+1 and −1 s) and find the value on the first column that corresponds to a value equal to or higher than the threshold value (sum of signed ones). This will denote the reproduced time for the corresponding trial. When one repeats

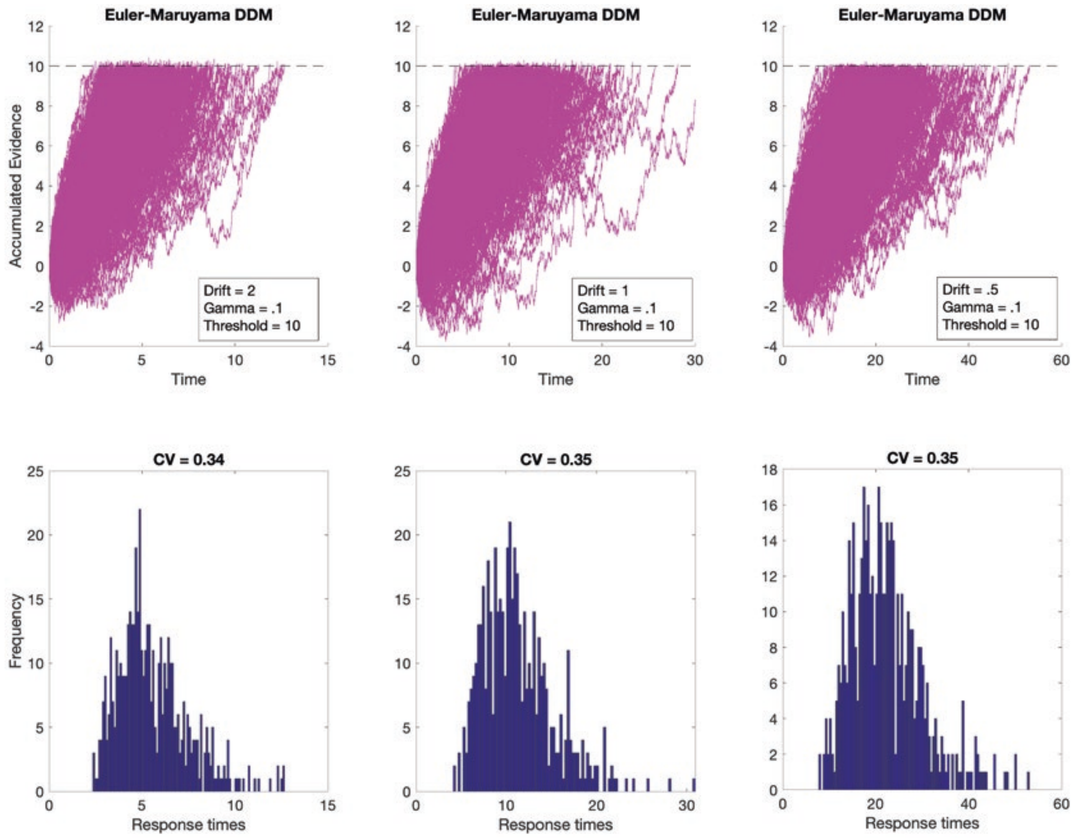


Fig. 6 Illustration of TopDDM through Euler–Maruyama methods for the approximate numerical solution of the stochastic differential equation. Note that the trajectory of the decision particles appears very similar across three targets (top panel—different columns) because the x-axis upper limits are adjusted such that they are inversely proportional to the drift rate. The same rescaling applies to the threshold first crossing times of the decision particle (bottom panel—different columns). The estimates of the coefficient of variation were nearly identical for three different durations. The drift rates of 2, 1, and .05 were used in these simulations with a γ (the proportion of inhibition to excitation) value of .1 and a threshold value of 10

this process, one can get the distribution of temporal reproductions, which will be best accounted for by the inverse Gaussian distribution (Simen et al., 2011a).

A simpler and more efficient way of simulating TopDDM is using the Euler–Maruyama methods for the approximate numerical solution of the stochastic differential equation by repeatedly computing the following sum:

$$x_{\text{new}} = x_{\text{old}} + (1-\gamma)\times I \times \Delta t + c \times \sqrt{(1+\gamma)\times I \times \Delta t} \times \mathrm{N}(0,1),$$

with Δt representing a small time step. We use the square root of Δt in the noise term because variance grows with *dt* and thus standard deviation grows with square root of Δt. For very low values of Δt, this equation approaches the continuum limit of the opponent Poisson DDM as the DDM approximates the continuum limit of SPRT.

Empirical support for the TopDDM comes from electrophysiological studies with nonhuman animals (Merchant & Averbeck, 2017; Komura et al., 2001; Leon & Shadlen, 2003; Jazayeri & Shadlen, 2015; Murakami et al., 2017—although some of these studies were not designed to test differential slopes of ramping activity to time different intervals) as well as humans (Macar & Vidal, 2003; but see Kononowicz & van Rijn, 2014). From these, Merchant and Averbeck

(2017) clearly demonstrated that TopDDM accounts for not only psychophysical properties of timing behavior of monkeys in a rhythmic timing task but also their higher order statistics of the response time distributions and the autocorrelation structure in neural representation predicted by TopDDM. To highlight the conceptual continuity between different models of interval timing, we would like to highlight that BET is a special case of TopDDM, where the γ (inhibition to excitation ratio) parameter is set to 0 (i.e., pure Poisson clock). But the problem with using BET (i.e., using γ of zero) is that the interspike intervals would have to be too long to be neurally plausible to produce reasonable CVs. Although TopDDM can also be criticized for the difficulty of sustaining an integrative state of neuronal populations over long periods, Simen et al. (2011b) showed that neural integration could take place over much longer time scales based on sequentially triggered interval timers.

Marr's levels of analysis provide a useful meta-theoretical framework for the conceptualization of interval timing models. Marr (1982) introduced three hierarchical levels of analysis for any cognitive system. The highest level of understanding of a system (the computational level) entails understanding what computational function the cognitive system is carrying out (e.g., computing the route to a destination). This corresponds to the evolutionary analysis of the cognitive system. The second level of analysis (the algorithmic level) focuses on the procedures carried out to perform the computation (e.g., searching routes with a depth-first search tree vs. randomly generating routes and assessing whether they are valid). The lowest level of analysis (the implementation level) focuses on how the representation and algorithmic solution can be realized at the hardware level (i.e., physical implementation). Under this meta-theoretical framework, the models outlined last constitute the algorithmic level account of interval timing based on noisy evidence integration processes (Marr, 1982—see Fig. 10).

Simen et al. (2011a) also developed an implementation-level account of this model based on a neural network with four layers, the Stochastic Ramp and Trigger (SRT) model. The input layer of this architecture is composed of detector neurons that fire with the onset of the timing stimulus. This function characterizes these units as sensory neurons (e.g., De Corte et al., 2022). The input from detector neurons activates the second-layer units with strong recurrent excitation, which results in sustained activity even when the neurons are excited by inputs below the threshold for switching on. This property of neurons in the second layer causes them to fire maximally throughout the timing epoch. The units at the third layer constitute the ramping units that can integrate inputs from the second layer units based on balanced recurrent excitation. It is this balanced self-excitation that enables these units to perfectly integrate information rather than exhibiting a bistable dynamic as displayed by the units of the second layer with stronger self-excitation. The rate at which ramping activity occurs depends on the number of active, second-layer tonic units that feed input into the third layer. More tonic units feeding input into the ramping units result in steeper ramps, which would be required for timing shorter intervals. Finally, the fourth layer of the architecture implements threshold crossing based on input received from the ramping units of the third layer and strong recurrent excitation. These units fire when the firing rate in the ramping units reaches a constant level. Figure 7 illustrates this functional architecture.

One of the open questions is what the implementational level counterpart of TopDDM is in the brain. For instance, although it was originally conceived as a balance of glutamatergic and GABAergic signaling within cortical populations, the balance between excitatory and inhibitory inputs may be underlain at the systems level by the dopamine-dependent balance between the direct and indirect pathways of the striatum. Within this framework, over or under-dopaminergic input would be expected to modulate the balance between excitatory and inhibitory inputs, which would, in turn, predict the clock speed effects of dopaminergic modulation. This hypothesis also predicts that the timing impreci-

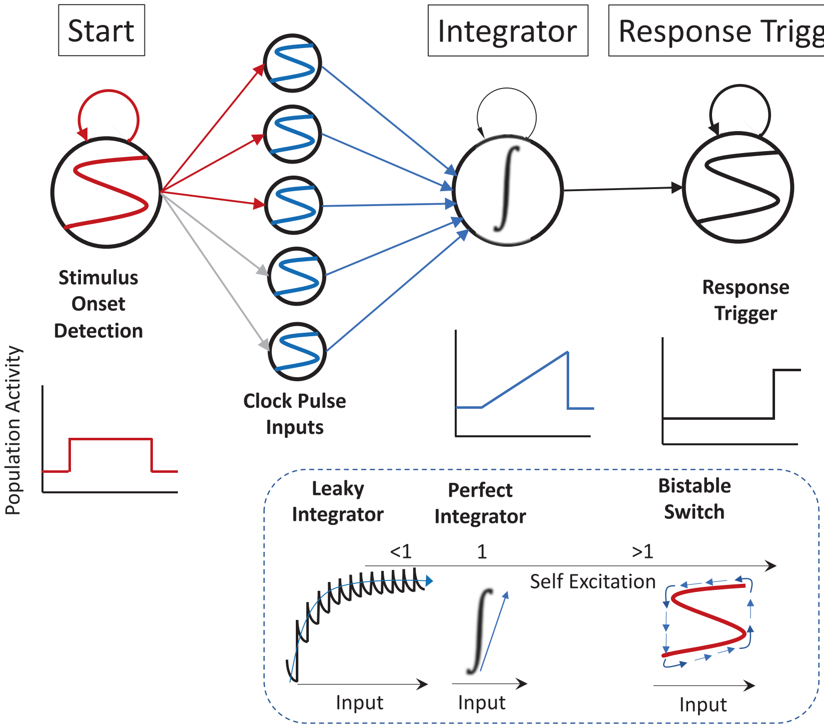


Fig. 7 Physical (neural) implementation level explanation of the time adaptive opponent Poisson drift-diffusion model of interval timing as a Stochastic Ramp and Trigger model. The sensory detection of the onset of timing stimulus and the switch units that feed inputs to the integrator is realized by bistable units. The integrator sums the excitatory and inhibitory input from the bistable switch units and in turn drives another bistable unit that implements the decision to respond. The lower dashed inset shows the excitation profile of these units based on the degree of self-excitation

sion should increase with higher dopamine levels and decrease with lower dopamine levels. These questions should guide future research to investigate the neural basis of TopDDM.

The TopDDM has been applied to model performance in a variety of interval timing tasks (e.g., temporal bisection—Church & Deluty, 1977; peak interval procedure—Roberts, 1981) as well as temporal error monitoring (e.g., Akdogan & Balci, 2016). In parallel to the divergence between the standard psychological treatment of choice accuracy and response times in an isolated fashion (remember the parable of blind men and elephant) vs. decision-theoretic analysis of these decision outputs in a unified fashion in two-alternative forced choice tasks, Balci and Simen (2014) offered a new analytical approach to the analysis of temporal bisection task by taking into account not only psychophysical accuracy functions (as typically done in standard psychophysical approach) but also the corresponding response times within the framework of noisy evidence accumulation (see also Ratcliff, 2014; Akbiyik et al., 2022; Akdogan & Balci, 2016).

In the temporal bisection task, subjects are initially trained to discriminate between short and long reference intervals by emitting two different responses for categorizing them based on their subjective similarity to short and long reference intervals (Church & Deluty, 1977). Once the categorization accuracy reaches a certain criterion, experimenters include intermediate test durations and do not provide any feedback for their categorization. When the proportion of long choices is expressed as a function of test durations, one typically gets a sigmoidal (s shape) psychophysical function that can be well-captured by a logistic function or cumulative density functions of unimodal distribution functions (e.g., Weibull given the different many shapes it can take based on its shape parameter). But the response times associated with these categorizations are typically not considered in the analysis of the performance. In exploring this valuable aspect of the data, Balci and Simen (2014) showed that the response times of short categorization responses slow down, whereas the response times of long categorization responses speed up with longer test durations. They offered a unified account of choice proportions and response times in terms of a sequential drift-diffusion process.

The first-stage drift-diffusion process was assumed to be adapted to hit a threshold at the long reference intervals (adhering to the principles of the TopDDM); timer stage. The state of the timer at the end of the test duration was assumed to set the starting point as well as the drift rate of the second stage decision process (based on the difference between the current starting point and unbiased starting point). If the state of the timing process was above the starting point at the end of the timing stimulus, the drift rate would be positive, and the decision variable would, on average, approach the upper threshold related to the long categorization. On the other hand, if the state of the timing process was below the starting point, the drift rate would be negative, and the decision variable would, on average, approach the lower threshold related to the short categorization. Note that what determines the state of the activity level at the end of the timing stimulus is not only the test duration but also the diffusion coefficient that perturbs the average path of the decision particle. This sequential decision process can account for both sigmoidal psychophysical curves and the observed modulation of response times as a function of test interval duration. This model demonstrated how interval timing and temporal decision making processes could be incorporated into a common theoretic framework. Recent work has provided empirical neural evidence in support of this model (e.g., Ofir & Landau, 2022; Wiener et al., 2018) (Fig. 8).

More recently, Luzardo et al. (2017) modeled timed, anticipatory responses in the peak interval procedure within the framework of TopDDM by assuming two response thresholds, the lower one for the initiation of timed responses (start time) whereas the higher one for the termination (stop time) of the timed responses (as opposed to a single threshold in the original TopDDM). Importantly, they showed that when the first threshold is allowed to be variable, this model accounts for the correlation patterns between the start and stop times as observed in the empirical data (Gibbon & Church, 1990). Finally, Akdogan and Balci (2017) showed that monitoring of the direction and magnitude of errors in timing behavior could also result from the comparison of the two TopDDMs, one underlying perceptual timing whereas the other underlying motor timing. They assumed that the order of threshold first-crossing times of these timers can determine the direction of the second-stage drift-diffusion process, which would in turn inform the agent regarding the direction and magnitude of errors in their time estimates. All these efforts demonstrate the explanatory power of the TopDDM framework in accounting for different variants of timing behaviors ranging from very simple timed responses to temporal awareness (Balcı, 2022).

A limitation of time adaptive models is that it is not clear how one can determine which clock speed to use during the first exposure to a new duration (sometimes referred to as the “infinitude of the possible” problem; Gallistel & King, 2010). The adaptive and fast learning rules of TopDDM and Rivest and Bengio (2011) can overcome this problem as they allow adaptation of the drift rate so that the threshold is hit at the correct time after experiencing a new interval (after hitting the threshold earlier or later than the occurrence of the new timing event). With a maximum learning rate, this rule requires only a single learning iteration so that the model can encode new durations as quickly as classical pacemaker-

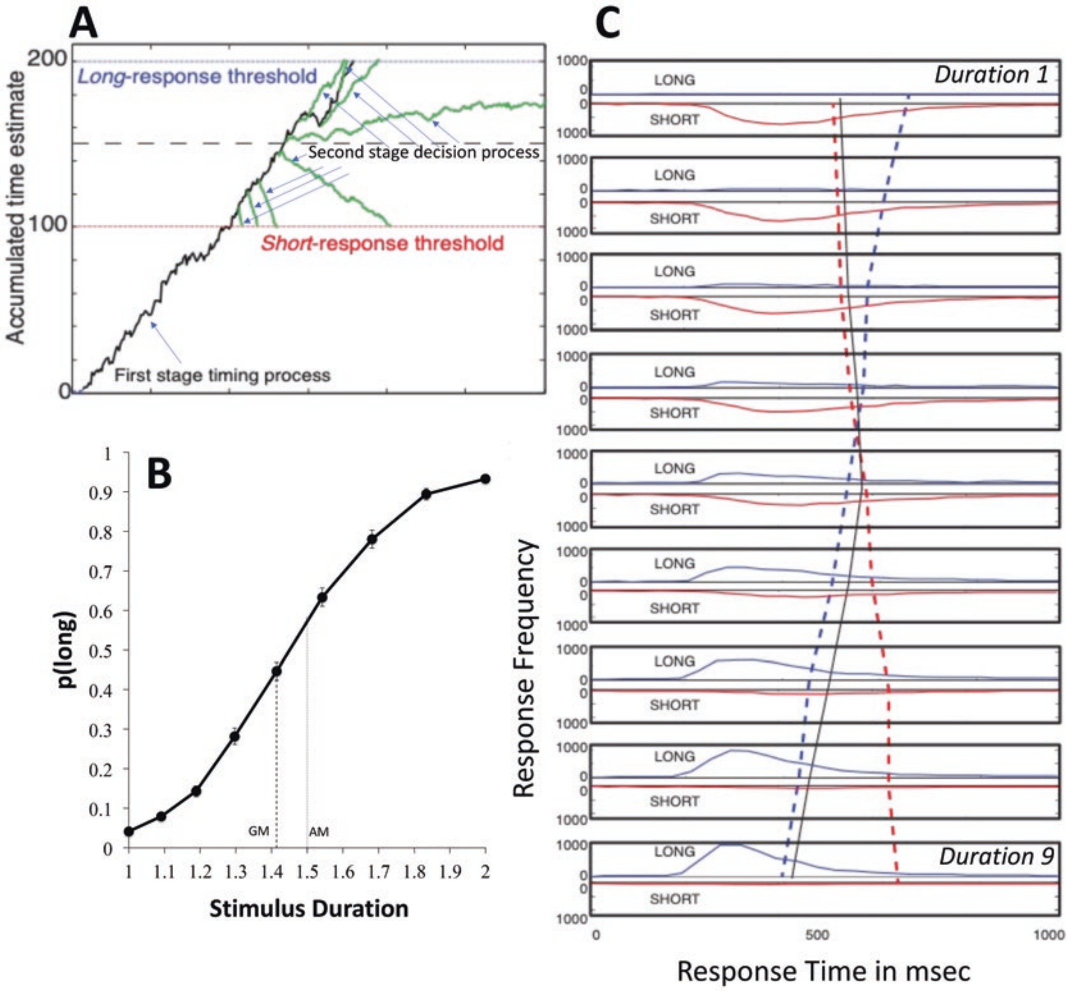


Fig. 8 (**A**) The timer (TopDDM) tracks time up to the duration of the long reference duration. The decision process begins when a stimulus duration ends. Sample decision processes are shown (each for a different stimulus interval). The starting point of the decision process corresponds to the location of the timer process at the end of the stimuli. Resultantly, the decision process starts at different locations depending on where the first process is at the end of the stimulus duration. Drift is toward the long threshold if the timer location exceeds the level of subjective equality (black horizonral dashed line). (**B**) Mean proportion of long choices as a function of the stimulus duration. AM: arithmetic mean, GM: geometric mean. (**C**) Response time predictions shown as dashed lines along with relative densities of short and long categorization judgments. The solid lines show the response time predictions for all judgments

accumulator models; with a smaller learning rate, it can adapt more gradually.

Retrospective (Episodic) Timing

All the models explained above target prospective timing. However, a large portion of our daily timing involves what is called retrospective timing. Unlike prospective timing, retrospective timing is treated as an automatic process based on implicit temporal associations between events and part of the episodic memory system (MacDonald, 2014). Similar to the place cell architecture with respect to the spatial representations (hippocampal neurons that fire when a particular location is occupied—O'Keefe & Nadel, 1978), Howard Eichenbaum and colleagues showed that a group of hippocampal neurons (now known as time cells) fire successively at successive moments in temporally structured events independent of the time-dependent behaviors and spatial information (Eichenbaum, 2014; MacDonald et al., 2011; Pastalkova et al., 2008—but also see Tiganj et al. (2017) for similar observations in the rodent medial prefrontal cortex (MPC) based on the reanalysis of data from Kim et al., 2013; and see Mello et al. (2015) and Akhlaghpour et al. (2016) for similar observations in the striatum).

Time cells are hypothesized to underlie the temporal/ordinal organization of retrospective memory representations by assigning “time stamps” or providing temporal context for memories (e.g., MacDonald et al., 2013; Manns et al., 2007) and even the temporal organization of classical conditioning (e.g., trace conditioning that requires hippocampal involvement—for review, MacDonald, 2014 but see Ahmed et al., 2020). Raster plots of spiking activity of time cells show that different cells fire maximally at different periods during the interval, and the spiking window of the later-peaking neurons is wider than those neurons that fire during the earlier epochs (Fig. 9)—they rescale the width of these high-activity windows according to the duration of the events. The wider window of spiking of neurons later in the chain may derive from slower signal propagation as a function of time (Kraus et al., 2013; Howard et al., 2014; Mello et al., 2015; Salz et al., 2016). This is effectively equivalent to the lower discriminability of longer time intervals (via more overlap in the spiking windows for different durations) or the representation of shorter intervals with more units compared to longer intervals (see Fig. 5—lower panel). This architecture would in turn lead to lower temporal precision by which events are reconstructed and has lower precision for longer times. Note that the sequential propagation of activity parallels the sequential activity in BET, including the rescaling property, but the signal propagation speed is assumed to be constant in BET (not matching the behavior of time cells).

Zeki and Balci (2019) proposed a model of time cells that was composed of a single layer of cells with integrate-and-fire characteristics feed-forward excitatory connections. An inhibitory cell was assumed to provide overall inhibition to time cells with each activation of a time cell, which increased with time. In other words, the inhibition is proportional to total time cell activ-

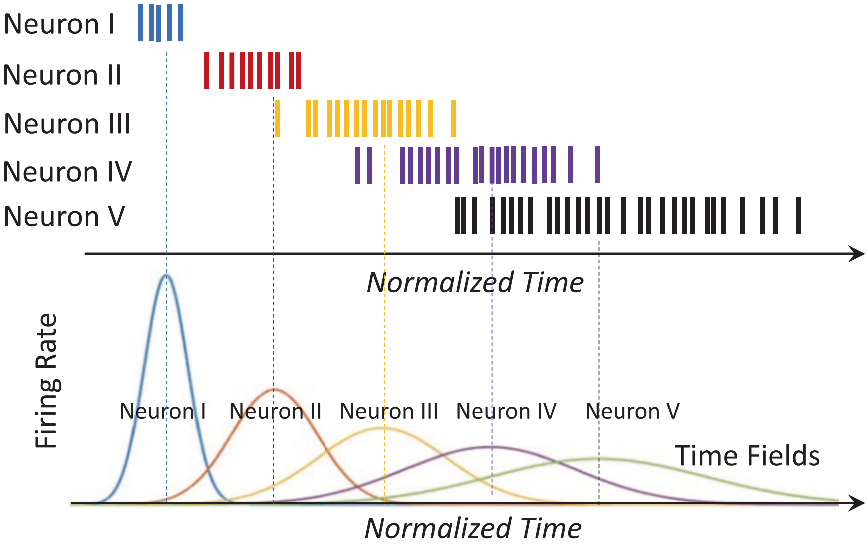


Fig. 9 Illustration of raster plots of temporally modulated spiking (top) and the resultant time fields of the corresponding neurons

ity, which as described earlier has resemblances to the core assumptions of the lowest-level of the TopDDM/SRT/Poisson-spiking hierarchy in Simen et al. (2011a). Thus, the same principle applies in two different cases, and natural selection will likely lead to the reuse of a successful approach. This increasing inhibition delays the firing of time cells later in the chain. Importantly, this prediction was shown to be robust to the variation in the key model parameters. Zeki and Balci (2023) recently showed that similar dynamics (including temporal scaling) can be generated by integrate-and-fire neurons with slow after-hyperpolarization currents and varying resting membrane potentials and without a need for a specialized network architecture (e.g., synaptic connections between time cells) other than self-excitation. This model relies on the fact that neurons with lower resting potentials will fire later in time and exhibit a wider activity window (e.g., right-shifted and broader tuning function) compared to neurons with higher resting potential. The initiation and termination of neurons' firing with self-excitation result from the initial decay and then reactivation of slow after-hyperpolarizing current (activated after multiple action potentials), respectively.

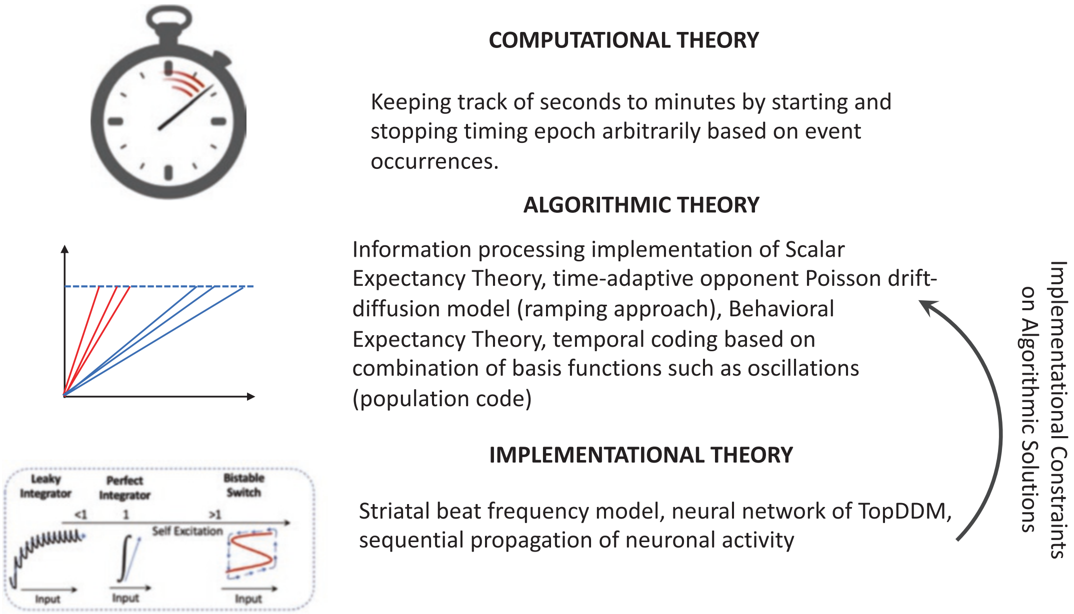
Other models have relied on exponentially decaying firing rates based on the dynamics of calcium-activated nonspecific cationic (CAN) currents using an integrate-and-fire model (Tiganj et al., 2015). In this model, the calcium clearance mechanisms (active at all times) and voltage-gated calcium currents (present during action potentials) determine the CAN. In other words, with each action potential, a fixed calcium concentration is added (resulting in calcium concentration above zero at the outset due to stimulus-driven activity) while otherwise, the change in calcium at each time point is dependent on the clearance of calcium, which depends on the calcium concentration in the cell. The latter property results in an exponential decay of the intracellular calcium concentration with time constants up to the minute range. Importantly, Tiganj et al. (2015) have also shown that the time constants of decay can be externally controlled in multiple ways (e.g., calcium clearance, the maximum conductivity of the CAN current channels, the amount of calcium influx during action potentials, and the charge needed to cause each spike).[1] Within this framework, it is not only the firing rate but also the identity of the active neuron that provides temporal information, similar to place cells (Howard et al., 2014). It is the rescalability feature of these models that puts them under time adaptive models of interval timing with the exception that the clock signals are not integrated in a linear but a negatively accelerating fashion.

Note that decay functions were used earlier in the Multiple Time Scales Model that uses a series of leaky integrators with decay functions according to the power law (Staddon & Higa, 1996, 1999—see also Temporal Context Model of Shankar & Howard, 2010). One of the advantages of using decay functions and thus the strength of memory traces as a proxy for time is that temporal representation becomes an emergent property of memory processes rather than requiring a dedicated timing mechanism. This kind of timing process may be primarily applicable to retrospective, but not prospective, timing since the strength of the memory trace is the proxy for time.

Finally, several models rely on repeatable stochastic processes, namely the neuronal activity propagation pathways in stochastically connected noisy neurons (Fig. 11) to encode and remember time intervals (e.g., Ahrens & Sahani, 2008; Buonomano, 2005; Buonomano & Merzenich, 1995; Haß et al., 2008; Karmarkar & Buonomano, 2007). The primary idea of these models is that when the initial unit in this chain is activated, its likelihood of following the same path will be higher than the likelihood of following different paths of activity propagation. One can liken this to the likelihood of people following earlier paths

---

[1] See Durstewitz D. (2003) ramping model of neuronal timing based on calcium dynamics. According to this model, the integration of input depends on the opening of the voltage-gated Ca2+ channels by the tonic synaptic input. The opening of Ca2+ activates depolarizing currents and generates a linear ramp in firing rate based on the resulting positive feedback (Fig. 10).



COMPUTATIONAL THEORY

Keeping track of seconds to minutes by starting and stopping timing epoch arbitrarily based on event occurrences.

ALGORITHMIC THEORY

Information processing implementation of Scalar Expectancy Theory, time-adaptive opponent Poisson drift-diffusion model (ramping approach), Behavioral Expectancy Theory, temporal coding based on combination of basis functions such as oscillations (population code)

IMPLEMENTATIONAL THEORY

Striatal beat frequency model, neural network of TopDDM, sequential propagation of neuronal activity

Fig. 10 Interval timing conceptualized at different levels of analysis (Marr, 1982)

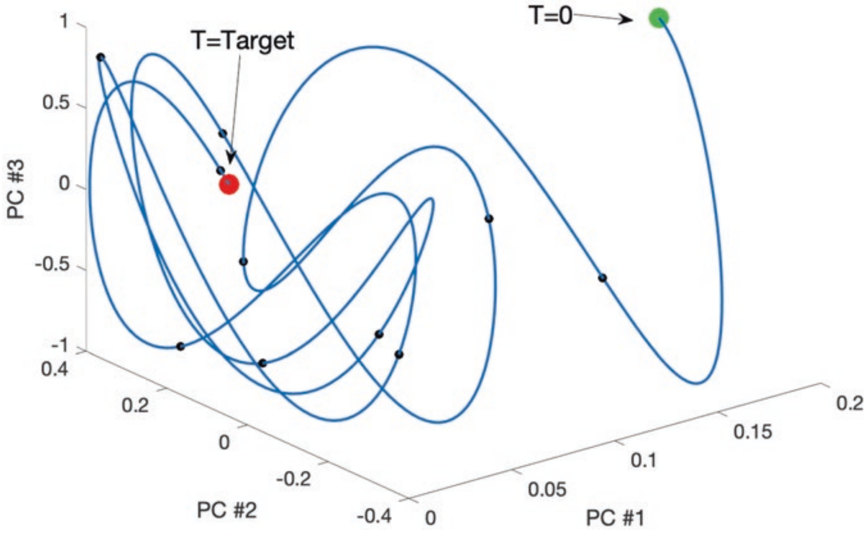


Fig. 11 Depiction of neural trajectory model of interval timing. With the onset of an event, the activation has a higher likelihood of following the same trajectory and the point on the trajectory codes for elapsed time

on a lawn as a detour (following a degraded lawn) to minimize the path length between the origin and the destination.

Relatedly, other researchers showed that they can train recurrent neural networks to encode and remember temporal information (e.g., Bi & Zhou, 2020). Specifically, they report stereotypical dynamic trajectories where temporal information is encoded by the state evolution along this stereotypical trajectory. Furthermore, these dynamics were shown to have the property of temporal scaling at least while producing the interval (e.g., stretched for longer intervals—referred to above). Pérez and Merchant (2018) showed that time-dependent properties of recurrent neural networks (e.g., synaptic properties) can result in time-selective activity in a fashion that can account for the statistical properties of interval timing. Finally, Hardy and Buonomano (2018) showed that recurrent neural networks can generate the experimentally observed sequential activity (feed-forward trajectories) up to several seconds (i.e., 5 s) based on the balance between excitation and inhibition and a supervised learning rule. The state of the activity in neural networks can be read out as elapsed time. Importantly, this model also produces scalar variability.

These models of interval timing can also be categorized as ramping models in which individual neurons keep track of time by adapting the ramping rate (e.g., TopDDM) and as population-clock models in which temporal representation results from the activity of multiple neurons, in which a distinct pattern that emerges from the population codes for different time intervals (e.g., SBF—see De Corte et al., 2022 for a detailed discussion of this alternative categorization). Population clock models require neurons to exhibit the same firing profile in each trial for their population-level activity patterns to be informative with respect to time (not in and of itself necessarily coding for time and thus its metric properties). This, in turn, requires the use of global signals to reset the activity state of individual units at the beginning of a timing epoch (e.g., the start-gun signal in the SBF). It is this repeatability (reliability) of each unit's activity pattern that results in being able to read a unique pattern at the population level. For instance, if individual instruments start playing their part at random times without following the conductor's start signal, the resulting music will not represent the desired and pleasant composition to listen to. SBF and, to a certain extent, repeatable stochastic process models and time-cell functional architecture fall within the category of neural population-based coding of time intervals.

Another issue relates to empirical data that indicate animals' ability to make computations based on time representations (i.e., affordances of temporal representations). For example, one informative type of empirical phenomenon that has been analyzed by De Corte et al. (2022) is findings related to temporal averaging. Specifically, when rats are trained to expect rewards at different delays after the onset of auditory vs. visual stimulus and their timing behavior is clustered around the average of these delays when presented with the compound stimulus (e.g., audiovisual—e.g., Delamater & Nicolas, 2015; De Corte & Matell, 2016; Swanton et al., 2009). These findings suggest that nonhuman animals can make swift computations based on previously established time representations. De Corte (2021) demonstrated that this interesting phenomenon could be explained within the TopDDM framework by assuming that ramping units increase their firing rates at intermediate levels (intermediate to the firing rates for short and long intervals signaled by two different stimuli) during compound trials. This would occur, for instance, when different cue units associated with different intervals compete over the number of layer 2 switch units (De Corte, 2021). These findings are harder to account for with population codes since there is no apparent mechanism that would lead to the control of the propagation of neuronal activity based on the assumed representational architecture. Similar computations are required to be able to account for Ralph Miller and colleagues' temporal map results (e.g., Barnet & Miller, 1996; Molet et al., 2012). Note that population code read out of time intervals does not preserve the metric structure of temporal information, which makes them ineligible for arithmetic computations on temporal representations. We refer the reader to a study by De Corte et al. (2022) for an excellent discussion of these theoretical issues that relate to the conceptual analysis of timing models.

Conclusion

In this chapter, we tried to cover as much as possible the (neuro)computational approaches to modeling interval timing and categorize them as threshold vs. clock-adaptive models as well as dedicated clock/ramping vs. emergent time/population code models. We hope that the coverage of these models and their conceptual categorization will help guide timing researchers in their modeling efforts as well as in interpreting their findings considering these models. Finally, we would like to remind the reader that “All models are wrong, but some are useful” (George Box, British statistician) no matter how much modelers tend to equate empirical reality with their models. Thus, in principle, we, as timing modelers, are all elegantly wrong.

References

Abner, R. T., Edwards, T., Douglas, A., & Brunner, D. (2001). Pharmacology of temporal cognition in two mouse strains. *International Journal of Comparative Psychology, 14*(3–4), 189–210.

Addyman, C., French, R. M., & Thomas, E. (2016). Computational models of interval timing. *Current Opinion in Behavioral Sciences, 8*, 140–146. https://doi.org/10.1016/j.cobeha.2016.01.004

Ahmed, M. S., Priestley, J. B., Castro, A., Stefanini, F., Solis Canales, A. S., Balough, E. M., Lavoie, E., Mazzucato, L., Fusi, S., & Losonczy, A. (2020). Hippocampal network reorganization underlies the formation of a temporal association memory. *Neuron, 107*(2), 283.

Ahrens, M. B., & Sahani, M. (2008). Inferring elapsed time from stochastic neural processes. *Advances in Neural Inform Processing, 20*, 1–8.

Akbiyik, S., Göksun, T., & Balcı, F. (2022). Elucidating the common basis for task-dependent differential manifestations of category advantage: A decision theoretic approach. *Cognitive science, 46*(1), e13078. https://doi.org/10.1111/cogs.13078

Akdogan, B., & Balci, F. (2016). Stimulus probability effects on temporal bisection performance of mice (Mus musculus). *Animal Cognition, 19*, 15–30. https://doi.org/10.1007/s10071-015-0909-6

Akdoğan, B., & Balcı, F. (2017). Are you early or late?: Temporal error monitoring. *Journal of Experimental Psychology: General, 146*(3), 347–361. https://doi.org/10.1037/xge0000265

Akhlaghpour, H., Wiskerke, J., Choi, J. Y., Taliaferro, J. P., Au, J., & Witten, I. B. (2016). Dissociated sequential activity and stimulus encoding in the dorsomedial striatum during spatial working memory. *eLife, 5*, e19507. https://doi.org/10.7554/eLife.19507

Augustine, S. (2008). *The confessions* (H. Chadwick, Trans.). Oxford University Press.

Balcı, F. (2014). Interval timing, dopamine, and motivation. *Timing & Time Perception, 2*(3), 379–410. https://doi.org/10.1163/22134468-00002035

Balcı, F. (2022). Tracing the shadow of time. *Proceedings of the National Academy of Sciences of the United States of America, 119*(10), e2201001119. https://doi.org/10.1073/pnas.2201001119

Balcı, F., & Freestone, D. (2020). The Peak interval procedure in rodents: A tool for studying the neurobiological basis of interval timing and its alterations in models of human disease. *Bio-protocol, 10*(17), e3735. https://doi.org/10.21769/BioProtoc.3735

Balcı, F., & Simen, P. (2014). Decision processes in temporal discrimination. *Acta psychologica, 149*, 157–168. https://doi.org/10.1016/j.actpsy.2014.03.005

Balci, F., Gallistel, C. R., Allen, B. D., Frank, K. M., Gibson, J. M., & Brunner, D. (2009). Acquisition of peak responding: What is learned? *Behavioural Processes, 80*(1), 67–75. https://doi.org/10.1016/j.beproc.2008.09.010

Balcı, F., Ünübol, H., Grondin, S. et al. (2023). Dynamics of retrospective timing: A big data approach. *Psychon Bull Rev 30*, 1840–1847. https://doi.org/10.3758/s13423-023-02277-3

Barnet, R. C., & Miller, R. R. (1996). Temporal encoding as a determinant of inhibitory control. *Learning and Motivation, 27*, 73–91.

Bevins, R. A., & Ayres, J. J. B. (1995). One-trial context fear conditioning as a function of the interstimulus interval. *Animal Learning & Behavior, 23*(4), 400–410. https://doi.org/10.3758/BF03198940

Bi, Z., & Zhou, C. (2020). Understanding the computation of time using neural network models. *Proceedings of the National Academy of Sciences of the United States of America, 117*(19), 10530–10540. https://doi.org/10.1073/pnas.1921609117

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological review, 113*(4), 700–765. https://doi.org/10.1037/0033-295X.113.4.700

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57*(3), 153–178.

Buhusi, C. V., & Oprisan, S. A. (2013). Time-scale invariance as an emergent property in a perceptron with realistic, noisy neurons. *Behavioural Processes, 95*, 60–70. https://doi.org/10.1016/j.beproc.2013.02.015

Buonomano, D. V. (2005). A learning rule for the emergence of stable dynamics and timing inrecurrent networks. *Journal of Neurophysiology, 94*(4), 2275–2283.

Buonomano, D. V., & Merzenich, M. M. (1995). Temporal information transformed into a spatial code by a neural network with realistic properties. *Science (New York, N.Y.), 267*(5200), 1028–1030. https://doi.org/10.1126/science.7863330

Çevik, M. O. (2003). Effects of methamphetamine on duration discrimination. *Behavioral Neuroscience, 117*(4), 774–784.

Church, R. M., & Broadbent, H. A. (1990). Alternative representations of time, number, and rate. *Cognition, 37*, 55–81.

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology: Animal Behavior Processes, 3*(3), 216–228. https://doi.org/10.1037/0097-7403.3.3.216

Creelman, C. D. (1962). Human discrimination of auditory duration. *Journal of the Acoustical Society of America, 34*(5), 582–593. https://doi.org/10.1121/1.1918172

De Corte, B. J. (2021). *What are the neural mechanisms of "Higher-Order" timing? Complex behavior from low-level circuits*. ProQuest Dissertations Publishing. https://doi.org/10.17077/etd.006279

De Corte, B. J., & Matell, M. S. (2016). Temporal averaging across multiple response options: Insight into the mechanisms underlying integration. *Animal Cognition, 19*, 329–342. https://doi.org/10.1007/s10071-015-0935-4

De Corte, B. J., Akdoğan, B., & Balsam, P. D. (2022). Temporal scaling and computing time in neural circuits: Should we stop watching the clock and look for its gears? *Frontiers in Behavioural Neuroscience, 16*, 1022713. https://doi.org/10.3389/fnbeh.2022.1022713

Delamater, A. R., & Nicolas, D.-M. (2015). Temporal averaging across stimuli signaling the same or different reinforcing outcomes in the peak procedure. *International Journal Comparative Psychology, 28*, uclapsych_ijcp_28552.

Drew, M. R., Fairhurst, S., Malapani, C., Horvitz, J. C., & Balsam, P. D. (2003). Effects of dopamine antagonists on the timing of two intervals. *Pharmacology, Biochemistry, and Behavior, 75*(1), 9–15. https://doi.org/10.1016/s0091-3057(03)00036-4

Durstewitz, D. (2003). Self-organizing neural integrator predicts interval times through climbing activity. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 23*(12), 5342–5353. https://doi.org/10.1523/JNEUROSCI.23-12-05342.2003

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. *Nature Reviews Neuroscience, 15*, 732–744. https://doi.org/10.1038/nrn3827

Emmons, E. B., De Corte, B. J., Kim, Y., Parker, K. L., Matell, M. S., & Narayanan, N. S. (2017). Rodent medial frontal control of temporal processing in the dorsomedial striatum. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 37*(36), 8718–8733. https://doi.org/10.1523/JNEUROSCI.1376-17.2017

Emmons, E. B., Kennedy, M., Kim, Y., & Narayanan, N. S. (2019). Corticostriatal stimulation compensates for medial frontal inactivation during interval timing. *Scientific Reports, 9*(1), 14371. https://doi.org/10.1038/s41598-019-50975-7

Emmons, E., Tunes-Chiuffa, G., Choi, J., Bruce, R. A., Weber, M. A., Kim, Y., & Narayanan, N. S. (2020). Temporal learning among prefrontal and striatal ensembles. *Cerebral Cortex Communications, 1*(1), tgaa058. https://doi.org/10.1093/texcom/tgaa058

Farrell, S., & Lewandowsky, S. (2018). *Computational modeling of cognition and behavior*. Cambridge University Press. https://doi.org/10.1017/CBO9781316272503

Ferster, C. B., & Skinner, B. F. (1957). *Schedules of reinforcement*. Appleton-Century-Crofts. https://doi.org/10.1037/10627-000

Gallistel, C. R., & King, A. P. (2010). *Memory and the computational brain: Why cognitive science will transform neuro-science*. Wiley-Blackwell Publishing.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*(3), 279–325. https://doi.org/10.1037/0033-295X.84.3.279

Gibbon, J., & Church, R. M. (1990). Representation of time. *Cognition, 37*(1–2), 23–54. https://doi.org/10.1016/0010-0277(90)90017-E

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. *Annals of the New York Academy of Sciences, 423*, 52–77. https://doi.org/10.1111/j.1749-6632.1984.tb23417.x

Grice, G. R. (1968). Stimulus intensity and response evocation. *Psychological Review, 75*, 359–373. https://doi.org/10.1037/h0026287

Gu, B.-M., van Rijn, H., & Meck, W. H. (2015). Oscillatory multiplexing of neural population codes for interval timing and working memory. *Neuroscience and Biobehavioral Reviews, 48*, 160–185. https://doi.org/10.1016/j.neubiorev.2014.10.008

Hardy, N. F., & Buonomano, D. V. (2018). Encoding time in feedforward trajectories of a recurrent neural network model. *Neural Computation, 30*(2), 378–396. https://doi.org/10.1162/neco_a_01041

Hass, J., & Durstewitz, D. (2014). Neurocomputational models of time perception. *Advances in Experimental Medicine and Biology, 829*, 49–71. https://doi.org/10.1007/978-1-4939-1782-2_4

Hass, J., Blaschke, S., Rammsayer, T., & Herrmann, J. M. (2008). A neurocomputational model for optimal temporal processing. *Journal of Computational Neuroscience, 25*(3), 449–464. https://doi.org/10.1007/s10827-008-0088-4

Howard, M. W., MacDonald, C. J., Tiganj, Z., Shankar, K. H., Du, Q., Hasselmo, M. E., & Eichenbaum, H. (2014). A unified mathematical framework for coding time, space, and sequences in the hippocampal region. *The Journal of Neuroscience, 34*(13), 4692–4707. https://doi.org/10.1523/JNEUROSCI.5808-12.2014

Jazayeri, M., & Shadlen, M. N. (2015). A neural mechanism for sensing and reproducing a time interval. *Current Biology, 25*, 2599–2609. https://doi.org/10.1016/j.cub.2015.08.038

Kalueff, A. V., Stewart, A. M., Song, C., Berridge, K. C., Graybiel, A. M., & Fentress, J. C. (2016). Neurobiology of rodent self-grooming and its value for translational neuroscience. *Nature Reviews Neuroscience, 17*(1), 45–59. https://doi.org/10.1038/nrn.2015.8

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: Encoding time in neural network states. *Neuron, 53*(3), 427–438. https://doi.org/10.1016/j.neuron.2007.01.006

Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological Review, 95*(2), 274–295. https://doi.org/10.1037/0033-295X.95.2.274

Kim, J., Ghim, J. W., Lee, J. H., & Jung, M. W. (2013). Neural correlates of interval timing in rodent prefrontal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33*(34), 13834–13847. https://doi.org/10.1523/JNEUROSCI.1443-13.2013

Komura, Y., Tamura, R., Uwano, T., Nishijo, H., Kaga, K., & Ono, T. (2001). Retrospective and prospective coding for predicted reward in the sensory thalamus. *Nature, 412*, 546–549. https://doi.org/10.1038/35087595

Kononowicz, T. W., & van Rijn, H. (2014). Decoupling interval timing and climbing neural activity: A dissociation between CNV and N1P2 amplitudes. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 34*(8), 2931–2939. https://doi.org/10.1523/JNEUROSCI.2523-13.2014

Kraus, B. J., Robinson, R. J., 2nd, White, J. A., Eichenbaum, H., & Hasselmo, M. E. (2013). Hippocampal "time cells": Time versus path integration. *Neuron, 78*, 1090–1101. https://doi.org/10.1016/j.neuron.2013.04.015

Leon, M. I., & Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron, 38*, 317–327. https://doi.org/10.1016/s0896-6273(03)00185-5

Link, S. W. (1992). *The wave theory of difference and similarity*. Psychology Press.

Luce, R. (1986). *Response times: Their role in inferring elementary mental organization* (Oxford Psychology Series) (Vol. 8). Clarendon (Oxford University Press). 562 pp.

Ludvig, E. A., Conover, K., & Shizgal, P. (2007). The effects of reinforcer magnitude on timing in rats. *Journal of the Experimental Analysis of Behavior, 87*(2), 201–218. https://doi.org/10.1901/jeab.2007.38-06

Luzardo, A., Rivest, F., Alonso, E., & Ludvig, E. A. (2017). A drift–diffusion model of interval timing in the peak procedure. *Journal of Mathematical Psychology, 77*, 111–123. https://doi.org/10.1016/j.jmp.2016.10.002

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

MacDonald, C. J. (2014). Prospective and retrospective duration memory in the hippocampus: Is time in the foreground or background? *Philosophical Transactions of the Royal Society of London. Series B, Biological sciences, 369*(1637), 20120463. https://doi.org/10.1098/rstb.2012.0463

MacDonald, C. J., Lepage, K. Q., Eden, U. T., & Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron, 71*(4), 737–749. https://doi.org/10.1016/j.neuron.2011.07.012

MacDonald, C. J., Carrow, S., Place, R., & Eichenbaum, H. (2013). Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33*(36), 14607–14616. https://doi.org/10.1523/JNEUROSCI.1537-13.2013

Machado, A. (1997). Learning the temporal dynamics of behavior. *Psychological Review, 104*(2), 241–265. https://doi.org/10.1037/0033-295X.104.2.241

Machado, A., Malheiro, M. T., & Erlhagen, W. (2009). Learning to time: A perspective. *Journal of the Experimental Analysis of Behavior, 92*(3), 423–458. https://doi.org/10.1901/jeab.2009.92-423

Manns, J. R., Howard, M. W., & Eichenbaum, H. (2007). Gradual changes in hippocampal activity support remembering the order of events. *Neuron, 56*(3), 530–540. https://doi.org/10.1016/j.neuron.2007.08.017

Maricq, A. V., Roberts, S., & Church, R. M. (1981). Methamphetamine and time estimation. *Journal of Experimental Psychology. Animal Behavior Processes, 7*(1), 18–30. https://doi.org/10.1037/0097-7403.7.1.18

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. W.H. Freeman.

Matell, M. S., & Meck, W. H. (2000). Neuropsychological mechanisms of interval timing behavior. *BioEssays: News and Reviews in Molecular, Cellular and Developmental Biology, 22*(1), 94–103. https://doi.org/10.1002/(SICI)1521-1878(200001)22:1<94::AID-BIES14>3.0.CO;2-E

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cognitive Brain Research, 21*, 139–170. https://doi.org/10.1016/j.cogbrainres.2004.06.012

Matell, M. S., Meck, W. H., & Nicolelis, M. A. (2003). Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behavioral Neuroscience, 117*(4), 760–773. https://doi.org/10.1037/0735-7044.117.4.760

Matell, M. S., Bateson, M., & Meck, W. H. (2006). Single-trials analyses demonstrate that increases in clock speed contribute to the methamphetamine-induced horizontal shifts in peak-interval timing functions. *Psychopharmacology, 188*(2), 201–212. https://doi.org/10.1007/s00213-006-0489-x

Meck, W. H. (1983). Selective adjustment of the speed of internal clock and memory processes. *Journal of Experimental Psychology: Animal Behavior Processes, 9*(2), 171–201. https://doi.org/10.1037/0097-7403.9.2.171

Meck, W. H. (1996). Neuropharmacology of timing and time perception. *Brain Research. Cognitive*

Brain Research, 3(3–4), 227–242. https://doi.org/10.1016/0926-6410(96)00009-2

Meck, W. H., & Church, R. M. (1987). Cholinergic modulation of the content of temporal memory. *Behavioral Neuroscience, 101*(4), 457–464. https://doi.org/10.1037/0735-7044.101.4.457

Meck W. H. (2006). Neuroanatomical localization of an internal clock: a functional link between mesolimbic, nigrostriatal, and mesocortical dopaminergic systems. *Brain research, 1109*(1), 93–107. https://doi.org/10.1016/j.brainres.2006.06.031

Meck, W. H., Penney, T. B., & Pouthas, V. (2008). Corticostriatal representation of time in animals and humans. *Current Opinion in Neurobiology, 18*, 145–152. https://doi.org/10.1016/j.conb.2008.08.002

Mello, G. B. M., Soares, S., & Paton, J. J. (2015). A scalable population code for time in the striatum. *Current Biology, 25*(9), 1113–1122.

Merchant, H., & Averbeck, B. B. (2017). The computational and neural basis of rhythmic timing in medial premotor cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 37*(17), 4552–4564. https://doi.org/10.1523/JNEUROSCI.0367-17.2017

Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Computation, 1*, 359–371. https://doi.org/10.1162/neco.1989.1.3.359

Molet, M., Miguez, G., Cham, H. X., & Miller, R. R. (2012). When does integration of independently acquired temporal relationships take place? *Journal of Experimental Psychology: Animal Behavior Processes, 38*, 369–380. https://doi.org/10.1037/a0029379

Murakami, M., Shteingart, H., Loewenstein, Y., & Mainen, Z. F. (2017). Distinct sources of deterministic and stochastic components of action timing decisions in rodent frontal cortex. *Neuron, 94*, e7908–e7919. https://doi.org/10.1016/j.neuron.2017.04.040

O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. Oxford University Press.

Odum, A. L., & Ward, R. D. (2007). Characterizing the effects of d-amphetamine on temporal discrimination. *Behavioural Processes, 75*(2), 156–166. https://doi.org/10.1016/j.beproc.2007.02.022

Ofir, N., & Landau, A. N. (2022). Neural signatures of evidence accumulation in temporal decisions. *Current Biology: CB, 32*(18), 4093–4100.e6. https://doi.org/10.1016/j.cub.2022.08.006

Okun, M., & Lampl, I. (2009). Balance of excitation and inhibition. *Scholarpedia, 4*(8), 7467., revision #150430. https://doi.org/10.4249/scholarpedia.7467

Oprisan, S. A., & Buhusi, C. V. (2011). Modeling pharmacological clock and memory patterns of interval timing in a striatal beat-frequency model with realistic, noisy neurons. *Frontiers in Integrative Neuroscience, 5*, 52. https://doi.org/10.3389/fnint.2011.00052

Oprisan, S. A., & Buhusi, C. V. (2013). Why noise is useful in functional and neural mechanisms of interval timing? *BMC Neuroscience, 14*, Article 84. https://doi.org/10.1186/1471-2202-14-84

Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsáki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science (New York, N.Y.), 321*(5894), 1322–1327. https://doi.org/10.1126/science.1159775

Pavlov, I. P. (1927). *Conditioned reflexes*. Oxford University Press.

Pérez, O., & Merchant, H. (2018). The synaptic properties of cells define the hallmarks of interval timing in a recurrent neural network. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 38*(17), 4186–4199. https://doi.org/10.1523/JNEUROSCI.2651-17.2018

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59–108. https://doi.org/10.1037/0033-295X.85.2.59

Ratcliff, R. (2014). Measuring psychometric functions with the diffusion model. *Journal of Experimental Psychology: Human Perception and Performance, 40*(2), 870–888. https://doi.org/10.1037/a0034954

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20*(4), 873–922. https://doi.org/10.1162/neco.2008.12-06-420

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences, 20*(4), 260–281. https://doi.org/10.1016/j.tics.2016.01.007

Rivest, F., & Bengio, Y. (2011). Adaptive drift-diffusion process to learn time intervals. Cornell University Library, *arXiv*, 1103.2382v1. https://doi.org/10.48550/arXiv.1103.2382

Roberts, S. (1981). Isolation of an internal clock. *Journal of Experimental Psychology: Animal Behavior Processes, 7*(3), 242–268. https://doi.org/10.1037/0097-7403.7.3.242

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review, 65*(6), 386–408. https://doi.org/10.1037/h0042519

Salz, D. M., Tiganj, Z., Khasnabish, S., Kohley, A., Sheehan, D., Howard, M. W., & Eichenbaum, H. (2016). Time cells in hippocampal area CA3. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 36*(28), 7476–7484. https://doi.org/10.1523/JNEUROSCI.0087-16.2016

Shankar, K. H., & Howard, M. W. (2010). Timing using temporal context. *Brain Research, 1365*, 3–17. https://doi.org/10.1016/j.brainres.2010.07.045

Simen, P., Balci, F., de Souza, L., Cohen, J. D., & Holmes, P. (2011a). A model of interval timing by neural integration. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 31*(25), 9238–9253. https://doi.org/10.1523/JNEUROSCI.3121-10.2011

Simen, P., Balci, F., deSouza, L., Cohen, J. D., & Holmes, P. (2011b). Interval timing by long-range temporal integration. *Frontiers in Integrative Neuroscience, 5*, Article 28. https://doi.org/10.3389/fnint.2011.00028

Simen, P., Rivest, F., Ludvig, E. A., Balci, F., & Killeen, P. (2013). Timescale invariance in the pacemaker-accumulator family of timing models. *Timing & Time Perception, 1*, 159–188. https://doi.org/10.1163/22134468-00002018

Simen, P., Vlasov, K., & Papadakis, S. (2016). Scale (in)variance in a unified diffusion model of decision making and timing. *Psychological Review, 123*(2), 151–181. https://doi.org/10.1037/rev0000014

Staddon, J. E., & Higa, J. J. (1996). Multiple time scales in simple habituation. *Psychological Review, 103*(4), 720–733. https://doi.org/10.1037/0033-295x.103.4.720

Staddon, J. E., & Higa, J. J. (1999). Time and memory: Towards a pacemaker-free theory of interval timing. *Journal of the Experimental Analysis of Behavior, 71*(2), 215–251. https://doi.org/10.1901/jeab.1999.71-215

Sundarapandian, V. (2009). 7. Queueing theory. *Probability, Statistics and Queueing Theory*. PHI Learning. ISBN 978-8120338449.

Swanton, D. N., Gooch, C. M., & Matell, M. S. (2009). Averaging of temporal memories by rats. *Journal of Experimental Psychology: Animal Behavior Processes, 35*, 434–439. https://doi.org/10.1037/a0014021

Tiganj, Z., Hasselmo, M. E., & Howard, M. W. (2015). A simple biophysically plausible model for long time constants in single neurons. *Hippocampus, 25*(1), 27–37. https://doi.org/10.1002/hipo.22347

Tiganj, Z., Jung, M. W., Kim, J., & Howard, M. W. (2017). Sequential firing codes for time in rodent medial prefrontal cortex. *Cerebral cortex (New York, N.Y. : 1991), 27*(12), 5663–5671. https://doi.org/10.1093/cercor/bhw336

Treisman, M. (1963). Temporal discrimination and the indifference interval: Implications for a model of the "internal clock". *Psychological Monographs: General and Applied, 77*(13), 1–31. https://doi.org/10.1037/h0093864

Wald, A. (1947). *Sequential analysis*. Wiley.

Wald, A., & Wolfowitz, J. (1940). On a test whether two samples are from the same population. *The Annals of Mathematical Statistics, 11*(2), 147–162. https://doi.org/10.1214/aoms/1177731909

Wiener, M., Parikh, A., Krakow, A., et al. (2018). an intrinsic role of beta oscillations in memory for time estimation. *Scientific Reports, 8*, 7992. https://doi.org/10.1038/s41598-018-26385-6

Zeki, M., & Balcı, F. (2019). A simplified model of communication between time cells: Accounting for the linearly increasing timing imprecision. *Frontiers in Computational Neuroscience, 12*, 111. https://doi.org/10.3389/fncom.2018.00111

Zeki, M., & Balcı, F. (2023). Modeling time cell neuron-level dynamics. *Computational Brain & Behavior*. https://doi.org/10.1007/s42113-023-00183-4

Part II

Neurophysiology of Timing

Neural Sequences and the Encoding of Time

Saray Soldado-Magraner
and Dean V. Buonomano

Abstract

Converging experimental and computational evidence indicate that on the scale of seconds the brain encodes time through changing patterns of neural activity. Experimentally, two general forms of neural dynamic regimes that can encode time have been observed: neural population clocks and ramping activity. Neural population clocks provide a high-dimensional code to generate complex spatiotemporal output patterns, in which each neuron exhibits a nonlinear temporal profile. A prototypical example of neural population clocks are neural sequences, which have been observed across species, brain areas, and behavioral paradigms. Additionally, neural sequences emerge in artificial neural networks trained to solve time-dependent tasks. Here, we examine the role of neural sequences in the encoding of time, and how they may emerge in a biologically plausible manner. We conclude that neural sequences may represent a canonical computational regime to perform temporal computations.

Keywords

Timing · Neural sequences · Ramping · Temporal processing · Neural population clocks

It can be said that one of the brain's main functions is to predict the future. Memory, for example, is biologically adaptive only to the extent that previous experiences and information are useful to better respond to future events. Indeed, the importance of prediction is highlighted in many theories of brain function (Helmholtz, 1860; Mumford, 1992; Rao & Ballard, 1999; Friston, 2005). Timing, of course, is a key element of prediction, as it is generally not sufficient to only predict what will happen, but when it will happen. Indeed, animals must make predictions across time scales, from anticipating the position of a moving object or the presence of predators or prey, to tracking the circadian and seasonal changes in their physical environment. In addition to the importance of timing to prediction, timing is a critical component of sensory and motor processing. A prominent example is the generation of speech and music, which exemplify the brains' sophisticated ability to discriminate

---

S. Soldado-Magraner
Department of Neurobiology, University of California, Los Angeles, Los Angeles, CA, USA

D. V. Buonomano (✉)
Department of Neurobiology, University of California, Los Angeles, Los Angeles, CA, USA

Department of Psychology, University of California, Los Angeles, Los Angeles, CA, USA
e-mail: dbuono@ucla.edu

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_5

and generate complex spatiotemporal patterns. To cope with the need to produce temporally accurate predictions, parse the temporal structure of sensory events, and generate appropriately timed motor outputs and behaviors, the brain evolved a range of distinct mechanisms to track time across scales ranging from milliseconds to days and beyond (Buonomano, 2007). Here, we focus primarily on the neural mechanisms underlying the ability to track time and generate timed motor outputs on the time scale of a few seconds. It is on this scale that timing is most sophisticated, for it is on the scale of seconds that we can recognize and generate temporal patterns at multiple speeds. Our ability to tell time and discern temporal patterns on the scale of approximately a second underlies speech and music processing. Outside the range of hundreds of milliseconds to a few seconds, the spatiotemporal patterns that comprise music and speech cease to be recognized as music or speech.

Theories of the Neural Mechanisms of Timing

Early theories of timing on the scale of seconds focused primarily on oscillator-based models, including pacemaker/accumulator models in which “tics” of an oscillator are integrated over time, providing a linear metric of elapsed time (Creelman, 1962; Gibbon, 1977; Killeen & Fetterman, 1988; Treisman et al., 1990). Later oscillator-based models included those in which multiple oscillators operating at different frequencies provided a readout of time by detecting the beats or relative phases of the population of oscillators (Miall, 1989; Matell et al., 2003; Buhusi & Meck, 2005; Buhusi et al., 2016). Over the past few decades, there has been a shift toward nonoscillator-based models, most notably those in which time is encoded in time-dependent changes in the firing rate of neurons. These models include ramping and neural population clocks (Ivry & Spencer, 2004; Mauk & Buonomano, 2004; Merchant et al., 2013; Balcı & Simen, 2016; Paton & Buonomano, 2018; Issa et al., 2020). Generally speaking, ramping models are based on a neural integrator which produces linear or monotonic changes in firing rate, thus providing a linear or nonlinear metric of elapsed time (Durstewitz, 2003; Simen et al., 2011; Balcı & Simen, 2016). Neural population clock models explicitly propose that neural circuits encode time through changing patterns of neural activity imposed by the recurrent connectivity of the circuits (Buonomano & Mauk, 1994; Mauk & Donegan, 1997; Buonomano & Laje, 2010). While both models rely on time-dependent changes in firing rate, in ramping models, a single neuron can, at least in principle, provide a continuous and linear metric of elapsed time. Whereas, in neural population clocks, a large population of neurons is, by definition, necessary to encode a continuous window of time.

In contrast to ramping models, the temporal profile of the firing rate of neurons in population clocks models is nonmonotonic. In part as a consequence of this, a further distinction between ramping and neural population clocks models pertains to the dimensionality of the neural dynamics of the population of neurons. Neural population clocks comprise a high-dimensional neural system, whereas ramping models are low-dimensional. That is, in contrast to a population of neurons that are ramping together and thus are highly correlated, the pairwise correlation between neurons within a neural population clock is relatively low.

Depending on the temporal profile of their firing rate, and the “sparsity” of the neural dynamics (how many neurons fire at the same moment), neural population clocks encompass a number of different dynamic regimes including complex patterns and neural sequences (Paton & Buonomano, 2018). Here, we refer to complex neural population clocks as those in which a single neuron can have multiple peak firing rates, and neural sequences as dynamic regimes in which each neuron has a single peak (often captured by a Gaussian-like increase and decrease in firing rate) in which these peaks approximately tile the time window (see Fig. 1). We stress, how-

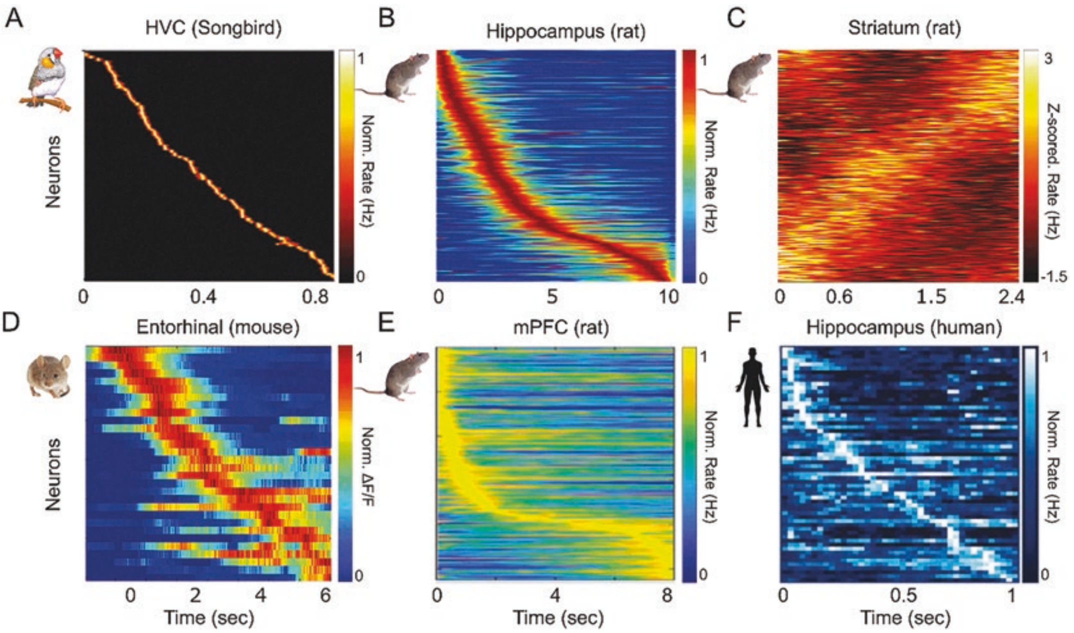


Fig. 1 **Neural sequences are ubiquitous across species and brain regions**. (**a**) Neurons in the premotor nucleus (HVC) of songbirds fire very sparsely while the animal is engaged in song vocalization, tiling the whole time span of the song (Lynch et al., 2016). (**b**) Neural sequences emerge in the hippocampus of rats during an interval discrimination task. Importantly, the sequences are scaled upon different interval durations (Shimbo et al., 2021). (**c**) The speed of striatal sequential dynamics of rats correlates with their subjective perception of the passage of time (Gouvea et al., 2015). (**d**) The entorhinal cortex of mice encodes elapsed time in the form of neural sequences during immobility (Heys & Dombeck, 2018). (**e**) Time can robustly be decoded in the medial prefrontal cortex of rats engaged in a delayed-match-to-sample task (Ning et al., 2022). (**f**) Recently, neural sequences that encode time have also been reported in humans performing an episodic memory task (Umbach et al., 2020)

ever, that the term *neural sequence* is often used in a broad and not carefully defined manner. Sometimes, the term *neural sequence* is used interchangeably with the term *synfire chain*. Here, we distinguish between these terms, because *synfire chain* was originally coined in reference to the sequential, but temporally discrete activity produced by synchronized “bursts” of activity that drive propagation within a feedforward network (Abeles, 1991; Diesmann et al., 1999; Gewaltig et al., 2001).

Below, we will focus primarily on the experimental and computational evidence that in many cases the brain relies on neural sequences to encode time on the scale of hundreds of milliseconds to many seconds, and discuss the potential mechanisms underlying the emergence of neural circuits that support the generation of multiple neural sequences.

Neural Sequences In Vivo

One of the earliest observations of neural sequences comes from the study of neural correlates of learned vocalization in songbirds (Hahnloser et al., 2002). Neurons in the premotor nucleus HVC burst at specific times during bird song production and tile the whole duration of the song, providing a very sparse timing code that drives downstream motor neurons responsible for sound generation (Fig. 1a). Subsequent studies provided a causal link between the dynamics underlying the neural sequences in HVC and song timing, by showing that cooling HVC, but not the motor nucleus, slowed song speed (Long & Fee, 2008).

One of the first observations of neural sequences in mammals was in the rat hippocampus (Pastalkova et al., 2008). Sequential activation of neurons had been observed extensively in the context of navigation as a result of the progressive activation of place cells as an animal traverses a spatial location. Pastalkova et al. (2008) demonstrated that in the same manner that the hippocampus encodes a cognitive map for space, it also contains information about elapsed time. Importantly, in contrast to the sequential activation of place cells, this "temporal map" was internally generated as it emerged in the absence of external cues—while the animal ran in place on a running wheel. Subsequent studies have consistently reported the presence of neural sequences that encode time in the hippocampus (MacDonald et al., 2011, 2013; Kraus et al., 2013; Taxidis et al., 2020; Umbach et al., 2020; Wang et al., 2020; Reddy et al., 2021; Shimbo et al., 2021; Ning et al., 2022) (Fig. 2b)—the cells that participate in these neural sequences and that have a characteristic time receptive field are sometimes referred to as "time cells" (MacDonald et al., 2011). Critically, neural sequences have also been observed in many brain regions other than the hippocampus (Fig. 1c–e), including parietal and motor cortices (Crowe et al., 2010; Harvey et al., 2012; Adler et al., 2019; Zhou et al., 2020), primary visual cortex (Xu et al., 2012; Gavornik & Bear, 2014), entorhinal cortex (Heys & Dombeck, 2018), prefrontal cortex (Tiganj et al., 2017; Ning et al., 2022), and striatum (Gouvea et al., 2015; Mello et al., 2015; Zhou et al., 2020).

The universality of neural sequences is further emphasized by the fact that they have been found across multiple species: songbirds (Hahnloser et al., 2002; Lynch et al., 2016), mice (Gavornik & Bear, 2014; Heys & Dombeck, 2018; Taxidis et al., 2020), rats (Pastalkova et al., 2008; MacDonald et al., 2011, 2013; Xu et al., 2012; Gouvea et al., 2015; Mello et al., 2015; Tiganj et al., 2017; Adler et al., 2019; Wang et al., 2020; Shimbo et al., 2021; Ning et al., 2022), monkeys (Crowe et al., 2010, 2014; Merchant et al., 2015; Gámez et al., 2019), and more recently humans (Umbach et al., 2020; Reddy et al., 2021) (see Fig. 1).

Neural sequences not only encode the passage of time, but can contain information about the context of how a particular behavioral event unfolds in time. Specifically, different neural sequences can be triggered by different stimuli or during different contexts (Pastalkova et al., 2008; MacDonald et al., 2013; Taxidis et al., 2020). For example, different neural sequences are triggered upon presentation of different odors in odor discrimination tasks and working memory tasks (MacDonald et al., 2013; Taxidis et al., 2020). Importantly, in these cases, a given neuron can participate in both neural sequences but at different moments in time—this observation is most consistent with recurrent circuit architectures as opposed to feed-forward circuits. Relative spatial location is also encoded by different neural sequences during object construction tasks (Crowe et al., 2010), and navigation-choice tasks (Harvey et al., 2012). This indicates that a fundamental role of neural sequences may be to bind temporal passage of time to relevant behavioral variables. However, neural sequences have also been found during immobility and tile the whole trial, indicating that internal states may be correlated with the pure passage of time (MacDonald et al., 2013; Lynch et al., 2016; Heys & Dombeck, 2018; Reddy et al., 2021). In fact, the speed of a neural sequence correlates with the subject perception of the passage of time (Gouvea et al., 2015).

To date, the neurons participating in neural time sequences seem to be anatomically intermingled, i.e., neurons that fire close in time during the sequence do not seem to be closer in space (Harvey et al., 2012; Heys & Dombeck, 2018). Additionally, the cross-trial precision of the time field decreases with time (Pastalkova et al., 2008; MacDonald et al., 2011; Tiganj et al., 2017; Umbach et al., 2020), an observation consistent with the robust psychophysical finding that the standard deviation of timed motor responses increases linearly with time (Gibbon, 1977; Merchant et al., 2008; Laje et al., 2011; Grondin, 2014)—Weber's law.

An important feature of motor behavior is the ability to produce well-timed motor patterns at

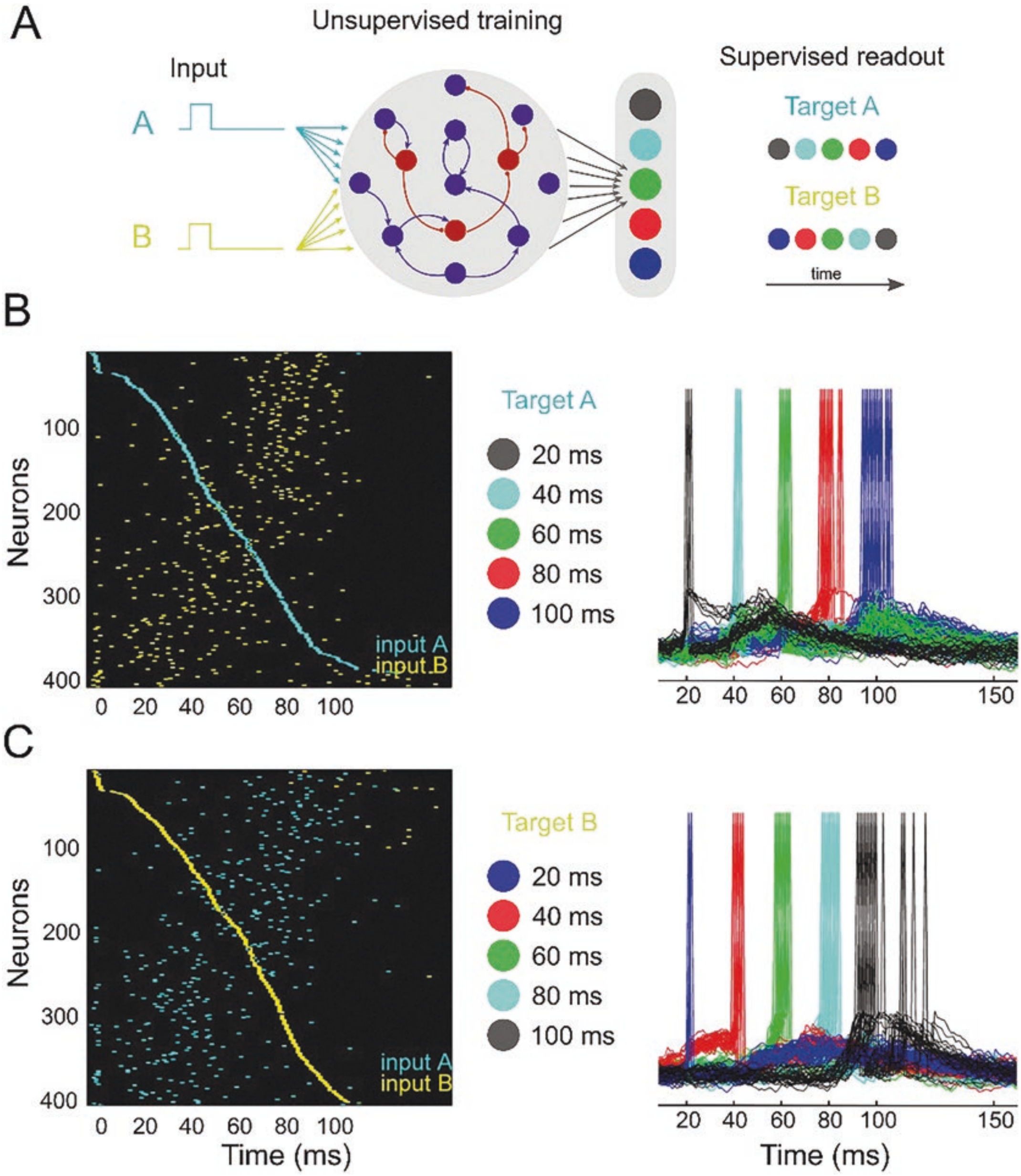


Fig. 2 **Neural sequences emerge in a self-organized manner guided by homeostatic plasticity rules**. (**a**) Schematic of a recurrent neural network with spiking excitatory and inhibitory units. The strength of excitatory synapses is governed by presynaptic-dependent homeostatic (PSD) learning rule. In this example, the recurrent circuit can receive two different inputs, each of which should drive a different sequence of activity in the same population of output units (right). (**b–c**) During unsupervised learning, different neural sequences emerge in the recurrent neural network. After this unsupervised learning stage, the connections onto the 5 output neurons are trained using a simple supervised learning rule. In response to Input A, the five output neurons are trained to fire into a specific time and order (Target A). The order is changed for Input B (Target B). After both stages of training, the different input patterns trigger robust neural sequences that tile the whole trial duration. Importantly, the sequences are independent from one another as can be observed when sorting the neurons of Input A based on Input B and vice versa. The different sequences can drive the five output neurons to fire reliably at different points in time and at a different order depending on the input pattern. (Adapted from Liu and Buonomano (2009))

varying speeds—e.g., such as playing a musical piece at different tempos, or tapping Morse code at speeds 10 or 20 words per minute (Hardy et al., 2018; Wang et al., 2018; Slayton et al., 2020). Thus, if neural sequences underlie motor timing, they should be able to undergo temporal scaling. Indeed, a number of studies have shown that depending on context the same neural sequence can unfold at different speeds (Mello et al., 2015; Wang et al., 2018; Zhou et al., 2020; Shimbo et al., 2021)—effectively implementing a mechanism for temporal scaling. Relatedly, temporal scaling can also be achieved by changing the amplitude of the neural trajectory in state space traveled by the neural sequence. For example, in monkeys trained on a rhythmic timing task, the sequential activation of neurons during each beat evolves in state space with circular dynamics, forming a regenerating loop that resets for every produced interval (Crowe et al., 2014; Merchant et al., 2015) (Gámez et al., 2019). A change in the task tempo (the timing in between beats) correlates with a change in the radii of the circular dynamics, with longer tempos corresponding to larger radii of the dynamics. The increase in radii is mediated by both, an increase in the number of neurons and an increase of duration of neural activation period within the sequence (Gámez et al., 2019).

Neural sequences are ubiquitous across species and brain regions, but there are fundamental differences in their dynamics. For example, the neural sequences in songbirds are very sparse (and thus have very high dimensionality), exhibiting time field widths on the order of tens of milliseconds. In contrast, the time field of neurons in the cortex and hippocampus may last hundreds of milliseconds and much higher "temporal density." Furthermore, there is some evidence that the neural dynamics in higher cortical areas present a lower degree of sequentially than areas located downstream. For example, during time-dependent tasks, the sequentiality index of neural activity in the striatum has been reported to be higher than that in the prefrontal and premotor cortex (Zhou et al., 2020; Handa et al., 2021), and the sequentiality in the hippocampus has been reported to be higher than in the prefrontal cortex (Ning et al., 2022). The higher degree of sequentiality is postulated to provide with a more efficient code to be read out by downstream biologically plausible mechanisms. As we will argue in the following sections, this higher sequentiality makes neural sequences an optimal dynamic regime for the decoding of time (Zhou et al., 2020).

Neurocomputational Models of Neural Sequences

The above experimental studies establish that neural sequences are observed across a wide range of brain areas, and at least in some cases exhibit specific features including embedding of multiple sequences and scaling. In parallel with these experimental studies, computational models have focused on the potential contribution of neural sequences to timing and other neurocomputational problems. In the context of encoding time, it has been proposed that neural sequences provide an ideal neural representation for downstream brain areas to flexibly generate either simple (e.g., a ramp or a timed step function) or complex spatiotemporal output patterns (Zhou et al., 2020, 2022).

As mentioned above, the activity patterns of RA-projecting HVC neurons provide one of the most compelling examples of sparse neural sequences underlying timing—specifically, of the temporal structure of birdsong (Hahnloser et al., 2002; Long & Fee, 2008; Long et al., 2010; Okubo et al., 2015; Lynch et al., 2016). Computational models have proposed that the neural sequences that underlie song production arise from feed-forward neural circuits in which neurons active at time t drive, through direct excitatory connections, the neurons that are active at time $t + 1$ (Jun & Jin, 2007; Long et al., 2010; Miller & Jin, 2013; Egger et al., 2020). Critically in these models, there is no recurrency, that is, units active at $t + 1$ do not synapse onto units active before $t + 1$. This feed-forward structure is consistent with the known circuitry of

HVC (Long et al., 2010; Egger et al., 2020); however, it is worth noting that intracellular recordings of RA-projecting HVC neurons reveal well-timed subthreshold events suggestive of presence of recurrent inputs (Long et al., 2010). At the computational level, at least, this lack of recurrency severely limits the number of neural sequences a population of neurons can encode.

Using either firing-rate or spike-based models of recurrent neural circuits, a number of models have demonstrated how single or multiple neural sequences can be embedded in recurrent neural networks (Rajan et al., 2016; Hardy & Buonomano, 2018). These models can be trained with supervised learning rules in which each unit in the recurrent neural network is given a target temporal profile from a template or an experimentally derived neuronal sequence (Laje & Buonomano, 2013; Rajan et al., 2016). In these multisequence models, different inputs trigger distinct neural sequences, and importantly, a given unit can participate in multiple sequences and have distinct time fields depending on which sequence is being played. Each of these neural sequences can be used to encode elapsed time from the onset of a cue, as well as represent a short-term memory that a given cue was presented.

As mentioned above, a fundamental feature of timing on the scale of seconds is its flexibility, i.e., the ability to generate a motor pattern, including playing an instrument or tapping out Morse code at multiple speeds (Jazayeri & Shadlen, 2010; Cicchini et al., 2012; Slayton et al., 2020). Thus, models of timing on the scale of seconds should account for the ability to generate the same pattern at multiple speeds. In the case of neural sequences, this means that a given neural sequence must be able to unfold at different speeds. A number of models have shown that under the proper conditions and training, it is indeed possible to trigger neural sequences at different speeds (Hardy et al., 2018; Remington et al., 2018; Liu et al., 2019; Zhou et al., 2022). This feature of temporal scaling generally has to be explicitly learned by a recurrent neural network and relies on separate inputs that govern speed, for example, by increasing the net input to all neuron which can reconfigure the dynamics in a manner that the trajectory flows more rapidly (or slowly) through neural state space (Goudar & Buonomano, 2018; Hardy & Buonomano, 2018).

Biologically Plausible Models of the Emergence of Neural Sequences

The above models have focused primarily on the circuit mechanisms potentially underlying the formation of neural sequences and have relied primarily on hard-wired or nonbiologically plausible learning rules to generate the appropriate architectures. Next, we address models that focus on how neural sequences may emerge in a biologically plausible manner.

Two of the first models addressing how neural sequences could emerge in a biologically plausible and unsupervised manner relied on a form of homeostatic plasticity termed presynaptic-dependent scaling (PSD). In these models (Buonomano, 2005; Liu & Buonomano, 2009), it was proposed that spiking neurons in a randomly connected recurrent network have a specific homeostatic setpoint of neural activity (e.g., one spike per some unit of time or trial). A specific subset of neurons within this population receives a brief external input at the onset of each “trial.” Neurons below their activity setpoint would homeostatically upregulate the strength of the synaptic weights coming from their presynaptic partners; critically, this homeostatic plasticity favored the potentiation of inputs from units that already had a high level of average activity. This is important because if an inactive neuron equitably potentiates inputs from presynaptic partners with both high and low levels of average activity, the system explodes when the silent partners eventually become active (Buonomano, 2005). It has been shown that this approach allows for the formation of multiple neural sequences in which a given unit can participate in many different neural sequences at different points in time (Fig. 2). Importantly, the network can produce multiple neural trajectories, and while it exhibited “functionally feed-forward activity,” the network architecture remained highly recurrent. However, as implemented in these models, the overall duration of the sequences was fairly short-lived (<300 ms).

Another early model of how neural sequences could emerge in a self-organizing manner proposed that spike-timing-dependent-plasticity (STDP) coupled with heterosynaptic competition could drive the formation of neural sequences (Fiete et al., 2010). The key component of this model was the presence of heterosynaptic competition in the form of presynaptic normalization, i.e., a ceiling on the total synaptic weights a given presynaptic unit could exert on its postsynaptic partners. This presynaptic normalization prevents units active early in the sequence from "capturing" many or all of the units active later in the sequence—and thus collapsing the neural sequence. This model, however, resulted primarily in purely feed-forward connectivity.

Other models of the emergence of neural sequences have also relied on STDP-based plasticity and proposed that newly generated neurons through adult neurogenesis contribute to the progressive growth of the sequence. One such model was developed in the context of song birds (Tupikov & Jin, 2021), animals in which adult neurogenesis can occur. In this model, spontaneously active neurons are randomly added to a growing neural chain, and because the new neurons exhibit higher spontaneous activity, they preferentially contribute to the growth of the chain. Runaway growth and plasticity are prevented by a passive decay term of synaptic weights and a cap on the number of postsynaptic units, and any presynaptic neuron can create strong synapses with a form of presynaptic normalization. Here again, the final neural architecture is primarily feed-forward.

Additional models have relied on associative plasticity, but are not self-organizing in the sense that during the learning phase a recurrent network is driven by a reproducible temporal pattern or a "tutor" that serves to embed a neural sequence into the RNN during training (Rajan et al., 2016; Murray & Escola, 2017; Gillett et al., 2020; Maes et al., 2020; Cone & Shouval, 2021). For example, Maes et al. (2020) demonstrated that separate clusters of recurrently connected neurons can be trained using voltage-dependent STDP, leading to the generation of neural sequences. Additionally, a spike-based model demonstrated that using a biological plausible LTP/LTD learning rule based on eligibility traces and trained on a target sequence also results in neural sequences that after training can be triggered by the initial component of the trained pattern (Cone & Shouval, 2021).

Whether in the presence or absence of tutor sequences, it is clear that biologically plausible learning rules capable of generating neural sequences are unlikely to be a consequence of simple forms of associative plasticity or STDP. This is particularly true when recurrent neural circuits in which a given neuron participates in multiple neural sequences (Buonomano, 2005; Fiete et al., 2010; Cone & Shouval, 2021; Tupikov & Jin, 2021). Indeed, some studies have demonstrated that associative plasticity effectively erases or collapses neural sequences (Liu & Buonomano, 2009; Bernacchia et al., 2022). Intuitively, one can see that STDP can result in a sequence collapse or latency reduction (Song et al., 2000; Liu & Buonomano, 2009; Lee & Buonomano, 2012): neurons that are sequentially activated strengthen the connections in the forward order, thus progressively decreasing the duration of the neural sequence, or even collapsing the neural sequence. It is thus clear that biologically plausible models of neural sequences at a minimum require mechanisms that counteract the strong correlations produced by neural sequences, and such counteracting forces may include presynaptic weight normalization, homeostatic plasticity, or inhibitory control.

Neural Sequences in Supervised Recurrent Neural Networks

Further evidence supporting the view that neural sequences comprise a canonical neural dynamic regime emerges from studies with supervised artificial neural networks. Neurocomputational and machine learning approaches often rely on training nonspiking recurrent neural networks

(RNN) using powerful (but nonbiological) gradient descent-based learning rules. Numerous studies have observed that neural sequences often emerge in RNNs trained to perform working memory or timing tasks (Orhan & Ma, 2019; Zhang et al., 2021; Zhou et al., 2022). The relevance of these findings is that supervised RNNs are trained to solve a specific task, but not to solve it any specific way. Thus, the dynamic regimes the RNNs exhibit emerge from the training algorithm per se, and the solutions are in a sense “assumption free.”

Here, we provide an example of a supervised firing-rate RNN trained to solve a simple timing task (Zhou et al., 2022, 2023): After a brief input *A* or *B*, the RNN must learn to produce a timed response at either 1 or 2.2 s, respectively (Fig. 3a). Thus, the RNN must encode elapsed time and trigger the output at the appropriate times in the absence of any external signals. In principle, this could be achieved a number of ways, including, ramping activity of the RNN units, complex patterns of neural activity, neural sequences, or even oscillatory activity. We show that the strategy RNNs use to encode time strongly depend on the temporal profile imposed on the output unit. If the output used was required to immediately step up at the target time, robust

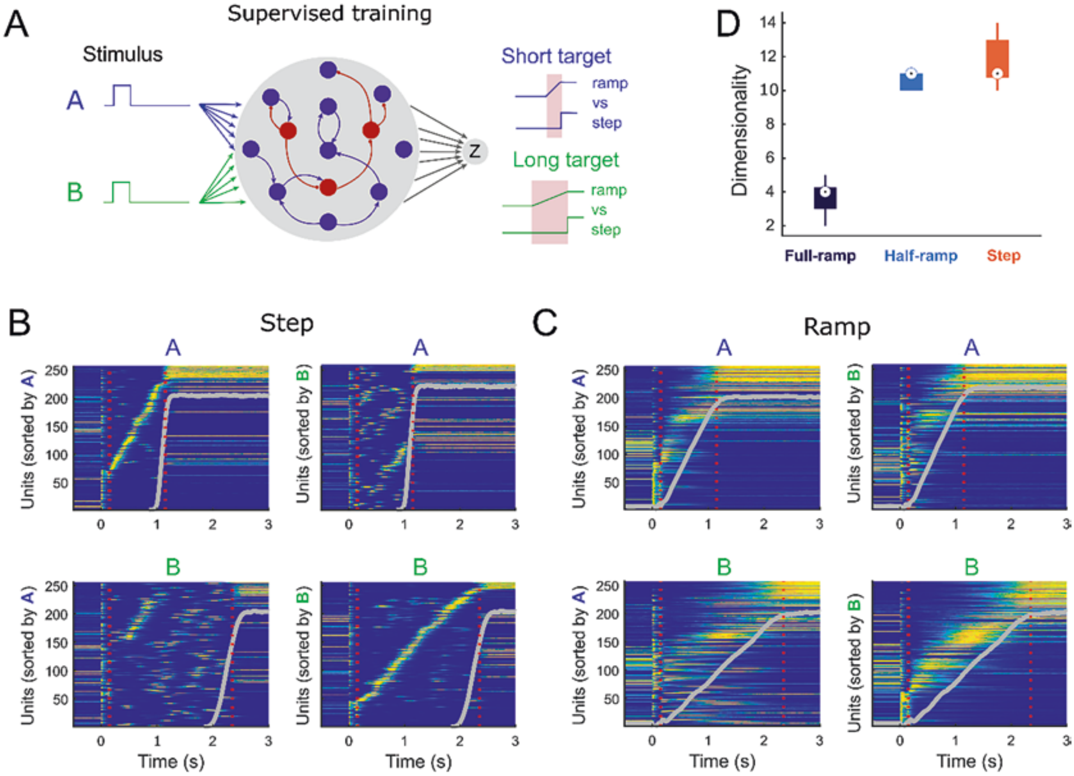


Fig. 3 Neural sequences emerge in supervised RNNs trained to solve a timing task. (**a**) A firing-rate RNN is trained via backpropagation through time to solve a stimulus-dependent timing task. After stimulus A, the network must produce an output that peaks after 1 second of elapsed time (short target). Under stimulus B, the network must produce the output at a longer interval, 2.2 s (long target). Two different target conditions for the output unit are used in the task: steps and ramps. For ramp targets, a continuous ramping of the output unit is imposed during the delay. For steps, the output unit must remain silent and only respond at the end of the delay. (**b**) Although the network learns to solve the task in all scenarios, different dynamic regimes emerge depending on the target output employed. During step outputs, robust neural sequences emerge. (**c**) During ramping outputs, ramping activity of the recurrent units is more prevalent. (**d**) The different dynamic regimes can be quantified by computing the dimensionality of the recurrent pull. With a step output target (or half-ramp that starts rising at the delay midpoint) higher dimensional and more sequential RNN dynamics is observed

and fairly sparse neural sequences emerged in the RNN (Fig. 3b). In contrast, if the RNN was trained to produce a full ramp that started at the offset of the stimulus and peaked at the target times (1 or 2.2 s), the neural dynamics within the RNN was significantly different, tending toward more ramp-like activity (Fig. 3c). This difference can be quantified by calculating the dimensionality of the RNN activity during the interval. We can define the dimensionality as the number of PCA components that are needed to account for 95% of the variability. If most units in the RNN are ramping, then the pairwise correlation between them will be high and the dimensionality low. In contrast, a sparse neural sequence will have high dimensionality because the average pairwise correlations between the units are low. We can see that as the output transition from a full ramp to a half-ramp or a step function, the dimensionality of the RNN activity dramatically increases (Fig. 3d). Thus, there is a relationship between the temporal profile of the target output and whether the RNN dynamics is low or high dimensional.

This observation is intuitive as it is straight forward to generate a ramping output unit from an RNN in which all units are also full ramps. But such ramping RNN activity cannot be used to generate an output such as a step function or a half-ramp (a ramp that starts at the interval midpoint). This is because—assuming a linear output unit—it is impossible to generate a step function from the linear sum of full ramps. Thus, the RNN must encode time through higher dimensional dynamics—which could in principle include a family of ramps with different onsets, a family of exponential decaying or rising activity profiles, oscillatory activity, or the actual neural sequences we observed.

These results are consistent with the notion that neural sequences represent a canonical dynamic regime because the high dimensionality and quasi-orthogonality of the RNN units provide a near optimal set of basis function that can be easily harnessed using biologically plausible learning rules to generate arbitrary spatiotemporal output patterns (Fiete et al., 2004; Zhou et al., 2020, 2022).

Conclusions

Over the past decades, there has been a strong shift away from notion that there may be a central timing circuit within the brain that is responsible for timing across scales. Today, we know that the brain uses multiple different mechanisms to tell time across different scales. And furthermore, even within the specific time scale of seconds, timing relies on multiple mechanisms and brain areas. Within this diversity, neural population clocks, in the form of neural sequences, are emerging as one of the main mechanisms underlying not only timing but a multitude of time-dependent computations.

Future studies, however, must address two general gaps in our current knowledge. First, a causal role for neural sequences in timing must be established, for example, by demonstrating that brief optogenetic perturbations that prematurely terminate a neural sequence impair timing. This approach may be hampered, however, by the possibility that many neural sequences are occurring in a distributed and interconnected manner throughout the brain. Second, the synaptic learning rules and principles that underlie the emergence of neural sequences in different neural circuits must be elucidated using cellular, systems, and computational approaches.

Acknowledgments We thank Shanglin Zhou for helpful discussions and figures.

Funding DB was supported by the National Institute of Health grant NS116589.

SSM was supported by the Swiss National Science Foundation grants P2ZHP3-187943 and P500PB-203133.

References

Abeles, M. (1991). *Corticonics*. Cambridge University Press.

Adler, A., Zhao, R., Shin, M. E., Yasuda, R., & Gan, W.-B. (2019). Somatostatin-expressing interneurons enable and maintain learning-dependent sequential activation of pyramidal neurons. *Neuron, 102*, 202–216.e207.

Balcı, F., & Simen, P. (2016). A decision model of timing. *Current Opinion in Behavioral Sciences, 8*, 94–101.

Bernacchia, A., Fiser, J., Hennequin, G., & Lengyel, M. (2022). Adaptive erasure of spurious sequences in sensory cortical circuits. *Neuron, 110*, 1857.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews: Neuroscience, 6*, 755–765.

Buhusi, C. V., Oprisan, S. A., & Buhusi, M. (2016). Clocks within clocks: Timing by coincidence detection. *Current Opinion in Behavioral Sciences, 8*, 207–213.

Buonomano, D. V. (2005). A learning rule for the emergence of stable dynamics and timing in recurrent networks. *Journal of Neurophysiology, 94*, 2275–2283.

Buonomano, D. V. (2007). The biology of time across different scales. *Nature Chemical Biology, 3*, 594–597.

Buonomano, D. V., & Laje, R. (2010). Population clocks: Motor timing with neural dynamics. *Trends in Cognitive Sciences, 14*, 520–527.

Buonomano, D. V., & Mauk, M. D. (1994). Neural network model of the cerebellum: Temporal discrimination and the timing of motor responses. *Neural Computation, 6*, 38–55.

Cicchini, G. M., Arrighi, R., Cecchetti, L., Giusti, M., & Burr, D. C. (2012). Optimal encoding of interval timing in expert percussionists. *The Journal of Neuroscience, 32*, 1056–1060.

Cone, I., & Shouval, H. Z. (2021). Learning precise spatiotemporal sequences via biophysically realistic learning rules in a modular, spiking network. *eLife, 10*, e63751.

Creelman, C. D. (1962). Human discrimination of auditory duration. *The Journal of the Acoustical Society of America, 34*, 582–593.

Crowe, D. A., Averbeck, B. B., & Chafee, M. V. (2010). Rapid sequences of population activity patterns dynamically encode task-critical spatial information in parietal cortex. *The Journal of Neuroscience, 30*, 11640–11653.

Crowe, D. A., Zarco, W., Bartolo, R., & Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *The Journal of Neuroscience, 34*, 11972–11983.

Diesmann, M., Gewaltig, M. O., & Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature, 402*, 529–533.

Durstewitz, D. (2003). Self-organizing neural integrator predicts interval times through climbing activity. *The Journal of Neuroscience, 23*, 5342–5353.

Egger, R., Tupikov, Y., Elmaleh, M., Katlowitz, K. A., Benezra, S. E., Picardo, M. A., Moll, F., Kornfeld, J., Jin, D. Z., & Long, M. A. (2020). Local axonal conduction shapes the spatiotemporal properties of neural sequences. *Cell, 183*, 537–548.e512.

Fiete, I. R., Hahnloser, R. H. R., Fee, M. S., & Seung, H. S. (2004). Temporal sparseness of the premotor drive is important for rapid learning in a neural network model of birdsong. *Journal of Neurophysiology, 92*, 2274–2282.

Fiete, I. R., Senn, W., Wang, C. Z. H., & Hahnloser, R. H. R. (2010). Spike-time-dependent plasticity and heterosynaptic competition organize networks to produce long scale-free sequences of neural activity. *Neuron, 65*, 563–576.

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

Gámez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biology, 17*, e3000054.

Gavornik, J. P., & Bear, M. F. (2014). Learned spatiotemporal sequence recognition and prediction in primary visual cortex. *Nature Neuroscience, 17*, 732–737.

Gewaltig, M. O., Diesmann, M., & Aertsen, A. (2001). Propagation of cortical synfire activity: Survival probability in single trials and stability in the mean. *Neural Networks, 14*, 657–673.

Gibbon, J. (1977). Scalar expectancy theory and Weber's law in animal timing. *Psychological Review, 84*, 279–325.

Gillett, M., Pereira, U., & Brunel, N. (2020). Characteristics of sequential activity in networks with temporally asymmetric Hebbian learning. *Proceedings of the National Academy of Sciences, 117*, 29948–29958.

Goudar, V., & Buonomano, D. V. (2018). Encoding sensory and motor patterns as time-invariant trajectories in recurrent neural networks. *eLife, 7*, e31134.

Gouvea, T. S., Monteiro, T., Motiwala, A., Soares, S., Machens, C., & Paton, J. J. (2015). Striatal dynamics explain duration judgments. *eLife, 4*, e11386.

Grondin, S. (2014). About the (non)scalar property for time perception. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of interval timing* (pp. 17–32). Springer New York.

Hahnloser, R. H., Kozhevnikov, A. A., & Fee, M. S. (2002). An ultra-sparse code underliesthe generation of neural sequences in a songbird. *Nature, 419*, 65–70.

Handa, T., Harukuni, R., & Fukai, T. (2021). Concomitant processing of choice and outcome in frontal corticostriatal ensembles correlates with performance of rats. *Cerebral Cortex, 31*, 4357–4375.

Hardy, N. F., & Buonomano, D. V. (2018). Encoding time in feedforward trajectories of a recurrent neural network model. *Neural Computation, 30*, 378–396.

Hardy, N. F., Goudar, V., Romero-Sosa, J. L., & Buonomano, D. V. (2018). A model of temporal scaling correctly predicts that motor timing improves with speed. *Nature Communications, 9*, 4732.

Harvey, C. D., Coen, P., & Tank, D. W. (2012). Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature, 484*, 62–68.

Helmholtz, H. (1860). *Handbuch der Physiologischen Optik*: English translation. Dover.

Heys, J. G., & Dombeck, D. A. (2018). Evidence for a subcircuit in medial entorhinal cortex representing elapsed time during immobility. *Nature Neuroscience, 21*, 1574–1582.

Issa, J. B., Tocker, G., Hasselmo, M. E., Heys, J. G., & Dombeck, D. A. (2020). Navigating through time: A spatial navigation perspective on how the brain may encode time. *Annual Review of Neuroscience, 43*, 43.

Ivry, R. B., & Spencer, R. M. C. (2004). The neural representation of time. *Current Opinion in Neurobiology, 14*, 225–232.

Jazayeri, M., & Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nature Neuroscience, 13*, 1020–1026.

Jun, J. K., & Jin, D. Z. (2007). Development of neural circuitry for precise temporal sequences through spontaneous activity, axon remodeling, and synaptic plasticity. *PLoS One, 2*, e723.

Killeen, P. R., & Fetterman, J. G. (1988). A behavioral theory of timing. *Psychological Review, 95*, 274–295.

Kraus, B. J., Robinson, R. J., II, White, J. A., Eichenbaum, H., & Hasselmo, M. E. (2013). Hippocampal "time cells": Time versus path integration. *Neuron, 78*, 1090–1101.

Laje, R., & Buonomano, D. V. (2013). Robust timing and motor patterns by taming chaos in recurrent neural networks. *Nature Neuroscience, 16*, 925–933.

Laje, R., Cheng, K., & Buonomano, D. V. (2011). Learning of temporal motor patterns: An analysis of continuous vs. reset timing. *Frontiers in Integrative Neuroscience, 5*, 61.

Lee, T. P., & Buonomano, D. V. (2012). Unsupervised formation of vocalization-sensitive neurons: A cortical model based on short-term and homeostatic plasticity. *Neural Computation, 24*, 2579–2603.

Liu, J. K., & Buonomano, D. V. (2009). Embedding multiple trajectories in simulated recurrent neural networks in a self-organizing manner. *The Journal of Neuroscience, 29*, 13172–13181.

Liu, Y., Tiganj, Z., Hasselmo, M. E., & Howard, M. W. (2019). A neural microcircuit model for a scalable scale-invariant representation of time. *Hippocampus, 29*, 260–274.

Long, M. A., & Fee, M. S. (2008). Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature, 456*, 189–194.

Long, M. A., Jin, D. Z., & Fee, M. S. (2010). Support for a synaptic chain model of neuronal sequence generation. *Nature, 468*, 394–399.

Lynch, G. F., Okubo, T. S., Hanuschkin, A., Hahnloser, R. H., & Fee, M. S. (2016). Rhythmic continuous-time coding in the songbird analog of vocal motor cortex. *Neuron, 90*, 877–892.

MacDonald, C. J., Lepage Kyle, Q., Eden Uri, T., & Eichenbaum, H. (2011). Hippocampal "time cells" bridge the gap in memory for discontiguous events. *Neuron, 71*, 737–749.

MacDonald, C. J., Carrow, S., Place, R., & Eichenbaum, H. (2013). Distinct hippocampal time cell sequences represent odor memories in immobilized rats. *The Journal of Neuroscience, 33*, 14607–14616.

Maes, A., Barahona, M., & Clopath, C. (2020). Learning spatiotemporal signals using a recurrent spiking network that discretizes time. *PLoS Computational Biology, 16*, e1007606.

Matell, M. S., Meck, W. H., & Nicolelis, M. A. (2003). Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behavioral Neuroscience, 117*, 760–773.

Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. *Annual Review of Neuroscience (Palo Alto, CA), 27*, 307–340.

Mauk, M. D., & Donegan, N. H. (1997). A model of Pavlovian eyelid conditioning based on the synaptic organization of the cerebellum. *Learning & Memory, 3*, 130–158.

Mello, G. B., Soares, S., & Paton, J. J. (2015). A scalable population code for time in the striatum. *Current Biology, 25*, 1113–1122.

Merchant, H., Zarco, W., & Prado, L. (2008). Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. *Journal of Neurophysiology, 99*, 939–949.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience, 36*, 313–336.

Merchant, H., Pérez, O., Bartolo, R., Méndez, J. C., Mendoza, G., Gámez, J., Yc, K., & Prado, L. (2015). Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. *European Journal of Neuroscience, 41*, 586–602.

Miall, C. (1989). The storage of time intervals using oscillating neurons. *Neural Computation, 1*, 359–371.

Miller, A., & Jin, D. Z. (2013). Potentiation decay of synapses and length distributions of synfire chains self-organized in recurrent neural networks. *Physical Review E, 88*, 062716.

Mumford, D. (1992). On the computational architecture of the neocortex. II. The role of cortico-cortical loops. *Biological Cybernetics, 66*, 241–251.

Murray, J. M., & Escola, G. S. (2017). Learning multiple variable-speed sequences in striatum via cortical tutoring. *eLife, 6*, e26084.

Ning, W., Bladon, J. H., & Hasselmo, M. E. (2022). Complementary representations of time in the prefrontal cortex and hippocampus. *Hippocampus, 32*, 577.

Okubo, T. S., Mackevicius, E. L., Payne, H. L., Lynch, G. F., & Fee, M. S. (2015). Growth and splitting of neural sequences in songbird vocal development. *Nature, 528*, 352–357.

Orhan, A. E., & Ma, W. J. (2019). A diverse range of factors affect the nature of neural representations underlying short-term memory. *Nature Neuroscience, 22*, 275–283.

Pastalkova, E., Itskov, V., Amarasingham, A., & Buzsaki, G. (2008). Internally generated cell assembly sequences in the rat hippocampus. *Science, 321*, 1322–1327.

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

Rajan, K., Harvey Christopher, D., & Tank David, W. (2016). Recurrent network models of sequence generation and memory. *Neuron, 90*, 128.

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

Reddy, L., Zoefel, B., Possel, J. K., Peters, J., Dijksterhuis, D. E., Poncet, M., van Straaten, E. C., Baayen, J. C., Idema, S., & Self, M. W. (2021). Human hippocampal neurons track moments in a sequence of events. *The Journal of Neuroscience, 41*, 6714–6725.

Remington, E. D., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible sensorimotor computations through rapid reconfiguration of cortical dynamics. *Neuron, 98*, 1005–1019.e1005.

Shimbo, A., Izawa, E.-I., & Fujisawa, S. (2021). Scalable representation of time in the hippocampus. *Science Advances, 7*, eabd7013.

Simen, P., Balci, F., de Souza, L., Cohen, J. D., & Holmes, P. (2011). A model of interval timing by neural integration. *The Journal of Neuroscience, 31*, 9238–9253.

Slayton, M. A., Romero-Sosa, J. L., Shore, K., Buonomano, D. V., & Viskontas, I. V. (2020). Musical expertise generalizes to superior temporal scaling in a Morse code tapping task. *PLoS One, 15*, e0221000.

Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neuroscience, 3*, 919–926.

Taxidis, J., Pnevmatikakis, E. A., Dorian, C. C., Mylavarapu, A. L., Arora, J. S., Samadian, K. D., Hoffberg, E. A., & Golshani, P. (2020). Differential emergence and stability of sensory and temporal representations in context-specific hippocampal sequences. *Neuron, 108*, 984–998.e989.

Tiganj, Z., Jung, M. W., Kim, J., & Howard, M. W. (2017). Sequential firing codes for time in rodent medial prefrontal cortex. *Cerebral Cortex, 27*, 5663–5671.

Treisman, M., Faulkner, A., Naish, P. L. N., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception, 19*, 705–743.

Tupikov, Y., & Jin, D. Z. (2021). Addition of new neurons and the emergence of a local neural circuit for precise timing. *PLoS Computational Biology, 17*, e1008824.

Umbach, G., Kantak, P., Jacobs, J., Kahana, M., Pfeiffer, B. E., Sperling, M., & Lega, B. (2020). Time cells in the human hippocampus and entorhinal cortex support episodic memory. *Proceedings of the National Academy of Sciences, 117*, 28463–28474.

Wang, J., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible timing by temporal scaling of cortical responses. *Nature Neuroscience, 21*, 102–110.

Wang, M., Foster, D. J., & Pfeiffer, B. E. (2020). Alternating sequences of future and past behavior encoded within hippocampal theta oscillations. *Science, 370*, 247–250.

Xu, S., Jiang, W., Poo, M.-m., & Dan, Y. (2012). Activity recall in a visual cortical ensemble. *Nature Neuroscience, 15*, 449–455.

Zhang, X., Liu, S., & Chen, Z. S. (2021). A geometric framework for understanding dynamic information integration in context-dependent computation. *iScience, 24*, 102919.

Zhou, S., Masmanidis, S. C., & Buonomano, D. V. (2020). Neural sequences as an optimal dynamical regime for the readout of time. *Neuron, 108*, 651–658.e655.

Zhou, S., Masmanidis, S. C., & Buonomano, D. V. (2022). Encoding time in neural dynamic regimes with distinct computational tradeoffs. *PLoS Computational Biology, 18*, e1009271.

Zhou, S., Seay, M., Taxidis, J., Golshani, P., & Buonomano, D. V. (2023). *Multiplexing working memory and time in the trajectories of neural networks* (Vol. 7, p. 1170). Nature Human Behaviour.

Temporal Information Processing in the Cerebellum and Basal Ganglia

Masaki Tanaka, Masashi Kameda,
and Ken-ichi Okada

Abstract

Temporal information processing in the range of a few hundred milliseconds to seconds involves the cerebellum and basal ganglia. In this chapter, we present recent studies on nonhuman primates. In the studies presented in the first half of the chapter, monkeys were trained to make eye movements when a certain amount of time had elapsed since the onset of the visual cue (time production task). The animals had to report time lapses ranging from several hundred milliseconds to a few seconds based on the color of the fixation point. In this task, the saccade latency varied with the time length to be measured and showed stochastic variability from one trial to the other. Trial-to-trial variability under the same conditions correlated well with pupil diameter and the preparatory activity in the deep cerebellar nuclei and the motor thalamus. Inactivation of these brain regions delayed saccades when asked to report subsecond intervals. These results suggest that the internal state, which changes with each trial, may cause fluctuations in cerebellar neuronal activity, thereby producing variations in self-timing. When measuring different time intervals, the preparatory activity in the cerebellum always begins approximately 500 ms before movements, regardless of the length of the time interval being measured. However, the preparatory activity in the striatum persists throughout the mandatory delay period, which can be up to 2 s, with different rate of increasing activity. Furthermore, in the striatum, the visual response and low-frequency oscillatory activity immediately before time measurement were altered by the length of the intended time interval. These results indicate that the state of the network, including the striatum, changes with the intended timing, which lead to different time courses of preparatory activity. Thus, the basal ganglia appear to be responsible for measuring time in the range of several hundred milliseconds to seconds, whereas the cerebellum is responsible for regulating self-timing variability in the subsecond range. The second half of this chapter presents studies related to periodic timing. During eye movements synchronized with alternating targets at regular intervals, different neurons in the cerebellar nuclei exhibit activity related to movement timing, predicted stimulus timing, and the temporal error of synchronization. Among these, the activity associated with target appearance is particularly enhanced during synchronized movements and may represent an internal model of the temporal structure of stimulus sequence.

M. Tanaka (✉) · M. Kameda · K. Okada
Department of Physiology, Hokkaido University
School of Medicine, Sapporo, Japan
e-mail: masaki@med.hokudai.ac.jp

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_6

We also considered neural mechanism underlying the perception of periodic timing in the absence of movement. During perception of rhythm, we predict the timing of the next stimulus and focus our attention on that moment. In the missing oddball paradigm, the subjects had to detect the omission of a regularly repeated stimulus. When employed in humans, the results show that the fastest temporal limit for predicting each stimulus timing is about 0.25 s (4 Hz). In monkeys performing this task, neurons in the cerebellar nuclei, striatum, and motor thalamus exhibit periodic activity, with different time courses depending on the brain region. Since electrical stimulation or inactivation of recording sites changes the reaction time to stimulus omission, these neuronal activities must be involved in periodic temporal processing. Future research is needed to elucidate the mechanism of rhythm perception, which appears to be processed by both cortico-cerebellar and cortico-basal ganglia pathways.

Keywords

Time production · Synchronized movement · Rhythm perception · Sensory prediction · Internal model · Cerebellar nucleus · Striatum · Motor thalamus · Nonhuman primate

Introduction

When we wait at a traffic light or ride to the beat of music, we unconsciously measure the elapsed time from the previous event and accurately predict the timing of the next event. Temporal information processing in the range of a few hundred milliseconds to several seconds is essential for our daily activities and involves multiple networks in the brain, including the cortico-basal ganglia-thalamocortical pathways and cortico-cerebellar-thalamocortical pathways (Buhusi & Meck, 2005; Coull et al., 2011; Merchant et al., 2013; Ivry & Schlerf, 2008). As experienced in everyday life, these processes are influenced by various internal and external factors (Wittmann, 2013; Lake et al., 2016). In recent years, the neural mechanisms of temporal information processing have been investigated in detail in laboratory animals that perform a variety of tasks, such as time production or reproduction (Jazayeri & Shadlen, 2015; Mita et al., 2009; Schneider & Ghose, 2012; Yumoto et al., 2011; Kunimatsu et al., 2018; Tanaka, 2007), temporal discrimination (Leon & Shadlen, 2003; Shimbo et al., 2021; Mendoza et al., 2018; Chiba et al., 2021), synchronized movement (Merchant et al., 2011; Gamez et al., 2019; Okada et al., 2022), and internalized rhythm (Cadena-Valencia et al., 2018; Ohmae et al., 2013; Kameda et al., 2019). In this chapter, we present the results of our behavioral experiments and neuronal recordings from the cerebellum, striatum, and motor thalamus of nonhuman primates, and discuss the roles of these subcortical structures. We first describe the neural mechanisms involved in measuring the time elapsed from external events, and then discuss the neural mechanisms involved in predicting the timing of periodic stimuli.

Production of Single Time Interval

Variability of Subjective Time Passage

One way to explore the mechanism of temporal information processing is to ask subjects to report the passage of a certain length of time by, for example, pressing a stopwatch at a specific time interval. Even for such a simple task, the length of the produced time varies from trial to trial, likely due to fluctuations in internal factors such as attention and arousal level. Suzuki et al. (2016) examined the relationship between the speed of subjective time passage and pupil diameter, an objective marker of internal state, in monkeys performing a time production task. In their task, a visual cue was presented briefly (0.1 s) while the animals looked at the fixation point on the screen (Fig. 1a). The animals were rewarded if they moved their eyes toward the location of the previous cue after a mandatory delay interval of 1 s. Since the fixation point disappeared only after

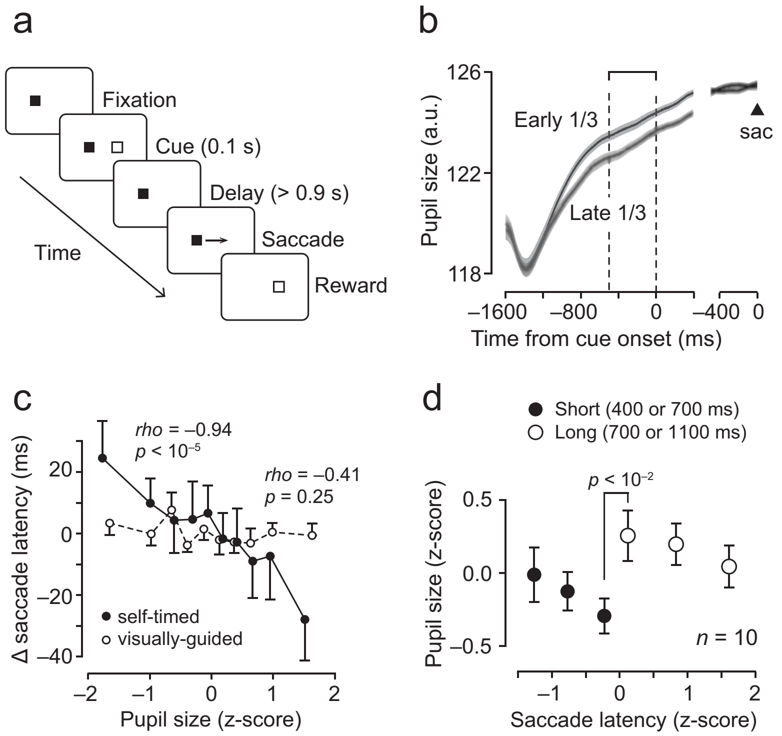


Fig. 1 Correlation between pupil size and the rate of subjective time lapse. (**a**) Self-timed saccade task. Monkeys made a memory-guided saccade to the location of a previously presented visual cue after a 1-second mandatory delay period. Since the fixation point disappeared only after the saccade, animals needed to measure the time elapsed from the visual cue. (**b**) Time courses of pupil size aligned with the cue onset. Trials are divided into three groups according to saccade latency, and the mean (± SE) pupil size for the earliest and latest saccade groups are shown. Two vertical dashed lines during eye fixation represent the time window for quantitative analysis. Note that the pupil size before the cue onset is larger for trials with shorter saccade latencies. (**c**) Relationship between pupil size and latency (relative to the mean) of self-timed (black circles) and visually guided (white circles) saccades. Error bars indicate ±95% CI. (**d**) Comparison between different mandatory delay intervals. Color of equiluminant fixation point indicated two different delay intervals. Data points represent the normalized mean (±95% CI) pupil size and saccade latency for each of the three saccade latency groups. Note that the pupil size is inversely correlated with the group-by-group variation in saccade latency under the same conditions, but not with actual saccade latency across conditions. Data summarize multiple experiments in two monkeys. (Adapted from Suzuki et al. (2016) under the terms of the Creative Commons Attribution 4.0 International License (CC-BY))

the saccade, the animals had to monitor the elapsed time in each trial. Data from several hundred trials were divided into three groups according to the length of the monkey's subjective 1-s (i.e., saccade latency), and the time course of pupil size during fixation was compared between the groups with the longest and shortest saccade latencies (Fig. 1b). In both groups of trials, pupils contracted immediately after the start of fixation and then gradually dilated, but interestingly, pupil diameters before the cue onset were clearly different depending on the subsequent self-timed saccade latency. Data from multiple experiments showed that pupil diameter before cue presentation was inversely correlated with the length of time the animals subsequently reported, with larger pupil diameters indicating shorter times (faster passage of time) and smaller pupil diameters indicating longer times (Fig. 1c). In contrast, pupil diameter did not correlate with the latency of visually guided saccades, where monkeys generated an immediate saccade toward the visual stimulus. Thus, internal factors influence temporal processing, and objective measures such as pupil diameter could be used to infer how subjects perceive the passage of time.

Then, when measuring longer duration, does the pupil size become much smaller? Suzuki et al. (2016) further trained animals to report two different time intervals (0.7 and 1.1 s, for example) according to the color of equiluminant fixation point. When the relationship between production time and pupil diameter was examined, pupil diameter was different under different instructions, even if the self-timed saccade latencies were similar (Fig. 1d). Furthermore, while saccade latency and pupil size were inversely correlated within each condition, pupil size tended to be larger when the intended time length was long. These results indicate that internal factors fluctuate temporal processing during individual trials under the same condition but may have no relationship with the length of the intended time interval, or that the relationship to the intended timing is the opposite of the trial-by-trial variation. These results suggest that, when reporting elapsed time, there are two signals, one related to the length of time to be measured and the other related to trial-by-trial variability, and pupil diameter correlates well with the latter.

Preparatory Activity for Self-Timing

How is the trial-by-trial variation in self-timing controlled in the brain? It is widely accepted that the timing of self-initiated movement is determined by the time course of the preparatory activity, which gradually increases over time (Jazayeri & Shadlen, 2015; Maimon & Assad, 2006; Janssen & Shadlen, 2005; Merchant & Averbeck, 2017; Lee & Assad, 2003; Ashmore & Sommer, 2013; Xu et al., 2014; Parker et al., 2014; Dacre et al., 2021). During motor preparation, there are a variety of neural activities with different time courses, but integration of these signals leads to the generation of a gradual ramp-like activity (Murakami et al., 2014). This indicates that neurons exhibiting activity at different times cooperate to integrate information into the network and generate ramp-like activity that ultimately controls the timing of movement. Therefore, trial-by-trial variations in self-timing are likely to be reflected in the time courses of the ramping activity.

Previous studies have shown that neurons in the supplementary eye field (SEF), the motor thalamus, the cerebellar dentate nucleus, and the caudate nucleus of the basal ganglia exhibit a significant preparatory activity during the self-timed saccade task (Kunimatsu et al., 2018; Tanaka, 2007; Ohmae et al., 2017; Kunimatsu & Tanaka, 2012). Figure 2a illustrates a neuron recorded from the ventrolateral (VL) thalamus, which showed a gradual increase in activity before self-initiated saccades (Tanaka, 2007). The time course of neuronal activity differed from trial to trial, peaking at the time of saccade initiation and then declining rapidly. When the data for each neuron were divided into five groups according to saccade latency, and the population activity was computed for each group, the variation in saccade timing correlated well with the slope of ramping activity (Fig. 2b). In addition, the firing rate immediately before saccade initiation was comparable between the groups, which was in good agreement with the rise-to-threshold model of decision making. Indeed, these neuronal activities appear to be closely linked to self-timing, as pharmacological inactivation of the recording sites in the VL thalamus delays saccades (Tanaka, 2006).

Similar preparatory activity has also been recorded in the cerebellar dentate nucleus, which projects to the VL thalamus. Ohmae et al. (2017) trained monkeys to generate self-initiated saccades and varied the mandatory delay interval from 0.4 to 2.4 s in different trial blocks. For short delay intervals (≤1.2 s), a significant correlation was found between the saccade latency and the rate of increase in preparatory activity (Fig. 2c). However, for the long delay interval (2.4 s), saccade latency did not correlate with the slope of the ramping activity (Fig. 2d). Instead, a correlation was found between saccade latency and the onset time of preparatory activity. These results suggest that during measurements of time interval shorter than ~1.2 s, trial-to-trial variations in response timing may reflect changes in the magnitude of cerebellar neuronal activity. On the other hand, during measurements of longer

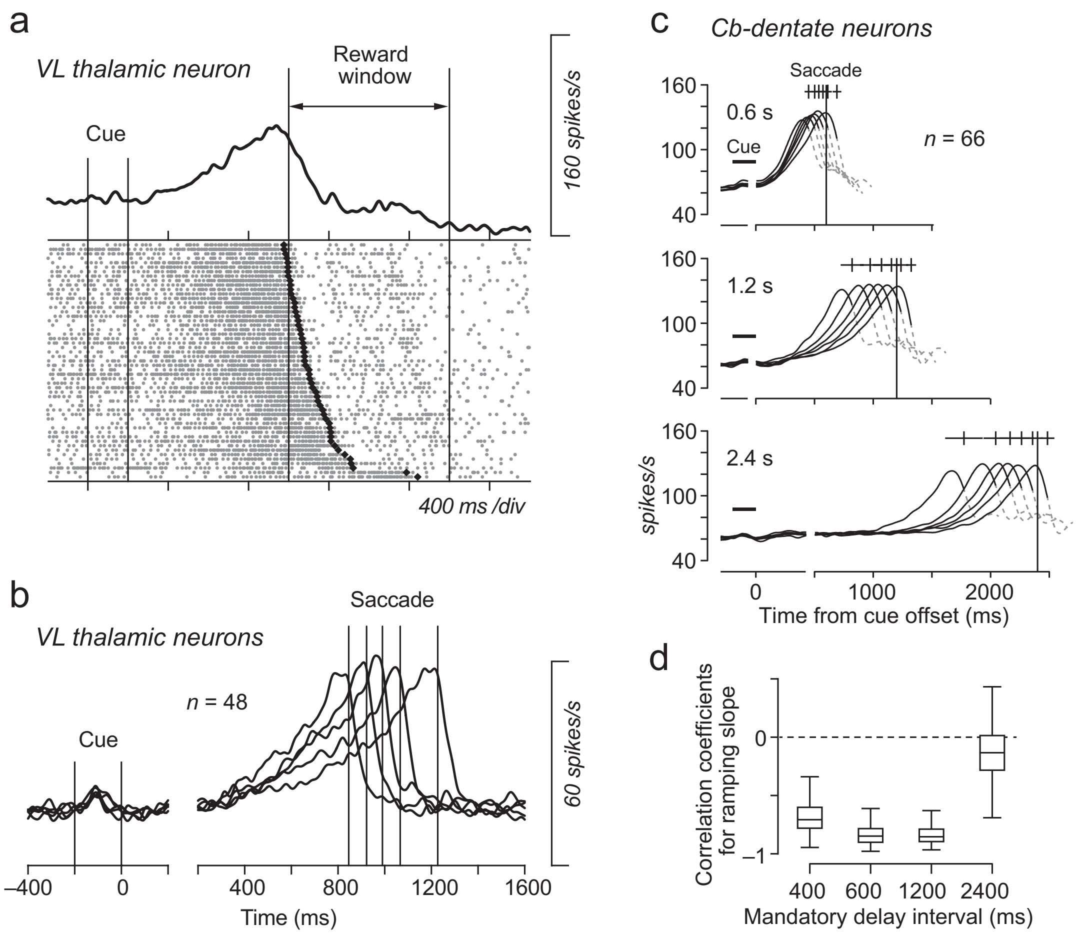


Fig. 2 Time courses of preparatory neuronal activity for self-timed saccades in the ventrolateral (VL) thalamus and the cerebellar dentate nucleus. (**a**) Activity of an example VL thalamic neuron during the self-timed saccade task. Data are aligned with the cue onset in the preferred direction, sorted according to saccade latency. Black symbol on each raster line indicates the time of saccade. Solid continuous line above the rasters indicates spike density. The animal received a reward for saccades generated within the reward window indicated by vertical lines. (**b**) Time courses of population activity of VL thalamic neurons. For each neuron, data were divided into five groups according to saccade latencies. Data were aligned either with the cue (left) or saccades (right), then were averaged across the population. In the right panel, traces are shifted in time so that the times of saccades (vertical lines) are placed at the means of the saccade latency relative to the cue offset. (**c**) Time courses of population activity of ramp-up neurons in the dentate nucleus of the cerebellum for different saccade latencies. For each neuron and mandatory delay interval, the data were divided into six groups according to saccade latencies. Crosses indicate the means and SDs of saccade latency for different groups. (**d**) Summary of correlation coefficients (Pearson's r) between the slope of ramping activity and saccade latency in the cerebellar dentate neurons. Box-whisker plots show the median, quartiles, and range of the results of the bootstrap analysis. (Adapted from Tanaka (2007) and Ohmae et al. (2017) under the terms of the Creative Commons Attribution 4.0 International License (CC-BY))

intervals, signals from the other areas that trigger preparatory activity in the cerebellum may be responsible for variations in self-timing.

One such candidate is the basal ganglia that are also involved in the generation of self-timed movements. In Parkinson's disease, the onset of self-initiated movements is delayed and the time measurements in seconds are inaccurate (Coull et al., 2011; Smith et al., 2007; Honma et al., 2016; Tokushige et al., 2018; Allman & Meck, 2012). Previous studies have suggested that the basal ganglia are mainly involved in measuring suprasecond time intervals, while the cerebellum is involved in subsecond timing (Buhusi & Meck, 2005; Lewis & Miall, 2003). The activities of single neurons in the cerebellar dentate nucleus and the striatum (caudate nucleus) were compared in monkeys trained to report 0.4, 1.0, and 2.2 s elapses according to the color of the fixation point (Fig. 3a), which varied from trial to trial (Kunimatsu et al., 2018). Neurons in both structures showed strong preparatory activity at all intervals, but their time courses were clearly different. The preparatory activity in the cerebellum started approximately 500 ms before the self-timed movements, while that in the striatum started at the onset of the delay period (i.e., immediately after the visual cue), and the rate of increase in neuronal activity depended on the length of time to be measured (Fig. 3b, c). Furthermore, when the time course of neuronal activity in each condition was compared across the three groups of trials sorted by saccade latency, the trial-by-trial variation started earlier in the cerebellum than in the striatum (Fig. 3d). Thus, for the measurements of the suprasecond interval, preparatory activity first started in the striatum, while the trial variation emerged earlier in the cerebellum. These results suggest that the striatum monitors the passage of time throughout

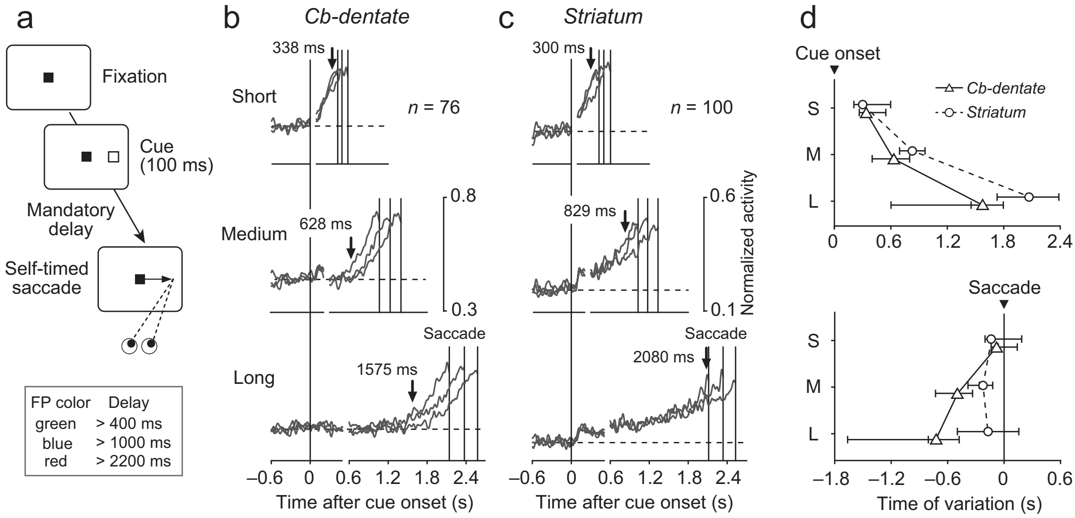


Fig. 3 Timing of trial-by-trial variation of ramping activity in the cerebellar dentate nucleus and the striatum (the caudate nucleus) during the self-timed saccade task. (**a**) Sequence of events in the self-timed saccade task. During central fixation, a cue flashed briefly (100 ms) in the peripheral visual field. Monkeys were required to remember the cue location and maintain fixation until expiration of the predetermined mandatory delay interval that was indicated by color of the fixation point (bottom inset). Animals received a reward if they correctly made a self-timed memory-guided saccade to the cue location after the mandatory delay period. (**b**, **c**) Time courses of population activity for neurons in the cerebellum (**b**) and striatum (**c**). For each delay condition, trials were divided into three groups according to saccade latency. Then, the data were normalized for each neuron, aligned with saccade initiation, and were shifted in time so that the times of saccades (vertical lines) were placed at the mean saccade latencies relative to the cue onset (right panels). On the left panels, data of the population activity were aligned with the cue onset (left vertical line). Downward arrows indicate the time when the traces of normalized neuronal activities started to diverge as detected by repeated measures ANOVAs ($p < 0.01$ for consecutive 40 ms, uncorrected for multiple comparisons). (**d**) Times of onset of trial-by-trial variation relative to the cue (upper panel) or saccade initiation (bottom). Each datapoint indicates the mean of the data derived from the analysis shown in b and c. Error bars with three tick marks denote 2.5, 50, and 97.5 percentiles obtained by the bootstrap analysis. Note that the trial-by-trial variation started earlier in the cerebellum than the striatum for medium and long delay intervals. (Adapted from Kunimatsu et al. (2018) under the terms of the Creative Commons Attribution 4.0 International License (CC-BY))

the delay period and reports the time relative to the entire interval. In contrast, the cerebellum may regulate movement timing within the range of several hundred milliseconds in each trial. Consistent with this, local inactivation of the cerebellar nucleus delayed only subsecond timing, whereas inactivation of the striatum altered both subsecond and suprasecond timing (Kunimatsu et al., 2018).

The recording studies described so far have examined only ramping activity, but other recent studies have examined the time course of neural trajectories in populations of neurons during timing behavior, under the assumption that individual neurons with different time courses are equally involved in timing. These studies showed that neurons in the medial frontal cortex and striatum flexibly change the shape and speed of neural trajectories depending on measured intervals (Wang et al., 2018; Gámez et al., 2019; Betancourt et al., 2023; Meirhaeghe et al., 2021). Striatal neurons with different slopes of ramping activity can be considered as part of the population producing these trajectory changes, and the sequential activity of many neurons may underlie the generation of ramping activity (Zhou & Buonomano, 2022). On the other hand, it is not yet known how the neural trajectories of cerebellar neuronal populations change in measurements ranging from several hundred milliseconds to a few seconds. Although a cerebellar model that generates ramping activity through learning has been proposed (Narain et al., 2018), its temporal limit has not been investigated. Future analysis is needed to determine how neural trajectory of population of cerebellar neurons can relate to subsecond fluctuations in self-timing.

Neural Correlates of Intended Timing

As mentioned above, in the striatum, the rate of increase in preparatory activity for different intervals varied from the beginning of the delay period (Fig. 3c). This indicates that neural excitability in the network including the striatum may flexibly change upon the instruction of the time length to be measured. Indeed, the response of striatal local field potentials (LFPs) to the peripheral visual cues (visual evoked potentials) scaled depending on the length of the mandatory delay interval (which was indicated at the trial start by the color of the fixation point), with large responses for short duration and small responses for long duration (Fig. 4a, b) (Suzuki & Tanaka, 2019). In addition, the LFP power of the low-frequency components (6–20 Hz) just before the cue presentation was proportional to the time interval to be measured, thereby indicating that the network state was altered during preparation of time measurement (Fig. 4c, d). However, when the data were analyzed separately by latency variation under the same conditions, no change in visual response or low-frequency power was found (Fig. 4e). Thus, the intention to measure a specific time interval alters the network state and excitability of striatal neurons, but the stochastic variation of the internal state under the same conditions may not.

The results are summarized in Table 1. When measuring time intervals, there are two sources of variation, one related to the intended length of time and the other to stochastic variation. The preparatory activity for self-timed movements began early in the striatum, regardless of the length of the intended time length, and began shortly (~500 ms) before the movement in the cerebellum. In contrast, the effects of trial-to-trial variability were first observed in the cerebellum and later found in the striatum. The visual response and the low-frequency power in the striatum scaled according to the length of the instructed time interval but not greatly by the stochastic variation of the produced interval. Pupil size, on the other hand, was strongly correlated with trial-to-trial variability in self-timing but not with the intended time interval to be measured.

These findings suggest that the cerebellum coordinates motor timing within a range of second or less, and that cerebellar activity has a significant influence on the intertrial variability of self-timing. The pupil diameter is known to correlate well with the activity of the central noradrenergic system (Murphy et al., 2014; Joshi et al., 2016; Aston-Jones & Cohen, 2005), and the noradrenergic neurons in the dorsal pons send pro-

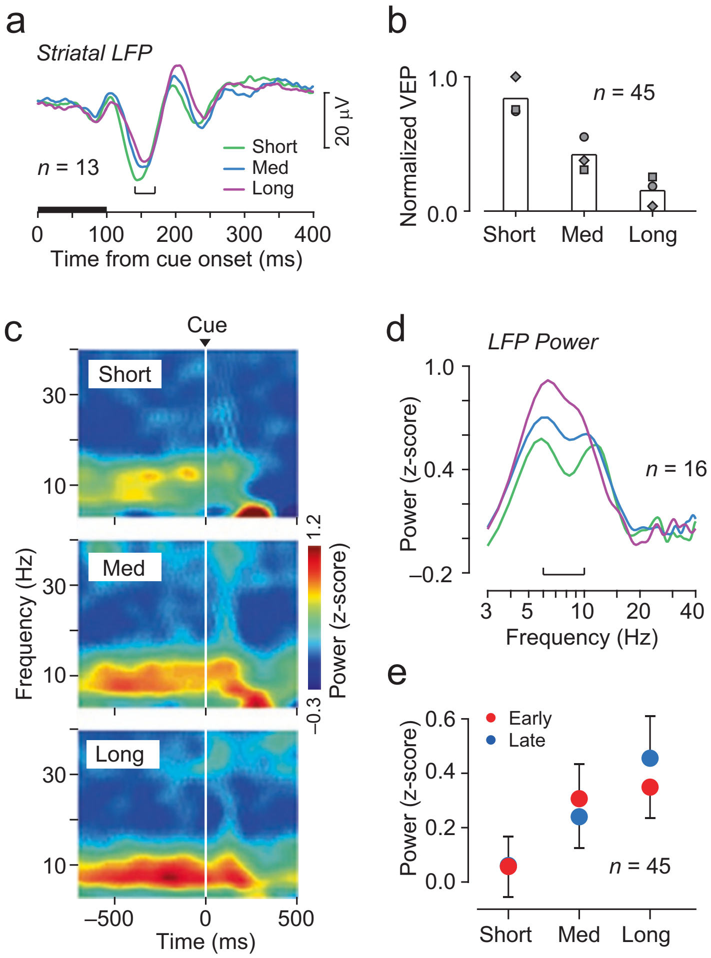


Fig. 4 Contextual modulation of local field potentials (LFPs) in the striatum during the preparation of time measurements. (**a**) Time courses of striatal LFPs aligned with the contralateral visual cue (horizontal black bar) in the self-timed saccade task. Colors indicate different interval conditions. Bracket denotes the time range of the maximum response in individual experiments. (**b**) Comparison of the magnitude of visually evoked response across conditions. Each bar summarizes the mean of normalized response obtained from 45 recording sites. Different symbols plot the data of different monkeys. (**c**) Color-coded power spectra of LFP for self-timed trials with different interval conditions. Vertical white line indicates the cue onset. Note that the power of the low-frequency component increases with longer delay intervals. (**d**) Spectral analysis of LFP during the precue period (500 ms). Bracket denotes the frequency range showing a significant difference across the interval conditions. (**e**) Power modulation associated with trial-to-trial variation in saccade latency. For each interval condition, data were divided into trials with early and late saccades, and the low-frequency power was computed for each. (Adapted from Suzuki and Tanaka (2019) under the terms of the Creative Commons Attribution 4.0 International License (CC-BY))

Table 1 Influence of variation of measured interval on neural and pupil response

	Striatum			Cerebellum	
Source of variation	Preparatory activity	Visual response	α-β power	Preparatory activity	Pupil size
Intended	Early	Scaled	Scaled	Late	–
Stochastic	Late	–	–	Early	Scaled

Minus sign indicates weak or no responsiveness to temporal variation jections to many brain regions, including the cerebellum, with exceptionally few projections to the striatum (Baldo et al., 2003; Jones & Yang, 1985). This suggests that stochastic changes in internal states may be accompanied by changes in the activity of neuromodulators, such as noradrenaline. This contributes to behavioral variations via changes in neuronal activity in the cerebellum. In contrast, intentional changes in the measured time length were accompanied by changes in the neural excitability of the striatum. It is well known that the low-frequency component of striatal LFP is related to dopamine, and low beta power (8–30 Hz) within the basal ganglia circuitry is significantly increased in Parkinson's disease (Hammond et al., 2007). Since dopamine in the striatum is also known to be involved in interval timing (De Corte et al., 2019; Hamilos et al., 2021; Soares et al., 2016; Kunimatsu & Tanaka, 2016), dopamine may contribute to changing the network state in the cortico-basal ganglia pathways on each trial, depending on the length of time being measured. In fact, patients with Parkinson's disease have a difficulty in varying the measurement time from trial to trial according to instructions, thereby showing a stronger central tendency than normal subjects (so-called "migration effect") (Malapani et al., 1998). However, the role of dopamine during normal conditions still needs to be explored, as a recent study has shown that the suppression of beta power associated with movement is not temporally coupled with dopamine release (Schwerdt et al., 2020). In addition, acetylcholine signaling in the striatum has also been shown to be involved in self-timing (Kunimatsu & Tanaka, 2016), and the relationship between low-frequency oscillations and acetylcholine awaits further clarification.

Rhythmic Timing

Role of the Cerebellum in Synchronized Movement

In addition to self-timed single movements, a series of repetitive movements synchronized to an external rhythm, such as tapping, have been used to investigate the neural mechanisms of temporal processing (Merchant et al., 2013; Kotz et al., 2018; Chauvigne et al., 2014; Gámez et al., 2019; Betancourt et al., 2023). Synchronized movement requires the learning of stimulus tempo, generation of an internal model for periodicity, precise control of movement timing, detection of temporal error, and update of the internal model (Repp, 2005). Although predictive synchronized movements such as musical dance are frequently observed in everyday life, spontaneous synchronization is thought to be unique to species with vocal learning abilities (Patel et al., 2009; Merchant & Honing, 2013). However, some vocal nonlearners, including monkeys, can be trained to generate predictive synchronized movements for immediate rewards (Takeya et al., 2017; Gamez et al., 2018).

As evidenced by the inability to ride excessively slow rhythms, predictive synchronization has a time limit. This appears to be 2–3 s for tapping (Tokushige et al., 2018; Mates et al., 1994) and approximately 1.5 s for eye movements (Shelhamer & Joiner, 2003; Takeya et al., 2018). In monkeys, this is close to the aforementioned time limit at which the rate of increase in neuronal activity in the cerebellum can regulate the timing of self-initiated movements (Fig. 2d). The cerebellum is also involved in eye-blink conditioning with a temporal limit of ~2 s (Medina & Mauk, 2000; Mauk & Buonomano, 2004), and the cerebellar cortex is known to be important in controlling its timing (Perrett et al., 1993). These findings suggest that the cerebellum plays a role in the temporal control of synchronized movements. Indeed, neural activity in the cerebellum, along with motor-related cortical areas and the basal ganglia, increases during motor synchrony (Chauvigne et al., 2014; Aso et al., 2010; Lee et al., 2016; Witt et al., 2008), and damage to the cerebellum increases the temporal variability of discrete periodic movements (Spencer et al., 2003).

Okada et al. (2022) trained monkeys to perform synchronized saccades to alternating left and right targets at regular intervals (Fig. 5a) and examined neuronal activity in the dentate nucleus of the cerebellum. Many neurons showed increased activity before eye movements, half of

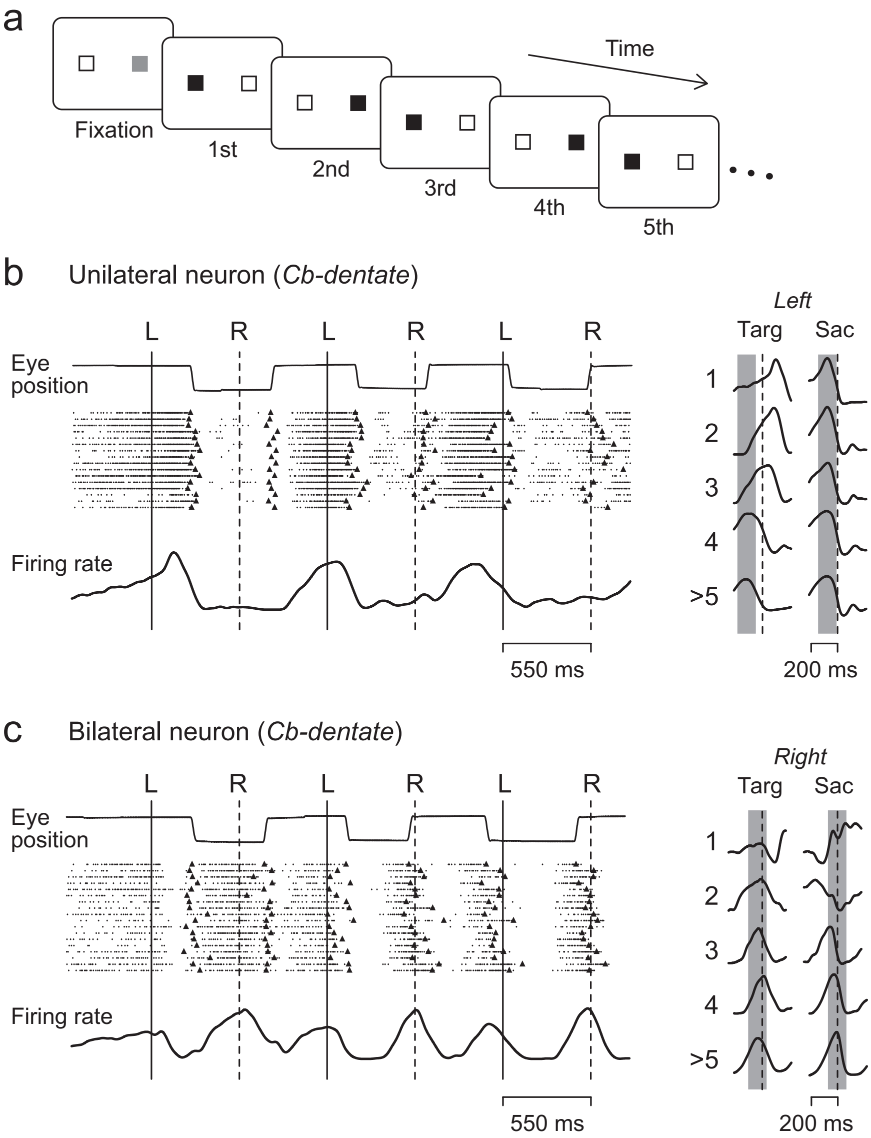


Fig. 5 Neuronal activity in the cerebellum during synchronized eye movements. (**a**) Saccade targets were presented alternately at landmark locations (white squares, 14° apart) for 400, 550, or 700 ms (constant in each trial). Monkeys were rewarded for every three predictive saccades. (**b**) Response of a unilateral neuron to the early stimulus sequence. Data are aligned with the target onset (solid and dashed vertical lines) in trials with a 550-ms stimulus interval. The black symbol on each raster line indicates saccade onset. Note that neuronal activity consistently precedes the leftward saccade. The right panel shows the traces of the mean firing rate aligned either to the target onset or saccade initiation (vertical dashed lines). Number on the left indicates the stimulus sequence. Shadings represent the 100 ms window centered at the peak activity during the later stimulus sequence. (**c**) Response of a bilateral neuron. Note that even during the initial few cycles of target presentation, the peak firing rate of the neuron roughly coincides with the target onset, whereas early saccades lag behind the target onset. (Adapted from Okada et al. (2022) under the terms of the Creative Commons Attribution 4.0 International License (CC-BY))

which had strong directional modulation, exhibiting a ramp-like preparatory activity mostly for ipsilateral synchronized saccades. However, these neurons also showed strong presaccade activity even when the animals made a series of reactive saccades to randomly timed targets. This suggests that they are involved in motor control, independent of the stimulus rhythm. In contrast, the remaining half of the presaccade neurons were active for eye movements in both directions, with enhanced activity during synchronized saccades compared with reactive saccades. These results indicate that neurons that are active in bilateral eye movements are important for motor synchrony.

Figure 5b, c illustrates examples of these neurons. The Unilateral neuron exhibited strong activity before saccades to the left, and the initial peak of activity and saccades occurred after the target onset. However, the timing of peak activity gradually advanced as saccade latency became shorter during the sequence. When neuronal activity was aligned with target onset or saccades in the order on each trial, the timing of peak activity varied with respect to the target onset, but consistently preceded saccades (Fig. 5b, right panel). This finding indicates that this unilateral neuron is related to motor control. In contrast, the representative bilateral neuron exhibited increased activity for saccades in both directions, and the timing of peak activity changed only slightly, although the saccade latency gradually shortened (Fig. 5c). When averaging neuronal activity in the order of target or saccade, the activity was more consistent with target onset, indicating that the neuron was activated in anticipation of stimulus timing rather than movement. Another class of neurons in the cerebellar nucleus showed increased activity immediately after saccades. About half of these neurons showed a significant correlation with saccade latency (or the time difference between target appearance and saccade), indicating that they detected temporal errors in synchronized saccades (Okada et al., 2022).

These results suggest that neurons in the cerebellar nuclei carry the signals necessary to predict the timing of alternating targets, adjust the timing of movements, and detect temporal errors to update the internal model of periodic events. Since these neurons were found in the output node of the cerebellum, the information was transmitted to the brainstem and thalamocortical pathways. Signals related to eye movements might be sent to the midbrain superior colliculus to regulate saccade timing (May et al., 1990; Prevosto et al., 2017), and those related to temporal errors might be sent to the inferior olive in the medulla to induce learning (De Zeeuw et al., 1998). Alternatively, these signals, along with those related to the stimulus timing, might be sent to different areas in the cerebral cortex via the thalamus. Future studies are needed to clarify how information is integrated within multiple cerebrocerebellar loops, including the dentate nucleus, during synchronized saccades. In addition, since the basal ganglia are also known to be involved in synchronized movements (Chauvigne et al., 2014; Witt et al., 2008; Hove et al., 2013; Rao et al., 1997; Bartolo et al., 2014), it is important to examine how the signals in the cortico-basal ganglia and the cortico-cerebellar loops cooperate and interact for synchronized movements. Signals processed in these subcortical loops must be integrated in the parietal and frontal cortices, which is key to a comprehensive understanding of the neural mechanisms of motor synchronization.

Temporal Limit of Sensory Prediction for Rhythmic Events

The premise of synchronized movements is to detect the periodicity of sensory events and predict their timing. As mentioned earlier, synchronized movement is difficult when the tempo is too slow, but it is easy to imagine that synchronization can also be impossible when the tempo is too fast. In the case of synchronized movements, this is largely due to the constraints of the motor system, but there is a paucity of information on the temporal limit of purely sensory rhythms. This has been examined using the missing (omission) oddball detection paradigm in healthy individuals (Ohmae & Tanaka, 2016). In this task, participants were asked to press a button as fast as possible when they detected the omission of isochronous repetitive auditory stimulus (Fig. 6a). To detect stimulus omission, it is necessary to predict the timing of the next stimulus. Figure 6b shows that the relationship between reaction time and the interstimulus interval (or stimulus tempo) differed after ~250 ms (4 Hz, vertical dashed line), suggesting that different mechanisms may underlie the detection of stimulus omissions for different tempos.

When listening to isochronous sound sequences with a short interstimulus interval, it is impossible to predict the timing of each stimulus in advance; rather, there is an impression of a

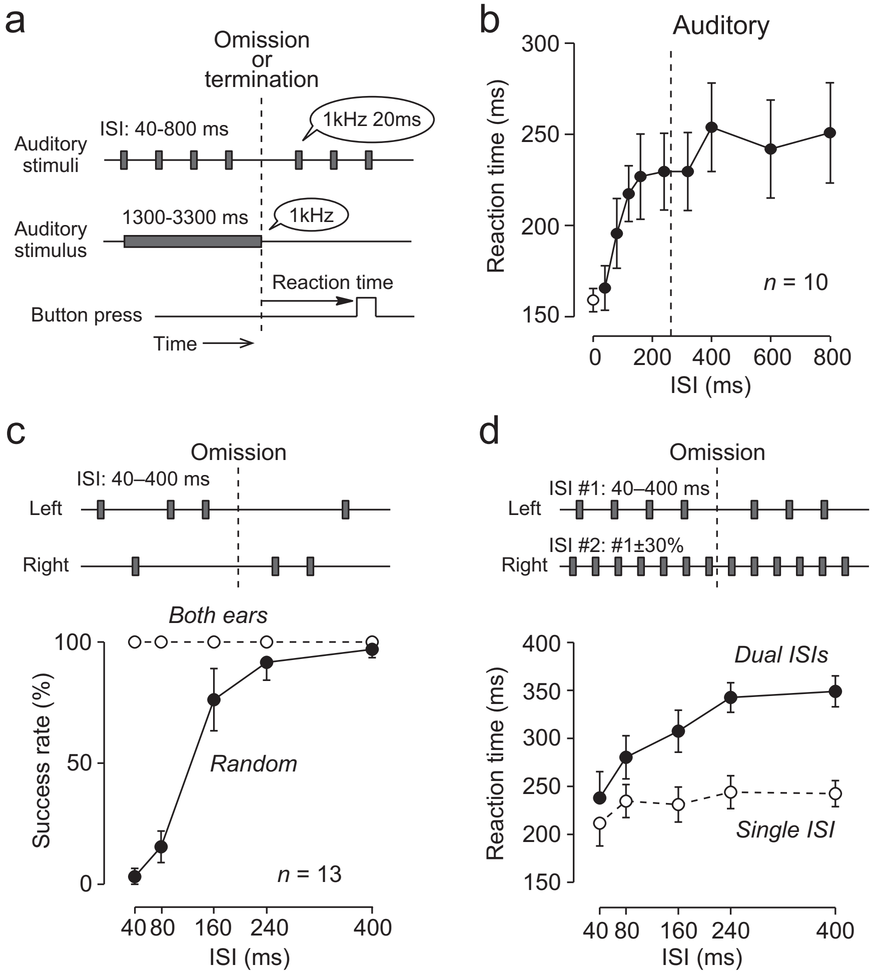


Fig. 6 Temporal limit of sensory prediction for rhythmic events. (**a**) Participants were asked to press the button as soon as possible in response to the omission of one repetitive auditory stimulus (upper) or the end of a continuous tone (middle). The interstimulus interval (ISI) ranged from 40 to 800 ms, and each repetitive sound lasted 20 ms. (**b**) Reaction time for different ISIs. Data for the continuous sound are shown at zero ISI (white circle). Error bars indicate 95% confidence intervals. Note that the reaction time highly depends on the ISI for the intervals shorter than 250 ms (vertical dashed line) but is relatively constant for longer intervals. (**c**) Each repetitive stimulus was randomly presented to either ear with a constant ISI (40–400 ms). Participants were asked to detect stimulus omission as accurately as possible. As a control, each sound was presented to both ears. The graph represents the proportions of correct trials in the random monaural (black circle) and binaural (white) conditions. (**d**) Repetitive stimuli with different ISIs were presented simultaneously to different ears. Stimulus omission occurred randomly in either ear. Black and white circles indicate the data for the dual- and single-ISI conditions, respectively. (Adapted from Ohmae and Tanaka (2016) under the terms of the Creative Commons Attribution 4.0 International License (CC-BY))

momentary interruption of a constant sound stream (Supplementary Audio files in (Ohmae & Tanaka, 2016)). This may be because auditory responses to regularly repeated sounds at a fast tempo are temporally grouped, and missing stimuli are detected as deviations from the steady state of neural activity. In fact, the data for responses to a pause in continuous sound are in line with the relationship between reaction time and stimulus tempo if the interstimulus interval of the continuous sound is assumed to be zero (Fig. 6b, white circle). When each sound in a sequence was presented randomly to each ear to prevent temporal grouping, the detection of stimulus omission became difficult only for stimulus intervals shorter than 250 ms (Fig. 6c). The same was true when sounds of different frequencies were presented at regular intervals (Ohmae & Tanaka, 2016).

On the other hand, the detection of stimulus omission for slower rhythms must rely on the temporal prediction of the next stimulus and may require higher-order cognitive processes that can be interfered with under dual-task conditions. To test this hypothesis, two sequences of sounds at different tempos were presented simultaneously to different ears so that attentional resources for temporal prediction were separately allocated to both ears. Even in this condition, the participants reliably detected the sound omission that occurred in either ear. However, compared to the response to a single sequence, the reaction time in the dual-task condition was prolonged, mainly for slower tempos (Fig. 6d). Taken together, these results suggest that a time window of approximately 4 Hz or slower is necessary to integrate auditory inputs from different sources to predict the next stimulus timing. Since a similar relationship between reaction time and interstimulus interval was also observed in sequences of visual and tactile stimuli (Ohmae & Tanaka, 2016), a 4-Hz limit of temporal prediction might be common to different sensory modalities. The 3–4 Hz limit for the integration of temporal information from different sources has also been demonstrated in simultaneity detection between different sensory modalities (Fujisaki & Nishida, 2010).

Subcortical Signals Underlying Rhythm Perception

Using a similar oddball detection paradigm with a visual metronome, neuronal activity during rhythm perception in the cerebellar nuclei, striatum, and motor thalamus has been examined in monkeys. In this task, a saccade target was initially presented either on the left or right side of the central fixation point (Fig. 7a). While maintaining eye fixation, visual stimuli surrounding the fixation point were repeatedly presented at an interstimulus interval that was constant in each trial but varied from trial to trial in the range of 100 to 600 ms. The animals were trained to make a saccade in response to an unexpected (random) omission of the repetitive stimulus to obtain a liquid reward (missing oddball condition). To detect stimulus omission, they were required to predict the timing of the next stimulus. In a variant of this task with a different color of the fixation point, the animals were required to detect the occurrence of color change in the repetitive stimulus (color oddball condition).

During this task, neurons in the posterior part of the cerebellar dentate nucleus, the head of the caudate nucleus, and the motor thalamus have been shown to exhibit periodic activity in the absence of movement (Fig. 7b) (Ohmae et al., 2013; Kameda et al., 2019, 2023; Matsuyama & Tanaka, 2021; Uematsu & Tanaka, 2022). Most of these neurons do not respond to the first few stimuli in sequence, but as repetition progresses, the firing modulation gradually increases, indicating neural entrainment to the stimulus rhythm. After several repetitions, the response to each stimulus reached a plateau, and at this moment the magnitude of the firing modulation was proportional to the interstimulus interval (Fig. 7c). Some neurons in the striatum also show preferences for a specific stimulus tempo (Kameda et al., 2019), indicating a tuned representation of interval timing (Heron et al., 2012; Protopapa et al., 2019). Importantly, periodic activity greatly decreases under the color oddball condition where temporal prediction is not needed but saccades are required, indicating that these activities are under top-down control and are involved in

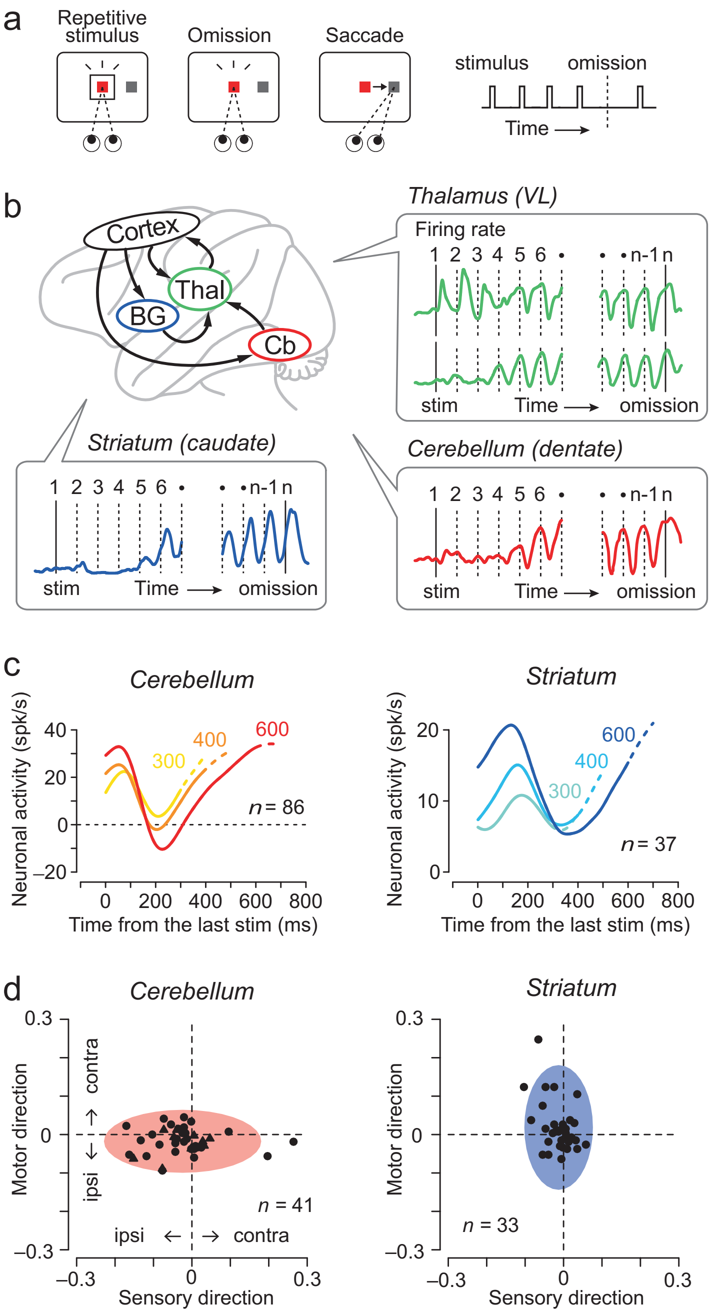


Fig. 7 Subcortical neuronal activities entrained to periodic visual stimuli. (**a**) In the missing oddball paradigm, a saccade target appeared horizontally during central fixation, and then a brief stimulus surrounding the fixation point was presented repeatedly at a fixed interstimulus interval (ISI). Animals were trained to make a saccade in response to the stimulus omission. (**b**) The colored solid line in the balloon represents the time course of neuronal activity recorded from the cerebellar dentate nucleus (red), striatum (caudate nucleus, blue), and motor thalamus (green), which showed periodic activity during eye fixation. In each panel, data are aligned with either the first stimulus (left) or stimulus omission (right). Note that most neurons gradually increase their periodic firing modulation with stimulus repetition. Also note that the timing of peak activity differs among brain regions. (**c**) The time courses of population activity in the cerebellum and striatum for three different ISIs aligned with the stimulus just

temporal processing. Indeed, neuronal activity at the time of stimulus omission correlates with reaction time in all three brain regions (Ohmae et al., 2013; Kameda et al., 2019; Matsuyama & Tanaka, 2021). Furthermore, electrical stimulation delivered to the cerebellar nuclei or the striatum shortened the reaction time to stimulus omission and inactivation prolonged it (Ohmae et al., 2013; Kameda et al., 2019; Uematsu et al., 2017). These effects disappeared under the color oddball condition or when the interstimulus interval was less than 200 ms, suggesting that these signals are important for predicting the timing of the next stimulus in advance.

Although the overall characteristics of neuronal activity are common in the cerebellum, striatum, and thalamus, the time course of periodic activity differs at each recording site. In the cerebellar dentate nucleus, most neurons show peak activity at the time of stimulus presentation, regardless of the stimulus interval, and cease firing during eye movements. Neuronal activity sharply decreases after the presentation of each stimulus, but its amplitude and the time course of recovery differs between stimulus intervals, resulting in anticipatory activity that peaks at the time of the next stimulus (Fig. 7b, c). In the striatum (caudate nucleus), on the other hand, neuronal activity increases after stimulus presentation and its peak remains constant at approximately 150–200 ms regardless of the stimulus interval, and the firing rate at the time of the next stimulus differed across intervals and the next stimulus timing cannot be accurately predicted from the time course of the population activity (Kameda et al., 2019). Furthermore, it has recently been shown that periodic neuronal activity in the cerebellum is modulated by the location of repetitive stimuli but not by the direction of prepared movements (Fig. 7d). Conversely, neuronal activity in the striatum is modulated by the direction of subsequent saccades but not by the location of the repetitive stimulus. These results suggest that the cerebellum is involved in the temporal prediction of sensory events while the striatum is involved in motor preparation during the task (Kameda et al., 2023). In the ventrolateral and the adjacent nuclei of the thalamus, which receive inputs from the cerebellum, many neurons exhibit anticipatory activity like those of the cerebellar nucleus. Some thalamic neurons reverse the direction of their response to each stimulus during repetition, showing the first reactive and later anticipatory activity (Fig. 7b). Since it has been suggested that the motor thalamus integrates information from subcortical and cortical sources (Galvan et al., 2016; Suzuki et al., 2021), the thalamic neurons showing such phase transitions may reflect signals from multiple input sources (Matsuyama & Tanaka, 2021; Sieveritz & Raghavan, 2021).

Thus, although subcortical networks are involved in rhythm processing, neural signals vary widely among brain regions, and further research is needed to clarify their specific roles. In particular, the time course of neural activity in the cerebellar nuclei during oddball tasks is similar to the activity that predicts target onset during saccade synchronization and may represent an internal model of periodic sensory events. Previous studies have shown that the cerebellum generates and updates internal models that predict sensory events even in the absence of movement (Cerminara et al., 2009; O'Reilly et al., 2008; Roth et al., 2013). Additionally, the time course of neuronal activity in the cerebellar nuclei during the oddball tasks closely resembles the time course of beta coherence in the cerebellum and the cerebral cortex during passive listening to isochronous sounds shown in humans (Fujioka et al., 2012), suggesting that the corticocerebellar loops may be essential for the generation of predictive signals. In relation to this, recent evidence suggests that the cerebellum is involved in temporal attention (Coull & Nobre,

◄

Fig. 7 (continued) before omission. Dashed lines indicate the data during 100 ms following the stimulus omission. (**d**) Comparison of sensory and motor components of neuronal activity in the cerebellum and striatum. The repetitive visual stimulus and saccade target were independently placed either left or right of the fixation point, and the modulation of periodic activity in individual neurons due to stimulus location was evaluated using the generalized linear model. Each dot represents beta coefficient of each neuron calculated for the location of the repetitive stimulus (sensory direction) or saccade target (motor direction). (Panels (**c**) and (**d**) are adapted from Kameda et al. (2019, 2023))

1998; Nobre & van Ede, 2018; Breska & Ivry, 2021), which appears to be controlled by low-frequency neural oscillations in the sensory cortex during rhythm perception (Lakatos et al., 2008). Since attention is likely allocated to predicted stimuli, the cortico-subcortical network, including the cerebellum, may play a role in temporal attention during rhythm processing.

Relation to Other Studies

A well-known hypothesis in the study of temporal processing is that the basal ganglia are involved in beat-based or relative timing, whereas the cerebellum is involved in measuring single absolute time intervals (Teki et al., 2011a; Grube et al., 2010). In fact, it has been clearly demonstrated that poor performance on timing tasks in Parkinson's disease and cerebellar degeneration is double dissociated when processing periodic and single time intervals (Breska & Ivry, 2018). However, many previous studies in experimental animals have employed behavioral tasks that require the measurement of single time intervals to reveal the involvement of the basal ganglia (Lee & Assad, 2003; De Corte et al., 2019; Wang et al., 2018; Mello et al., 2015). Functional imaging studies have shown increased neural activity in the cerebellum and basal ganglia during both single and periodic time processing (Wiener et al., 2010), suggesting that these two subcortical systems interact strongly (Petter et al., 2016; Teki et al., 2011b). As described above, time-specific neuronal activity has been found in the cerebellum and striatum of monkeys during the self-timing task and the missing oddball detection task, which require single and periodic time processing, respectively. In both cases, manipulation of neuronal activity at the recording sites by electrical stimulation or pharmacological inactivation alters the task performance.

However, it is important to review these data in terms of relative and absolute timing to further understand the role of these subcortical regions. For example, the time course of preparatory activity in the striatum during self-timed tasks has been shown to be temporally scaled in the range of a few hundred milliseconds to seconds (Kunimatsu et al., 2018; Wang et al., 2018; Mello et al., 2015). This can be viewed as the neural activity encoding the passage of time relative to the length of time being measured. Similar temporal scaling has recently been shown for midbrain dopamine neurons (Hamilos et al., 2021) and hippocampus (Shimbo et al., 2021). In contrast, neurons in the cerebellum are also temporally scaled when the time length is less than a second (Okada et al., 2022; Ohmae et al., 2017), but at longer time intervals, they do not have information relative to the intended time length to be measured. Furthermore, in the missing oddball task, neuronal activity in the cerebellar nuclei for each repetitive stimulus appears to encode the absolute elapsed time from the previous stimulus, rather than the relative timing to the previous beat (Ohmae et al., 2013). This might mean that the cerebellum locally measures the time between immediately preceding events even when events have a periodic time structure.

Nevertheless, monkeys are capable of maintaining internalized rhythm, as they continue periodic saccades when targets are presented in conjunction with eye movements (error-clamp condition) during the synchronized saccade task (Takeya et al., 2017). Similarly, monkeys can continue predictive tapping even after periodic external stimuli have disappeared (Gamez et al., 2018) and can covertly shift the direction of motor preparation in the internalized rhythms (Cadena-Valencia et al., 2018; Garcia-Garibay et al., 2016). Further investigation is needed to understand how periodic time is encoded under these conditions, and how the roles of the cerebellum and basal ganglia are different.

Summary and Conclusions

This chapter presents recent studies in nonhuman primates that have examined neural mechanisms during the measurement of single and periodic time intervals. The series of studies described earlier in this chapter focuses on pupil diameter and subcortical preparatory activity during a self-timed saccade task and reveals two mechanisms that regulate movement timing. The trial-to-trial variation in the produced time interval under the same conditions correlates well with pupil diameter (Suzuki et al., 2016), which reflects internal states such as attention and arousal level. When producing time intervals of 1.2 s or less, these timing variations can be explained by the rate of increasing activity in the cerebellum and motor thalamus, and inactivation of these areas delays self-timing. When measuring longer intervals, preparatory activity in the cerebellum begins approximately 500 ms before movement, and the stochastic variability in self-timing correlates well with the time of onset rather than with the slope of increasing activity (Fig. 2d, Ohmae et al., 2017).

On the other hand, when the length of time to be measured is explicitly indicated, the rate of increase in preparatory activity in the striatum changes accordingly. In the striatum, the preparatory activity persists throughout the delay period, but the effects of intertrial variations in self-timing appear only immediately before movement, following changes in the cerebellum (Fig. 3, Kunimatsu et al., 2018). This suggests that fluctuations in cerebellar activity are primarily responsible for the stochastic variation in timing. Because striatal inactivation changes self-timing regardless of the length of the mandatory delay, the striatum may measure elapsed time over the entire range from a few hundred milliseconds to several seconds.

In the striatum, the magnitude of the visual response to the cue was negatively correlated with the length of time to be measured, and the power of the low-frequency component before the cue was positively correlated (Fig. 4, Suzuki & Tanaka, 2019). This indicates that the length of the intended timing changes the state of the network, including the striatum, which may lead to changes in the time course of the preparatory activity. In Parkinson's disease, in which beta oscillation within the basal ganglia circuitry is abnormally enhanced, the central tendency is stronger when different time lengths are generated in each trial (Malapani et al., 1998). This may be because of the inability to flexibly change the state of the network according to the intended time length. In contrast, the trial-to-trial variation in self-timing is correlated with pupil diameter, which is known to reflect the activity of the central noradrenergic system. Strong noradrenergic innervation from the brainstem to the cerebellum may contribute in part to stochastic changes in self-timing. Thus, when producing time intervals, (1) intentional signals that depend on the length of time to be measured and, (2) stochastic signals that fluctuate according to the internal state, exist in the brain which might control self-timing by modulating neuronal activity in the basal ganglia and cerebellum, respectively.

The second half of this chapter discusses recent research on periodic timing. Stimuli that are regularly repeated on the order of hundreds of milliseconds produce rhythm perception, often accompanied by synchronized movements. Predictive synchronized movements are thought to be specific to vocal-learning species (Patel et al., 2009), but recent studies have shown that monkeys exhibit this ability when given immediate rewards (Takeya et al., 2017; Gamez et al., 2018). They can make predictive eye movements synchronized to alternating visual stimuli if the stimulus interval is 0.4 to 1.2 s. As movement becomes reactive when the stimulus interval is 1.8 s, the time limit for predictive synchronization may be approximately 1.5 s (Takeya et al., 2018).

When monkeys perform synchronized saccades, there are three types of neurons in the cerebellar nuclei that exhibit periodic activity related to movement timing, predicted target timing, and synchronization errors (Fig. 5, Okada et al., 2022). Of these, only neurons that show the peak activity at the target appearance is more active than during reactive saccades and may be particularly important in synchronizing movements. These neurons were active regardless of the target location, representing an internal model of periodic stimulation. To achieve synchronized movements, it is necessary to generate an internal model for periodic stimuli, adjust the movement timing, and update the internal model by detecting errors. The cerebellum contains all this information and may send signals to the brainstem and thalamocortical pathways, serving a part of the global network that enables motor synchrony. As it is also well known that the basal ganglia are involved in synchronized movements, it is necessary to compare the roles of these subcortical regions in future studies.

When perceiving rhythm, we predict the timing of the next stimulus and focus our attention on that moment. Its time limit has been investigated using the missing oddball paradigm (Fig. 6, Ohmae & Tanaka, 2016). When detecting an omission of isochronously repeated sounds, the reaction time changes after about 0.25 s (4 Hz). At faster tempos, it is impossible to integrate temporal information from the opposite ear or from sounds of different frequencies. Conversely, at slower tempos, the reaction times to stimulus omission were significantly prolonged under dual-task condition. Taken together with the results of the synchronized movements, accurate temporal prediction is possible for periodic stimuli in the range of 0.25 to 1.5 s (0.6 to 4 Hz).

Periodic neuronal activity has been recorded from the cerebellum (Ohmae et al., 2013), striatum (Kameda et al., 2019), and motor thalamus (Matsuyama & Tanaka, 2021) in monkeys performing a similar task using a visual metronome (Fig. 7). At each site, the magnitude of neuronal modulation increased gradually with repeated stimuli, reaching a plateau within a few seconds. Under these conditions, the size of the response to each stimulus depended on the stimulus tempo. The time course of periodic neuronal activity varies between brain regions; in particular, many neurons in the cerebellar nuclei exhibit anticipatory activity that peaks at the stimulus timing and the magnitude of periodic modulation depends on the location of the repetitive stimulus (Kameda et al., 2023). The time course of neuronal activity in primates resembles the time course of beta-band coherence seen in the cortex and the cerebellum during passive listening to isochronous sound in humans (Fujioka et al., 2012), which may represent an internal model predicting the stimulus timing. Manipulation of neuronal activity in the cerebellum and striatum changes the reaction time of omission detection, suggesting that these areas are involved in periodic temporal processing. In the future, the neural mechanisms of the cortico-cerebellar and cortico-basal ganglia pathways in rhythm perception should be clarified.

Acknowledgments This work was supported partly by a Grant-in-Aid for Scientific Research on Innovative Areas "Chronogenesis" from the Ministry of Education, Sports, Science, and Technology of Japan (MEXT), and the Core Research for Evolutionary Science and Technology (CREST) of the Japan Science and Technology Agency (JST).

References

Allman, M. J., & Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain, 135*(Pt 3), 656–677.

Ashmore, R. C., & Sommer, M. A. (2013). Delay activity of saccade-related neurons in the caudal dentate nucleus of the macaque cerebellum. *Journal of Neurophysiology, 109*(8), 2129–2144.

Aso, K., Hanakawa, T., Aso, T., & Fukuyama, H. (2010). Cerebro-cerebellar interactions underlying temporal information processing. *Journal of Cognitive Neuroscience, 22*(12), 2913–2925.

Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience, 28*, 403–450.

Baldo, B. A., Daniel, R. A., Berridge, C. W., & Kelley, A. E. (2003). Overlapping distributions of orexin/hypocretin- and dopamine-beta-hydroxylase immunoreactive fibers in rat brain regions mediating arousal, motivation, and stress. *The Journal of Comparative Neurology, 464*(2), 220–237.

Bartolo, R., Prado, L., & Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *The Journal of Neuroscience, 34*(11), 3910–3923.

Betancourt, A., Pérez, O., Gámez, J., Mendoza, G., & Merchant, H. (2023). Amodal population clock in the primate medial premotor system for rhythmic tapping. *Cell Reports, 42*, 113234.

Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease. *Proceedings of the National Academy of Sciences of the United States of America, 115*(48), 12283–12288.

Breska, A., & Ivry, R. B. (2021). The human cerebellum is essential for modulating perceptual sensitivity based on temporal expectations. *eLife, 10*, e66743.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience, 6*(10), 755–765.

Cadena-Valencia, J., Garcia-Garibay, O., Merchant, H., Jazayeri, M., & de Lafuente, V. (2018). Entrainment and maintenance of an internal metronome in supplementary motor area. *eLife, 7*, e38983.

Cerminara, N. L., Apps, R., & Marple-Horvat, D. E. (2009). An internal model of a moving visual target in the lateral cerebellum. *The Journal of Physiology, 587*(Pt 2), 429–442.

Chauvigne, L. A., Gitau, K. M., & Brown, S. (2014). The neural basis of audiomotor entrainment: An ALE meta-analysis. *Frontiers in Human Neuroscience, 8*, 776.

Chiba, A., Morita, K., Oshio, K. I., & Inase, M. (2021). Neuronal activity in the monkey prefrontal cortex during a duration discrimination task with visual and auditory cues. *Scientific Reports, 11*(1), 17520.

Coull, J. T., & 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. *The Journal of Neuroscience, 18*(18), 7426–7435.

Coull, J. T., Cheng, R. K., & Meck, W. H. (2011). Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology, 36*(1), 3–25.

Dacre, J., Colligan, M., Clarke, T., Ammer, J. J., Schiemann, J., Chamosa-Pino, V., et al. (2021). A cerebellar-thalamocortical pathway drives behavioral context-dependent movement initiation. *Neuron, 109*(14), 2326–38 e8.

De Corte, B. J., Wagner, L. M., Matell, M. S., & Narayanan, N. S. (2019). Striatal dopamine and the temporal control of behavior. *Behavioural Brain Research, 356*, 375–379.

De Zeeuw, C. I., Simpson, J. I., Hoogenraad, C. C., Galjart, N., Koekkoek, S. K., & Ruigrok, T. J. (1998). Microcircuitry and function of the inferior olive. *Trends in Neurosciences, 21*(9), 391–400.

Fujioka, T., Trainor, L. J., Large, E. W., & Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *The Journal of Neuroscience, 32*(5), 1791–1802.

Fujisaki, W., & Nishida, S. (2010). A common perceptual temporal limit of binding synchronous inputs across different sensory attributes and modalities. *Proceedings of the Biological Sciences, 277*(1692), 2281–2290.

Galvan, A., Hu, X., Smith, Y., & Wichmann, T. (2016). Effects of optogenetic activation of corticothalamic terminals in the motor thalamus of awake monkeys. *The Journal of Neuroscience, 36*(12), 3519–3530.

Gamez, J., Yc, K., Ayala, Y. A., Dotov, D., Prado, L., & Merchant, H. (2018). Predictive rhythmic tapping to isochronous and tempo changing metronomes in the nonhuman primate. *Annals of the New York Academy of Sciences, 1423*(1), 396–414.

Gamez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biology, 17*(4), e3000054.

Gámez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biology, 17*, e3000054.

Garcia-Garibay, O., Cadena-Valencia, J., Merchant, H., & de Lafuente, V. (2016). Monkeys share the human ability to internally maintain a temporal rhythm. *Frontiers in Psychology, 7*, 1971.

Grube, M., Cooper, F. E., Chinnery, P. F., & Griffiths, T. D. (2010). Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. *Proceedings of the National Academy of Sciences of the United States of America, 107*(25), 11597–11601.

Hamilos, A. E., Spedicato, G., Hong, Y., Sun, F., Li, Y., & Assad, J. A. (2021). Slowly evolving dopaminergic activity modulates the moment-to-moment probability of reward-related self-timed movements. *eLife, 10*, e62583.

Hammond, C., Bergman, H., & Brown, P. (2007). Pathological synchronization in Parkinson's disease: Networks, models and treatments. *Trends in Neurosciences, 30*(7), 357–364.

Heron, J., Aaen-Stockdale, C., Hotchkiss, J., Roach, N. W., McGraw, P. V., & Whitaker, D. (2012). Duration channels mediate human time perception. *Proceedings of the Biological Sciences, 279*(1729), 690–698.

Honma, M., Kuroda, T., Futamura, A., Shiromaru, A., & Kawamura, M. (2016). Dysfunctional counting of mental time in Parkinson's disease. *Scientific Reports, 6*, 25421.

Hove, M. J., Fairhurst, M. T., Kotz, S. A., & Keller, P. E. (2013). Synchronizing with auditory and visual rhythms: An fMRI assessment of modality differences and modality appropriateness. *NeuroImage, 67*, 313–321.

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences, 12*(7), 273–280.

Janssen, P., & Shadlen, M. N. (2005). A representation of the hazard rate of elapsed time in macaque area LIP. *Nature Neuroscience, 8*(2), 234–241.

Jazayeri, M., & Shadlen, M. N. (2015). A neural mechanism for sensing and reproducing a time interval. *Current Biology, 25*(20), 2599–2609.

Jones, B. E., & Yang, T. Z. (1985). The efferent projections from the reticular formation and the locus coeruleus studied by anterograde and retrograde axonal transport in the rat. *The Journal of Comparative Neurology, 242*(1), 56–92.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron, 89*(1), 221–234.

Kameda, M., Ohmae, S., & Tanaka, M. (2019). Entrained neuronal activity to periodic visual stimuli in the primate striatum compared with the cerebellum. *eLife, 8*, e48702.

Kameda, M., Niikawa, K., Uematsu, A., & Tanaka, M. (2023). Sensory and motor representations of internalized rhythms in the cerebellum and basal ganglia. *Proceedings of the National Academy of Sciences of the United States of America, 120*, e2221641120.

Kotz, S. A., Ravignani, A., & Fitch, W. T. (2018). The evolution of rhythm processing. *Trends in Cognitive Sciences, 22*(10), 896–910.

Kunimatsu, J., & Tanaka, M. (2012). Alteration of the timing of self-initiated but not reactive saccades by electrical stimulation in the supplementary eye field. *The European Journal of Neuroscience, 36*(9), 3258–3268.

Kunimatsu, J., & Tanaka, M. (2016). Striatal dopamine modulates timing of self-initiated saccades. *Neuroscience, 337*, 131–142.

Kunimatsu, J., Suzuki, T. W., Ohmae, S., & Tanaka, M. (2018). Different contributions of preparatory activity in the basal ganglia and cerebellum for self-timing. *eLife, 7*, e35676.

Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science, 320*(5872), 110–113.

Lake, J. I., LaBar, K. S., & Meck, W. H. (2016). Emotional modulation of interval timing and time perception. *Neuroscience and Biobehavioral Reviews, 64*, 403–420.

Lee, I. H., & Assad, J. A. (2003). Putaminal activity for simple reactions or self-timed movements. *Journal of Neurophysiology, 89*(5), 2528–2537.

Lee, S. M., Peltsch, A., Kilmade, M., Brien, D. C., Coe, B. C., Johnsrude, I. S., et al. (2016). Neural correlates of predictive saccades. *Journal of Cognitive Neuroscience, 28*(8), 1210–1227.

Leon, M. I., & Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron, 38*(2), 317–327.

Lewis, P. A., & Miall, R. C. (2003). Brain activation patterns during measurement of sub- and supra-second intervals. *Neuropsychologia, 41*(12), 1583–1592.

Maimon, G., & Assad, J. A. (2006). A cognitive signal for the proactive timing of action in macaque LIP. *Nature Neuroscience, 9*(7), 948–955.

Malapani, C., Rakitin, B., Levy, R., Meck, W. H., Deweer, B., Dubois, B., et al. (1998). Coupled temporal memories in Parkinson's disease: A dopamine-related dysfunction. *Journal of Cognitive Neuroscience, 10*(3), 316–331.

Mates, J., Muller, U., Radil, T., & Poppel, E. (1994). Temporal integration in sensorimotor synchronization. *Journal of Cognitive Neuroscience, 6*(4), 332–340.

Matsuyama, K., & Tanaka, M. (2021). Temporal prediction signals for periodic sensory events in the primate central thalamus. *The Journal of Neuroscience, 41*(9), 1917–1927.

Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. *Annual Review of Neuroscience, 27*, 307–340.

May, P. J., Hartwich-Young, R., Nelson, J., Sparks, D. L., & Porter, J. D. (1990). Cerebellotectal pathways in the macaque: Implications for collicular generation of saccades. *Neuroscience, 36*(2), 305–324.

Medina, J. F., & Mauk, M. D. (2000). Computer simulation of cerebellar information processing. *Nature Neuroscience, 3*(Suppl), 1205–1211.

Meirhaeghe, N., Sohn, H., & Jazayeri, M. (2021). A precise and adaptive neural mechanism for predictive temporal processing in the frontal cortex. *Neuron, 109*, 2995–3011.

Mello, G. B., Soares, S., & Paton, J. J. (2015). A scalable population code for time in the striatum. *Current Biology, 25*(9), 1113–1122.

Mendoza, G., Mendez, J. C., Perez, O., Prado, L., & Merchant, H. (2018). Neural basis for categorical boundaries in the primate pre-SMA during relative categorization of time intervals. *Nature Communications, 9*(1), 1098.

Merchant, H., & Averbeck, B. B. (2017). The computational and neural basis of rhythmic timing in medial premotor cortex. *The Journal of Neuroscience, 37*(17), 4552–4564.

Merchant, H., & Honing, H. (2013). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience, 7*, 274.

Merchant, H., Zarco, W., Perez, O., Prado, L., & Bartolo, R. (2011). Measuring time with different neural chronometers during a synchronization-continuation task. *Proceedings of the National Academy of Sciences of the United States of America, 108*(49), 19784–19789.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience, 36*, 313–336.

Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., & Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nature Neuroscience, 12*(4), 502–507.

Murakami, M., Vicente, M. I., Costa, G. M., & Mainen, Z. F. (2014). Neural antecedents of self-initiated actions in secondary motor cortex. *Nature Neuroscience, 17*(11), 1574–1582.

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping, 35*(8), 4140–4154.

Narain, D., Remington, E. D., Zeeuw, C. I., & Jazayeri, M. (2018). A cerebellar mechanism for learning prior distributions of time intervals. *Nature Communications, 9*, 469.

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience, 19*(1), 34–48.

Ohmae, S., & Tanaka, M. (2016). Two different mechanisms for the detection of stimulus omission. *Scientific Reports, 6*, 20615.

Ohmae, S., Uematsu, A., & Tanaka, M. (2013). Temporally specific sensory signals for the detection of stimulus omission in the primate deep cerebellar nuclei. *The Journal of Neuroscience, 33*(39), 15432–15441.

Ohmae, S., Kunimatsu, J., & Tanaka, M. (2017). Cerebellar roles in self-timing for sub- and supra-second intervals. *The Journal of Neuroscience, 37*(13), 3511–3522.

Okada, K. I., Takeya, R., & Tanaka, M. (2022). Neural signals regulating motor synchronization in the primate deep cerebellar nuclei. *Nature Communications, 13*(1), 2504.

O'Reilly, J. X., Mesulam, M. M., & Nobre, A. C. (2008). The cerebellum predicts the timing of perceptual events. *The Journal of Neuroscience, 28*(9), 2252–2260.

Parker, K. L., Chen, K. H., Kingyon, J. R., Cavanagh, J. F., & Narayanan, N. S. (2014). D1-dependent 4 Hz oscillations and ramping activity in rodent medial frontal cortex during interval timing. *The Journal of Neuroscience, 34*(50), 16774–16783.

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology, 19*(10), 827–830.

Perrett, S. P., Ruiz, B. P., & Mauk, M. D. (1993). Cerebellar cortex lesions disrupt learning-dependent timing of conditioned eyelid responses. *The Journal of Neuroscience, 13*(4), 1708–1718.

Petter, E. A., Lusk, N. A., Hesslow, G., & Meck, W. H. (2016). Interactive roles of the cerebellum and striatum in sub-second and supra-second timing: Support for an initiation, continuation, adjustment, and termination (ICAT) model of temporal processing. *Neuroscience and Biobehavioral Reviews, 71*, 739–755.

Prevosto, V., Graf, W., & Ugolini, G. (2017). The control of eye movements by the cerebellar nuclei: Polysynaptic projections from the fastigial, interpositus posterior and dentate nuclei to lateral rectus motoneurons in primates. *The European Journal of Neuroscience, 45*(12), 1538–1552.

Protopapa, F., Hayashi, M. J., Kulashekhar, S., van der Zwaag, W., Battistella, G., Murray, M. M., et al. (2019). Chronotopic maps in human supplementary motor area. *PLoS Biology, 17*(3), e3000026.

Rao, S. M., Harrington, D. L., Haaland, K. Y., Bobholz, J. A., Cox, R. W., & Binder, J. R. (1997). Distributed neural systems underlying the timing of movements. *The Journal of Neuroscience, 17*(14), 5528–5535.

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review, 12*(6), 969–992.

Roth, M. J., Synofzik, M., & Lindner, A. (2013). The cerebellum optimizes perceptual predictions about external sensory events. *Current Biology, 23*(10), 930–935.

Schneider, B. A., & Ghose, G. M. (2012). Temporal production signals in parietal cortex. *PLoS Biology, 10*(10), e1001413.

Schwerdt, H. N., Amemori, K., Gibson, D. J., Stanwicks, L. L., Yoshida, T., Bichot, N. P., et al. (2020). Dopamine and beta-band oscillations differentially link to striatal value and motor control. *Science Advances, 6*(39), eabb9226.

Shelhamer, M., & Joiner, W. M. (2003). Saccades exhibit abrupt transition between reactive and predictive; predictive saccade sequences have long-term correlations. *Journal of Neurophysiology, 90*(4), 2763–2769.

Shimbo, A., Izawa, E. I., & Fujisawa, S. (2021). Scalable representation of time in the hippocampus. *Science Advances, 7*(6), eabd7013.

Sieveritz, B., & Raghavan, R. T. (2021). The central thalamus: Gatekeeper or processing hub? *The Journal of Neuroscience, 41*(23), 4954–4956.

Smith, J. G., Harper, D. N., Gittings, D., & Abernethy, D. (2007). The effect of Parkinson's disease on time estimation as a function of stimulus duration range and modality. *Brain and Cognition, 64*(2), 130–143.

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

Spencer, R. M., Zelaznik, H. N., Diedrichsen, J., & Ivry, R. B. (2003). Disrupted timing of discontinuous but not continuous movements by cerebellar lesions. *Science, 300*(5624), 1437–1439.

Suzuki, T. W., & Tanaka, M. (2019). Neural oscillations in the primate caudate nucleus correlate with different preparatory states for temporal production. *Communications Biology, 2*, 102.

Suzuki, T. W., Kunimatsu, J., & Tanaka, M. (2016). Correlation between pupil size and subjective passage of time in non-human primates. *The Journal of Neuroscience, 36*(44), 11331–11337.

Suzuki, T. W., Inoue, K. I., Takada, M., & Tanaka, M. (2021). Effects of optogenetic suppression of cortical input on primate thalamic neuronal activity during goal-directed behavior. *eNeuro, 8*(2), ENEURO.0511-20.2021.

Takeya, R., Kameda, M., Patel, A. D., & Tanaka, M. (2017). Predictive and tempo-flexible synchronization to a visual metronome in monkeys. *Scientific Reports, 7*(1), 6127.

Takeya, R., Patel, A. D., & Tanaka, M. (2018). Temporal generalization of synchronized saccades beyond the trained range in monkeys. *Frontiers in Psychology, 9*, 2172.

Tanaka, M. (2006). Inactivation of the central thalamus delays self-timed saccades. *Nature Neuroscience, 9*(1), 20–22.

Tanaka, M. (2007). Cognitive signals in the primate motor thalamus predict saccade timing. *The Journal of Neuroscience, 27*(44), 12109–12118.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011a). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience, 31*(10), 3805–3812.

Teki, S., Grube, M., & Griffiths, T. D. (2011b). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Frontiers in Integrative Neuroscience, 5*, 90.

Tokushige, S. I., Terao, Y., Matsuda, S., Furubayashi, T., Sasaki, T., Inomata-Terada, S., et al. (2018). Does the clock tick slower or faster in Parkinson's disease? Insights gained from the synchronized tapping task. *Frontiers in Psychology, 9*, 1178.

Uematsu, A., & Tanaka, M. (2022). Effects of GABAergic and glutamatergic inputs on temporal prediction signals in the primate cerebellar nucleus. *Neuroscience, 482*, 161–171.

Uematsu, A., Ohmae, S., & Tanaka, M. (2017). Facilitation of temporal prediction by electrical stimulation to the primate cerebellar nuclei. *Neuroscience, 346*, 190–196.

Wang, J., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible timing by temporal scaling of cortical responses. *Nature Neuroscience, 21*(1), 102–110.

Wiener, M., Turkeltaub, P., & Coslett, H. B. (2010). The image of time: A voxel-wise meta-analysis. *NeuroImage, 49*(2), 1728–1740.

Witt, S. T., Laird, A. R., & Meyerand, M. E. (2008). Functional neuroimaging correlates of finger-tapping task variations: An ALE meta-analysis. *NeuroImage, 42*(1), 343–356.

Wittmann, M. (2013). The inner sense of time: How the brain creates a representation of duration. *Nature Reviews Neuroscience, 14*(3), 217–223.

Xu, M., Zhang, S. Y., Dan, Y., & Poo, M. M. (2014). Representation of interval timing by temporally scalable firing patterns in rat prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 111*(1), 480–485.

Yumoto, N., Lu, X., Henry, T. R., Miyachi, S., Nambu, A., Fukai, T., et al. (2011). A neural correlate of the processing of multi-second time intervals in primate prefrontal cortex. *PLoS One, 6*(4), e19168.

Zhou, S., & Buonomano, D. V. (2022). Neural population clocks: Encoding time in dynamic patterns of neural activity. *Behavioral Neuroscience, 136*, 374–382.

Diverse Time Encoding Strategies Within the Medial Premotor Areas of the Primate

Hugo Merchant, Germán Mendoza,
Oswaldo Pérez, Abraham Betancourt,
Pamela García-Saldivar, and Luis Prado

Abstract

The measurement of time in the subsecond scale is critical for many sophisticated behaviors, yet its neural underpinnings are largely unknown. Recent neurophysiological experiments from our laboratory have shown that the neural activity in the medial premotor areas (MPC) of macaques can represent different aspects of temporal processing. During single interval categorization, we found that preSMA encodes a subjective category limit by reaching a peak of activity at a time that divides the set of test intervals into short and long. We also observed neural signals associated with the category selected by the subjects and the reward outcomes of the perceptual decision. On the other hand, we have studied the behavioral and neurophysiological basis of rhythmic timing. First, we have shown in different tapping tasks that macaques are able to produce predictively and accurately intervals that are cued by auditory or visual metronomes or when intervals are produced internally without sensory guidance. In addition, we found that the rhythmic timing mechanism in MPC is governed by different layers of neural clocks. Next, the instantaneous activity of single cells shows ramping activity that encodes the elapsed or remaining time for a tapping movement. In addition, we found MPC neurons that build neural sequences, forming dynamic patterns of activation that flexibly cover all the produced interval depending on the tapping tempo. This rhythmic neural clock resets on every interval providing an internal representation of pulse. Furthermore, the MPC cells show mixed selectivity, encoding not only elapsed time, but also the tempo of the tapping and the serial order element in the rhythmic sequence. Hence, MPC can map different task parameters, including the passage of time, using different cell populations. Finally, the projection of the time varying activity of MPC hundreds of cells into a low dimensional state space showed circular neural trajectories whose geometry represented the internal pulse and the tapping tempo. Overall, these findings support the notion that MPC is part of the core timing mechanism for both single interval and rhythmic timing, using neural clocks with different encoding principles, probably to flexibly encode and mix the timing representation with other task parameters.

Keywords

Interval timing · Medial premotor areas · SMA · Categorical boundary · Ramping activity · Population dynamics

---

H. Merchant (✉) · G. Mendoza · O. Pérez ·
A. Betancourt · P. García-Saldivar · L. Prado
Instituto de Neurobiología, UNAM, Campus
Juriquilla, Querétaro, Mexico
e-mail: hugomerchant@unam.mx

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental
Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_7

Introduction

Time is a crucial parameter in life, and organisms have developed different mechanisms to quantify and predict events within the continuous flow of change in the environment (Tsao et al., 2022). Even if the central nervous system does not have a time sensory organ, animals are able to extract temporal information from stimuli of all sensory modalities and use it to generate timed behaviors. This paper focuses on the neural underpinnings of temporal processing during the perception and production of intervals in the hundreds of milliseconds range. This time scale is involved in basic but highly important behaviors observed since the invertebrates, such as collision avoidance and moving target interception (Merchant et al., 2001, 2003a, 2004a, b; Merchant & Georgopoulos, 2006). In addition, the range of hundreds of milliseconds is the scenario of complex behaviors such as the perception and production of speech (Assaneo et al., 2021), the execution and appreciation of music and dance (Merchant et al., 2015a; Mendoza & Merchant, 2014), and the performance of a large variety of sports (Merchant et al., 2003b; Merchant & Pérez, 2009). Researchers have found evidence of two distinct timing mechanisms in this scale: interval- and beat-based timing (Grube et al., 2010a, b; Teki et al., 2011; Breska & Ivry, 2018). In theory, interval-based timing implies the measurement of the absolute duration of discrete time intervals. In contrast, beat-based timing implies the quantification of relative durations with respect to the temporal regularity of the beat present in a stream of stimuli, such as a piece of music (Merchant et al., 2015a). Functional imaging and patient experiments have shown that the olivocerebellar and cortico-thalamic-basal ganglia circuit (CTBGc) are involved in the interval and beat-based timing, respectively (Cadena-Valencia et al., 2018; Grahn & Rowe, 2009; Sánchez-Moncada et al., 2020; Teki et al., 2012). Nevertheless, the medial premotor areas (MPC), composed of the supplementary motor area proper (SMA) and the presupplementary motor area (preSMA), are the common output to both the cerebellar and basal ganglia circuits (Rajendran et al., 2018; Schwartze et al., 2012), conferring them the ability to process temporal information during both single interval and rhythmic perception and production tasks.

In this paper, we argue that neural populations in the primate MPC show flexible and multiplexed time encoding strategies that support both interval-based and beat-based timing, largely drawing from observations from our laboratory (see Merchant & Honing, 2014 for initial view on this subject). In section "Neurophysiology of Interval-Based Perception in MPC", we describe the strategies by which preSMA encodes interval duration, category (long or short), and reward outcome based on a single-interval categorization task. In section "Neurophysiology of Beat-Based Perception in MPC", we show how neural populations in the MPC encode beat-based timing through a metronome synchronization-continuation task. We conclude with a brief summary and outlook for future work on timing neurophysiology in the primate brain.

Neurophysiology of Interval-Based Perception in MPC

Different psychophysical tasks have been developed to study how the brain processes temporal information from single intervals. We can classify these tasks according to their sensory and motor requirements or by the implicated sensory modality (Merchant et al., 2008a, b, c). In the motor domain, for example, one of the most employed paradigms is interval reproduction, in which subjects are first shown one interval or duration and then are asked to reproduce the duration with some motor response (Bartolo & Merchant, 2009; Jazayeri & Shadlen, 2015; Woodrow, 1930). Interval discrimination is one of the most employed paradigms in the sensory domain (Wearden, 1992; Kononowicz & van Rijn, 2014; Di Fabio et al., 2011). In this task, subjects are presented with two single, consecutive intervals and then are asked to emit a relative judgment about the two durations (Kim et al., 2013). The categorization of intervals as short or long is another instance of perceptual, interval-based timing (Ng et al., 2011; Méndez et al., 2014). This paradigm, also known as time bisection, is widely used since the late 1970s of the past century to explore the mechanism by which animals perceive the passage of time (Church & Deluty, 1977; Gibbon, 1981; Wearden, 1991; Allan & Gibbon, 1991). In these tasks, subjects are initially trained to make one or another action in response to prototypic short or long intervals. When subjects learn to differentiate the short from the long intervals, intermediate test durations are presented. The subject's goal is to categorize each test interval as short or long using the appropriate behavioral response. Despite being the subject of a long history of psychophysical studies, the neural basis of interval categorization remained unknown until recently.

The Ability of Monkeys and Humans to Categorize Intervals as Short or Long Is Similar

To provide information on the psychophysics and neural mechanisms of interval categorization in primates, we developed a paradigm where human subjects and Rhesus monkeys (*Macaca mulatta*) classify single intervals in the range of hundreds of milliseconds as short or long according to an arbitrary category limit (Mendez et al., 2011; Mendoza et al., 2018). In each task trial, the subjects were shown a test interval, with the onset and offset indicated by a brief visual stimulus displayed on a screen. After the interval offset and a fixed delay, the subjects communicated their perceptual decision by moving a cursor displayed on the screen into an orange circle, if the interval was categorized as short, or into a blue circle if the interval was classified as long. Crucially, subjects categorized three sets of eight intervals in every experiment. The short–long limit was the mean of the intervals of the set. Consequently, the shorter four intervals in a set should be categorized as short and the remaining as long. Importantly, the interval corresponding to the actual category limit of each set was never presented to the subjects as a test interval. As a result of this experimental design, the subjects had to change their subjective category boundary to classify the intervals of the different sets correctly.

We found that humans and monkeys perform similarly in these categorization tasks. Both species showed sigmoid psychometric functions, with the probability of long responses increasing as a function of the length of the intervals. In addition, both species got more correct answers for each block's shortest and longest intervals and more decision errors for the intermediate intervals. Furthermore, for both humans and monkeys, the bisection point, the interval at which the probability of "long" response is 0.5, was close to the mean of the test durations of each set. Also, both species showed similar relative thresholds and an increase in temporal variability as a function of the implicit interval, following the Scalar Property for Timing (Gibbon, 1981; Allan & Gibbon, 1991). All these results were concordant with previous behavioral data from monkeys and humans categorizing time intervals and showed that both species have similar abilities for the categorical perception of time (Mendez et al., 2011; Kopec & Brody, 2010; Merritt et al., 2010). When human subjects were tested in our interval categorization task, we found that even though there was a 1-second-long delay between interval presentation and decision communication, categorization difficulty affected subjects' performance, as well as their reaction and movement time. In addition, reaction and movement times were also influenced by the distance between the targets. This implies that not only perceptual, but also movement-related considerations were incorporated into the decision process (Méndez et al., 2014). In addition, we conducted a beta burst TMS study in humans and found that decreasing the excitability of MPC produced a clear disruption of timing during the same categorization task (Méndez et al., 2017). Then we decided to study the neural properties of the primate preSMA cells during this paradigm.

The Subjective Limit Between the Short and Long Categories Is Encoded by Neurons of the Monkey's MPC

We recorded the activity of neurons in the preSMA of Rhesus monkeys working in the relative interval categorization task. The monkeys had to change their subjective boundary within three blocks to correctly classify the intervals of the different sets. Interestingly, we found that this internal criterion is encoded by neurons in preSMA (Mendoza et al., 2018). We observed that some preSMA neurons, henceforth called "boundary neurons," generated a peak of activity at a relatively constant time after the interval onset, regardless of the actual duration of the presented interval. Thus, the peak of activity tended to occur after the interval offset when the test interval was short but before the interval offset when the test interval was long (Fig. 1a). This timed activity could serve as a mental reference to delimit the duration of the short and long intervals. Such a mechanism would only require information of the relative time of occurrence of the peak activity and the interval offset. This hypothesis was supported by the fact that the time of occurrence of the peak activity changed according to the interval set being tested. Hence, the time at which neurons reached the peak activity was close to the actual category boundary of the current set. We observed this effect at the population and single neuron levels. For the population of boundary neurons, the mean of peak activity times was close to the actual boundary of each interval set. At the single neuron level, some neurons changed their time of peak activity across the different sets, occurring earlier for the shorter set and later for the longer set but always resembling the actual category limit (Fig. 1a, b).

If boundary neurons encoded the subjective boundary, a correlation between the category selected by the monkeys and the category decoded from the activity of the boundary neurons should exist. Therefore, we determined whether the relative temporal occurrence of the peak activity and the end of the interval signaled by the presentation of the second stimulus could be used to decode the categorical decision of the monkeys. For each boundary neuron and trial, we quantified the time elapsed (τ) between the peak of activity and the interval offset (Fig. 1c). Next, we looked for the τ that minimized the error in classifying each interval as short or long (the best decoding criterion; Fig. 1d). Finally, we constructed neurometric curves with the probability of a particular interval being categorized as long obtained from the proportion of trials in which τ was smaller than the best decoding criterion (Fig. 1e). We found that the best criterions allowed the decoding of the monkey's responses with high precision. Consequently, the relative thresholds and the points of subjective equality from the neurometric and the corresponding psychometric curves were similar. Significant trial-by-trial correlations between the category selected by the monkeys and the category decoded from the peak activity were found for a large group of neurons. This is remarkable since limit was never shown as a test interval to the monkeys. Therefore, the neural representation of the boundary was likely computed subjectively from the actually presented intervals.

PreSMA Neurons Explicitly Represent the Categorical Response of the Monkeys

Since activity of boundary neurons can be used to decode the category selected by the monkeys. These results strongly support the notion that this signal encodes a mental criterion to judge the intervals as short or long. Interestingly, we also found a subpopulation of preSMA neurons that explicitly encode the category selected by the monkeys (Fig. 2a). These category-encoding neurons showed selective activity to the short or long responses: their activity was similar for all the intervals assigned by the monkeys to the same category and different for the intervals classified in the opposite category. We found neurons with higher activity for intervals categorized as short and neurons with higher activity for intervals categorized as long. These responses are similar to previously reported neurons encoding the cate-

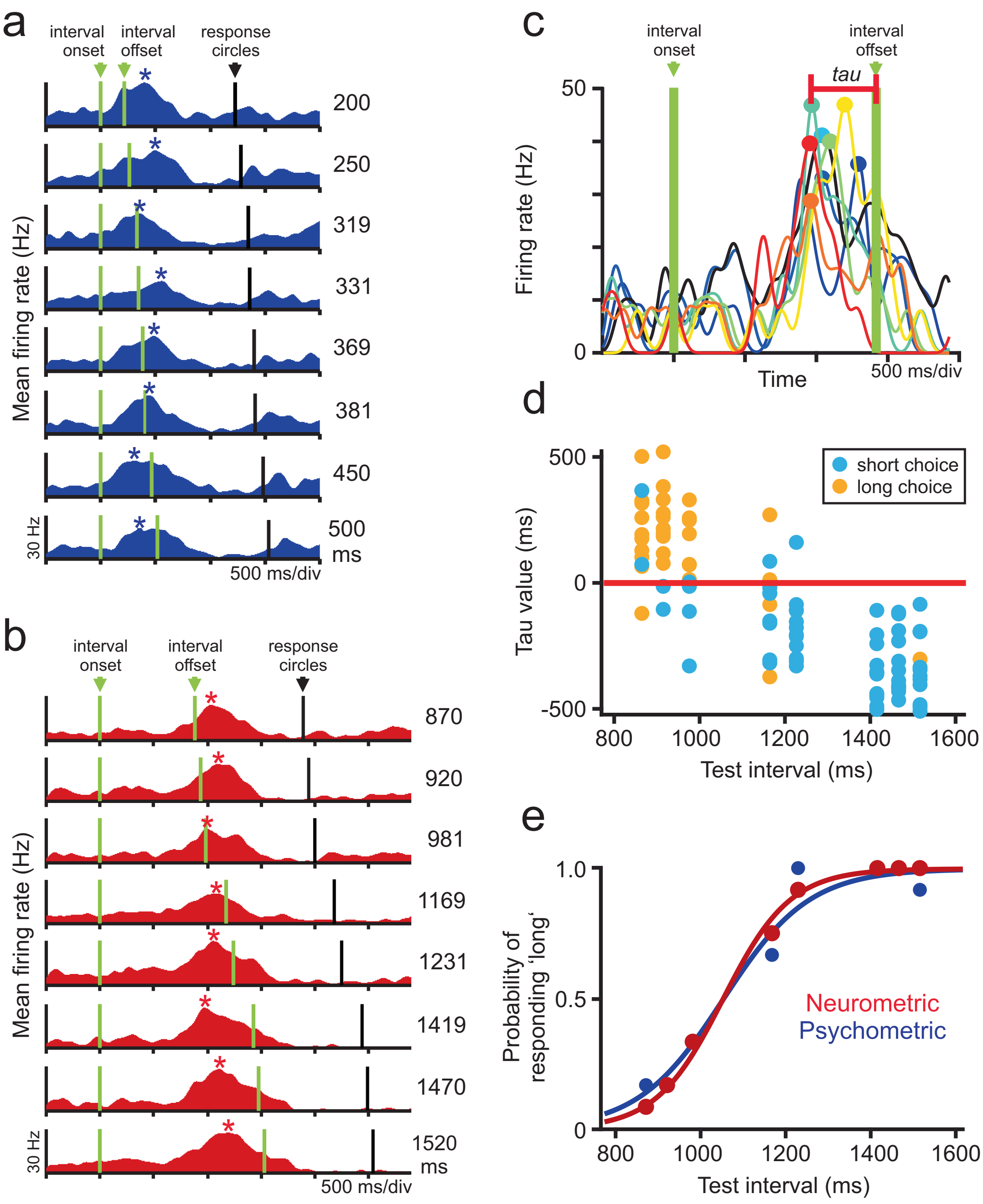


Fig. 1 Neurons in the preSMA of the Rhesus monkey encode the subjective category boundary in the interval categorization task. (**a**) Some preSMA neurons generate peaks of activity (indicated by the asterisks) with a relatively fixed latency. Consequently, the peaks tend to occur after the offset of the short but before the offset of the long intervals. The activity is aligned to the interval onset and is segregated by the intervals being categorized (indicated to the right of the panels). (**b**) Activity of the neuron shown in a) but during the categorization of a different set of longer intervals. Note that the peak activity occurs now around 1100 ms after interval onset. (**c**) For each trial, the time difference (tau) between the peak of neuronal activity and the end of the test interval was measured. (**d**) These measures tended to be positive for trials in which short intervals were presented and negative for long-interval trials and segregated short from long intervals in a way that correlated with the decision of the monkeys (indicated by the dot color; see the inset). The red line indicates the best criterion for this boundary neuron and set of intervals (see main text). (**e**) Psychometric and neurometric functions for the data shown in (**d**). (Figures were adapted from Mendoza et al., 2018)

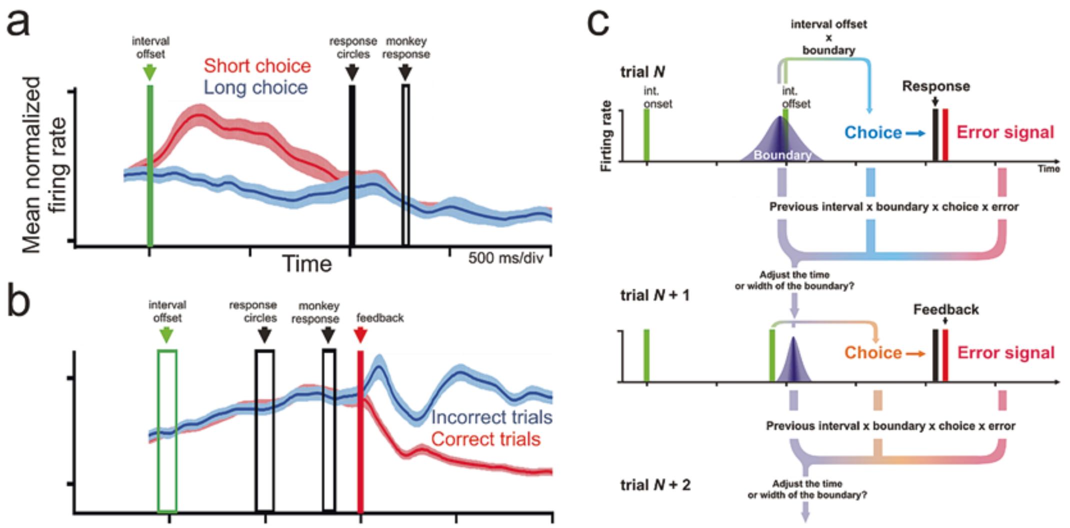


Fig. 2 (**a**) Mean (+ SEM) activity of a population of preSMA neurons encoding the category selected by the monkeys. In the short-response trials, these neurons showed higher activity before the actual monkey's response. Note that the activity is aligned to the interval offset; the interval onset is not shown. (**b**) Mean activity of a population of neurons encoding the outcome of the monkey's decision. These neurons decreased their activity after the reward delivery (feedback time) in the correct trials. Activity is aligned to the time of feedback. (**c**) Hypothetical trial-by-trial adjustment of the neural category boundary. In every trial of the task, intervals are judged to be short or long based on the subjective boundary (upper panel). In subsequent trials, the time or width of the peak activity is adjusted depending on the task's variables in the previous trials (lower panels). (Figures a and b were adapted from Mendoza et al., 2018)

gorical decision during tactile (Romo et al., 1995, 1996; Merchant et al., 1997) and visual categorization tasks (Freedman et al., 2001; Merchant et al., 2011a). An additional difference with the activity of "boundary neurons" was that category-selective activity occurred between the interval offset and the monkeys' responses (Mendoza et al., 2018). In our task, this period comprised a fixed time delay between the interval offset and the presentation of the response targets. During this delay, the monkeys had to remember their perceptual decision until the presentation of the response circles. Hence, the time-persistent activity of category-selective cells maybe the neural substrate of the memory trace of the categorical decision.

The Reward Outcome of the Perceptual Decisions Is Encoded by the preSMA

The consequences of the perceptual decision were also represented in the activity of preSMA neurons. In this case, some neurons showed activity selective to whether the monkeys' responses were correct (Fig. 1g). This outcome-selective activity was observed after the monkeys' responses, i.e., after the delivery, or absence, of reward. Hence, some outcome-encoding neurons were more active for correct responses, and other neurons showed higher activity for incorrect responses (Mendoza et al., 2018). This error signal may be essential in adjusting the subjective boundary or the choice during the learning or execution of the task. Further modeling studies are needed to reveal how the encoding of categorization variables emerges and is optimized in the preSMA neural circuits.

Overall, our results show a sequence of information encoding in preSMA during the categorization of intervals. First, the learned mental category boundary is evoked and used as a time reference to judge the ongoing interval. Next, an explicit representation of the selected category emerges and is maintained as a memory trace during the delay previous to the actual response of the monkey. Finally, a strong representation of the outcome of the perceptual decision is generated and maintained during the intertrial time (Fig. 2c). Importantly, although the category boundary is encoded from the level of single neurons, the actual monkey's criterion must result from the combined neural activity of all the population of boundary neurons. In fact, we found that the mean times of peak activity of all the boundary neurons correlated with the mean behavioral bisection point of the monkeys for each set of the categorization task (Mendoza et al., 2018). The concept of a population code may also apply to the neural representations of the selected category and the outcome of the perceptual decisions.

Behavior and Neurophysiology Suggest Interval Categorization Is Solved by Similar Neural Mechanisms in Monkeys and Humans

A key question is whether our experimental observations in the Rhesus monkey can be extrapolated to other animals, including humans. The concept of an internal criterion was proposed in 1981 to explain the behavior of different species in interval categorization (Gibbon, 1981). Subsequent behavioral and theoretical studies proposed that such a hypothetical criterion corresponds to the bisection point in humans and other animals and that is determined by the range of the test intervals (Killeen et al., 1997; Allan, 2002a). Nevertheless, the criterion hypothesis remained one of several possible alternative explanations (Gibbon, 1981; Maddox & Ashby, 1993; Allan, 2002b). Neurophysiologic studies in humans provided additional support to the idea of an internal criterion in humans. Lindbergh and Kieffaber (Lindbergh & Kieffaber, 2013) observed significant differences in the ERPs time-locked to probe offset between intervals judged to be short and long. In another study, Ng and colleagues (2011) reported an ERP, the Contingent Negative Variation, to increase in amplitude up to the value of the short prototype to remain at a constant level until about the mean of the short and long anchors and then, to return to baseline. These observations are consistent with a decision mechanism based on a category limit; i.e., during a trial of the categorization task, once the subjective time limit between the short and long intervals is exceeded, there is no need to continue attending the interval duration. Therefore, the category membership of the long intervals can be determined as early as the interval reaches the internal time criterion. This is consistent with the finding of shorter reaction times for the longer test durations in categorization tasks compared to other perceptual timing tasks, and with the observation that interval categorization is less demanding than other timing tasks that are supposed to require the tracking of the complete intervals (Bannier et al., 2019). Consequently, these behavioral observations, models, and neurophysiologic studies suggest that the neural mechanism operating in humans during interval categorization is similar to the mechanism we described in the monkey working in equivalent tasks.

Factors Influencing the Subjective Boundary

Experiments with human and nonhuman primates and other species performing interval categorization found choice biases resulting from manipulations of several task variables, including the range of the test intervals, the long–short ratio, the interval spacing and the probability of stimulus or reward occurrence (Allan, 2002a; Elsmore, 1972; Stubbs, 1976; Akdoğan & Balcı, 2016; Cambraia et al., 2019). Commonly, these manipulations produce systematic shifts of the psychometric functions with consequent changes in the bisection point. Notably, some of these variables were reported to produce similar changes in other perceptual tasks. In detection tasks, the psychometric thresholds are affected by manipulating the payoff or stimulus contingencies. This phenomenon, called the response bias, was proposed to depend on changes in the observer's criterion (Gescheider, 1997). A related phenomenon is the effect of the trial-by-trial history on the current subject's performance. In interval categorization, the current judgments were reported to be sensitive to the choices in the previous trials (Wehrman et al., 2020). Thus, in order to learn the task or after changes in the task's variables, the subjective limit must be fine-tuned by the events in the previous trials.

We suspect that the recent task history can bias the subject's decisions and consequently can affect the peak activity of the boundary neurons. We propose that the time or width of the peak activity is modulated according to the variables in the previous trials (Fig. 2c). With training, the peak activity of the boundary neurons would adapt to the statistics of the tasks, for example, to the range or probability of the test intervals or the reward contingencies. Therefore, adjusting the internal boundary would be an iterative process with critical importance during learning or after changes in task contingencies. An alternative explanation to the response bias or the "history effect" is that they result from changes in the balance of the short–long choice-related neural activity, not from changes in the activity of boundary neurons. As we demonstrated, different neural codes in preSMA represent the subjective boundary and the selected category. Due to this functional segregation, our paradigm is well suited to determine whether specific task variables affect the internal criterion or the choice—ongoing studies in our laboratory attempt to provide data on these topics.

No Ramping Activity in preSMA During Interval Categorization?

To end, we compare our observations during interval categorization with the neurophysiologic observations made in other timing tasks. Critically, we found that neurons in the monkey's preSMA encode the main variables needed to solve the categorization task, but we did not find cells encoding elapsed time during the interval presentation (Mendoza et al., 2018). This is remarkable since different research groups have reported these types of activity in the preSMA of monkeys. Tanji and colleagues reported that neurons in the SMA and preSMA of the Japanese monkey increased their activity as a function of the elapsed time (ramping activity) during the production of single intervals (Mita et al., 2009). Experiments in our laboratory with Rhesus monkeys trained in multiple-interval, synchronization-continuation tapping tasks also demonstrated ramping activity in preSMA during the interval production (Merchant et al., 2011b). Interestingly, ramping activity related to elapsed time is also observed during the perceptual phase of interval reproduction tasks in other motor-related cortical areas. Each trial of these tasks has perceptual and reproduction phases. In the perceptual phase, a target interval is presented, and the subjects must attend to its duration but avoid movement. Then, in the subsequent reproduction phase of the trial, the subjects reproduce, with any motor action, the previously perceived interval (Jazayeri & Shadlen, 2015; Henke et al., 2021).

We hypothesize that ramping activity is mainly recruited by motor-related areas during tasks requiring the quantification of the whole intervals, such as the perceptual or motor phases of motor timing tasks or perceptual discrimination tasks (Merchant et al., 1997, 2011b). Not only does interval categorization lack an interval reproduction phase, but our decoding analysis also demonstrated that it could be solved without quantifying the total interval duration. These ideas are consistent with most neurophysiologic studies in monkeys (Jazayeri & Shadlen, 2015; Mendoza et al., 2018; Perrett, 1998; Maimon & Assad, 2006; Renoult et al., 2006; Tanaka, 2007; Leon & Shadlen, 2003; Lebedev et al., 2008; Genovesio et al., 2009; Oshio et al., 2008). Nevertheless, we cannot discard the possibility that elapsed time can be decoded from the activity of neurons in other brain areas (see for example Gouvêa et al., 2015). Additional studies in different animal species, with a more diverse battery of timing paradigms and recording the simultaneous activity of several cortical areas, are needed to clarify these issues.

In conclusion, our data suggest that the monkey preSMA solves interval categorization using a simple mechanism that does not require exhaustive quantifying each interval but fine-tuning a time boundary. Initial evidence from different research groups suggests that the same mechanism might operate in the human brain. Therefore, our observations support classical psychophysical hypotheses suggesting that subjects compare the intervals to be categorized with an internal criterion that depends on the distributions of the test intervals, the choice biases, and the reward contingencies.

Neurophysiology of Beat-Based Perception in MPC

Psychophysics of Rhythmic Tapping

Music and dance depend on intricate loops of perception and action, where temporal processing can be engaged during the synchronization of movements with sensory information or during the internal generation of movement sequences (Janata & Grafton, 2003). Thus, beat induction is the cognitive ability that allows humans to hear a regular pulse in music and to move in synchrony with it. Importantly, without beat induction there is no music perception and, hence, is considered a universal human trait (Honing, 2012; Honing et al., 2012, 2018). These cognitive abilities depend on an internal brain representation of pulse that involves the generation of regular temporal expectations, which is the core of the beat-based timing mechanism (Balasubramaniam et al., 2021). Several studies have shown that the internal pulse is directly mapped to the timing of the entrained movements, typically measured as tapping movements that occur few milliseconds before the beat (Lenc et al., 2021; Nozaradan et al., 2016), 2017. A classical task used to study beat induction is the synchronization-continuation tapping task (SCT), which can be considered a simplified version of beat perception and entrainment in music. In the SCT, subjects tap in sync with periodic sensory cues (synchronization epoch), and then keep tapping after the metronome is extinguished using an internal beat representation (continuation epoch) (Wing, 2002). Performance in this task shows two features present in other timing tasks: a linear increase in variability of produced intervals with the mean that is a form of Weber's law, called scalar property (Merchant et al., 2008a; Gibbon et al., 1997; García-Garibay et al., 2016), and an over- and underestimation of intervals for shorter and longer intervals, termed regression toward the mean or bias effect (Woodrow, 1934; McAuley & Jones, 2003; Pérez & Merchant, 2018; Yc et al., 2019). Furthermore, subjects use an error correction mechanism that maintains tap synchronization with the metronome, since a longer produced interval tends to be followed by a shorter interval and vice versa, to avoid error accumulation and losing the metronome (Repp, 2005; Iversen et al., 2015; Pérez et al., 2023). In contrast, during continuation, there is a drift in the duration of produced intervals (Madison, 2001; Collier & Ogden, 2004).

Humans possess a remarkable flexibility for beat-based timing, recognizing the beat from a wide range of complex rhythms, and with the natural tendency to predictively entrain to the beat by moving different body parts, such as finger or foot taps (Patel, 2018). Recently, it has been demonstrated that macaques possess the neural machinery to perceive and entrain to the simplest form of beat: an isochronous auditory metronome. On one side, beat perception has been measured with mismatch negativity (MMN), an auditory event-related EEG potential that can be used as an index of a violation of temporal expectation. Notably, MMN is sensitive to violations of the beat using complex or simple rhythms in humans but only for isochronous metronomes in monkeys (Honing et al., 2012, 2018). On the other, psychophysical experiments showed that, when immediate feedback about the timing of each movement is provided, monkeys can predictively entrain to an isochronous beat, generating tapping movements in anticipation of the metronome (Gámez et al., 2018; see also García-Garibay et al., 2016; Takeya et al., 2017). In fact, macaques can flexibly change their movement tempo from trial to trial covering a range from 400 to 1000 ms (unpublished observations). Furthermore, monkeys can superimpose accentuation patterns onto an isochronous auditory sequence, suggesting that they can generate a simple subjective rhythm on a regular auditory sequence (Ayala et al., 2017; Criscuolo et al., 2023). Thus, both primate species can produce negative asynchronies, show an error correction mechanism, produce precise and accurate interval during synchronization to an isochronous (Betancourt et al., 2023). Nevertheless, humans perform better with auditory but monkeys with visual metronomes (Gámez et al., 2018; Betancourt et al., 2023; Zarco et al., 2009). These findings support the gradual audio-motor hypothesis that suggests that beat-based timing emerged gradually in primates, peaking in humans due to a sophisticated audio-motor circuit, and that is also present for isochrony in macaques where it depends on the close interaction between MPC, CTBGc, and the auditory cortex (Merchant & Honing, 2014; Honing & Merchant, 2014). Indeed, in both primate species, it has been shown that MPC plays a critical role in beat extraction and entrainment (Merchant et al., 2015a; Mendoza & Merchant, 2014; Rao et al., 1997; Chen et al., 2008).

Ramping Activity as a Single-Cell Timing Signal for Rhythmic Tapping

We recorded the activity of MPC cells during a version of the SCT where monkeys produced three intervals in the synchronization and three intervals in the continuation epochs (Fig. 3a). Brief auditory or visual interval markers were used during the synchronization phase and the range of target intervals was from 450 to 1000 ms (Zarco et al., 2009). The monkeys were able to produce the target intervals accurately, showing an average underestimation of ~50 ms across interval durations during the synchronization and continuation phases of the SCT. In addition, we analyzed the temporal variability of the monkeys' tapping performance, defined as the SD of the individual interresponse intervals (Merchant et al., 2008a, b). The temporal variability followed the scalar property, with a linear increase as a function of interval duration in both phases of SCT (Zarco et al., 2009). Furthermore, the analysis of the tapping hand kinematics revealed that monkeys temporalize the pauses or dwell between movements, while producing stereotypic downward and upward movements with a similar duration across the tested metronome tempos (450–1000 ms). These findings suggest that monkeys use an explicit timing strategy to perform the SCT, where the timing mechanism controls the dwell duration, while also triggering the execution of stereotyped tapping movements across each produced interval in the rhythmic sequence (Gámez et al., 2018; Donnet et al., 2014).

The extracellular activity of single MPC neurons was recorded during task performance. A large population of neurons showed ramping activity before or after each sensory or motor event in the SCT (Merchant et al., 2011b). Consequently, we developed a warping algorithm to determine whether the responses of the cells were aligned to the sensory or motor aspects of the SCT, and we found that most MPC cells were aligned to the tapping movements instead of the stimuli used to drive the temporal behavior (Perez et al., 2013; Merchant et al., 2015b).

Next, we used an iterative algorithm to find the best regression model to explain the increase or decrease of instantaneous activity over time with respect to the tapping times using the spike density function. With this method, we defined for each ramp the following parameters: duration, slope, peak magnitude, and the time τ from the peak to the button press. Using this information, we classified different cell populations with ramping activity into four groups: motor, relative-timing, absolute-timing and time-accumulator, and swinging cells (Merchant et al., 2011b). A large group of cells shows ramps before the movement onset that are similar across produced durations and the sequential structure of the task and, therefore, are considered motor ramps (Fig. 3b). Interestingly, another cell population showed an increase in ramp duration but a decrease in slope as a function of the animals' produced duration, reaching a similar discharge magnitude at a specific time before the button press. These cells are called relative-timing cells since their ramping profile could

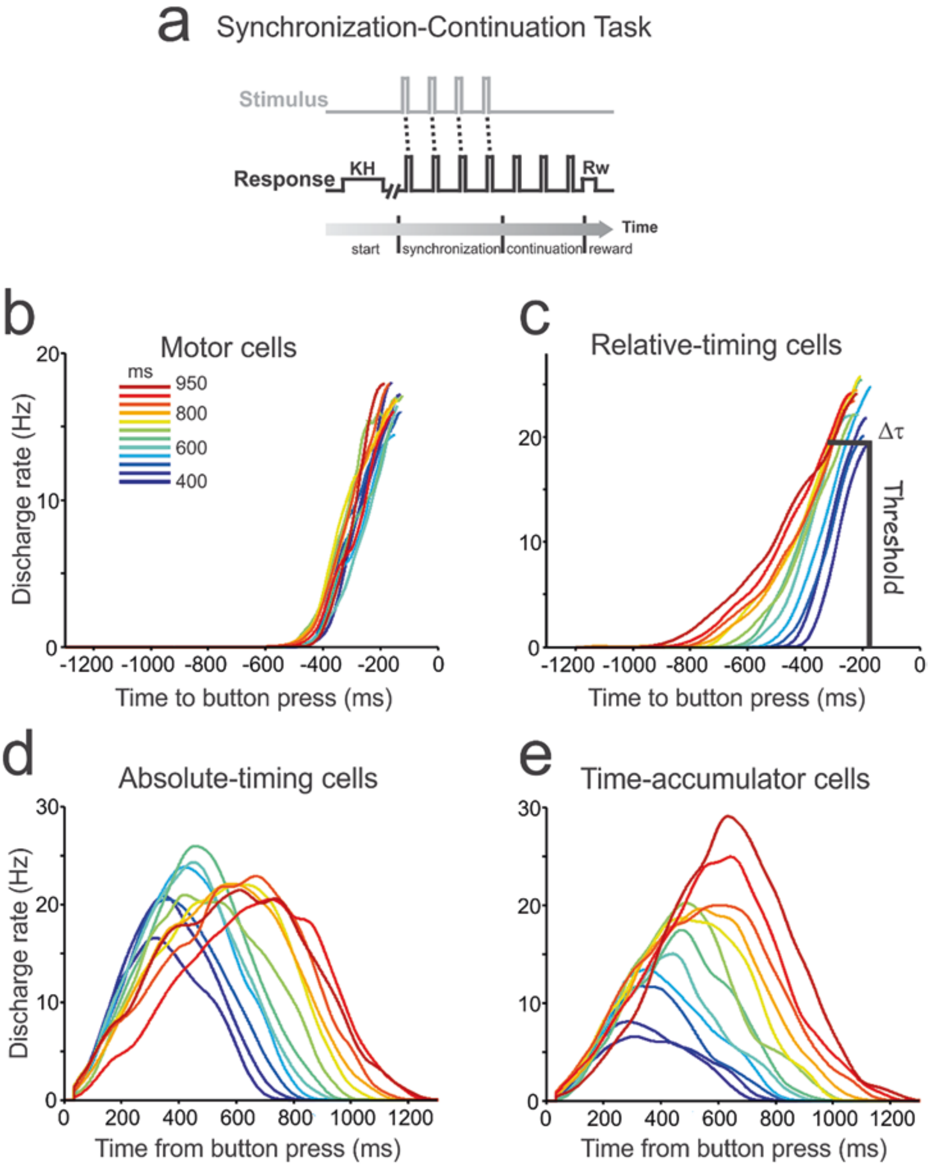


Fig. 3 (**a**) Synchronization-Continuation Task (SCT). Monkeys were required to push a button (R, black line) each time stimuli with a constant interstimulus interval (S, gray line) were presented, which resulted in a stimulus-movement cycle. After four consecutive synchronized movements, the stimuli stopped, and the monkeys continued tapping with similar pacing for three additional intervals. The target intervals, defined by brief auditory or visual stimuli, were 450, 550, 650, 850, and 1000 ms, and were chosen pseudo-randomly within a repetition. Different ramp populations: motor (**b**), relative-timing (**c**), absolute-timing (**d**), and time-accumulator (**e**) cells. b and c are aligned to the next button press, while c and d are aligned to the previous button press. The duration of produced intervals is color-coded in b. All these ramp population functions correspond to the addition of individual ramps over time. (Modified from Merchant et al., 2011b)

signal how much time is left to trigger the button press in the task sequence (Fig. 3c). Therefore, this population of MPC neurons has the response properties to encode the time remaining for an event or time-to-contact cells, and once the population reaches a firing magnitude threshold, it could trigger the internal beat signal (Merchant et al., 2011b).

Other groups of cells show a consistent increase followed by a decrease in their instantaneous discharge rate when their activity was aligned to the previous button press rather than to the next one. These cells showed an up-down activation profile whose duration increased as a function of the produced interval (Fig. 3d) and were called absolute-timing cells. In addition, we found cells that responded as sand clocks since their activity was accumulated as a function of the passage of time, with final peaks of activity that increased linearly with the produced interval and hence were denominated time-accumulator cells (Fig. 3e). Therefore, these cells could be representing the elapsed time since the previous movement, using two different encoding strategies: one functioning as an accumulator of elapsed time where the peak magnitude and the duration of the activation period are directly associated with the time passed, and another where only the duration of the activation period is encoding the length of the time passed since the previous movement (Merchant et al., 2011b; Merchant & Averbeck, 2017).

Cell activity changes associated with temporal information processing in behaving monkeys have been reported in the cerebellum (Perrett, 1998; Okada et al., 2022; Ohmae et al., 2017), the putamen (Bartolo et al., 2014; Bartolo & Merchant, 2015; Merchant & Bartolo, 2018), the caudate (Kameda et al., 2019; Kunimatsu et al., 2018), the thalamus (Tanaka, 2007; Wang et al., 2018), the posterior parietal cortex (Jazayeri & Shadlen, 2015; Renoult et al., 2006; Leon & Shadlen, 2003), and the prefrontal cortex (Henke et al., 2021; Oshio et al., 2008; Brody, 2003), as well as in the motor cortex (Merchant et al., 2004a; Maimon & Assad, 2006; Lebedev et al., 2008), and MPC (Mita et al., 2009; García-Garibay et al., 2016; Merchant et al., 2014; Sohn et al., 2019). These areas form different circuits linked to sensorimotor processing using the skeletomotor or oculomotor effector systems. Most of these studies have described climbing activity during different timing contexts, which include discrimination of time, time estimation, time categorization, single interval reproduction, and rhythmically produced saccades and hand movements. Therefore, the increase or decrease in instantaneous activity as a function of the passage of time is a property present in many cortical and subcortical areas of the CTBGc and the cerebellum that may be involved in different aspects of temporal processing in the hundreds of milliseconds scale. In fact, the ubiquitous presence of cells' increases or decreases in discharge rate as a function of time across different timing tasks and areas of a potential core timing circuit suggests that ramping activity is a fundamental element of the timing mechanism.

The recently reported temporal scaling as a mechanism to encode time remaining for action is very evident in ramping cells that behave as time-to-contact cells or the absolute timing cells reported by us (Fig. 3c, d) (Merchant & Averbeck, 2017). These cells show a similar instantaneous pattern of activations that is contracted for short and expanded for long intervals and can be the single-cell primordium for the neural population temporal scaling observed in neural trajectories in state space (Wang et al., 2018; Remington et al., 2018). In contrast, the time-accumulator cells do not scale in time; their main feature is the magnitude of their activity peak (Gámez et al., 2019). This type of cells might be very common during tasks that demand elapsed time encoding instead of predicting a sensory, cognitive, or motor event (Bi & Zhou, 2020; Merchant & Pérez, 2020).

Interval Tuning: A Circuit Signal for Context-Dependent Flexibility

Psychophysical studies on learning and generalization of time intervals support the notion that neurons in the timing circuit are tuned to specific interval durations but can be activated in a modality- and context-independent fashion (Sánchez-Moncada et al., 2020; Bartolo & Merchant, 2009; Meegan et al., 2000; Nagarajan et al., 1998). Accordingly, we found a graded modulation in the discharge rate of cells as a function of interval duration during the SCT in cells of MPC (Merchant et al., 2013a). Figure 4a, b shows the profile of activation of a cell in the preSMA of a

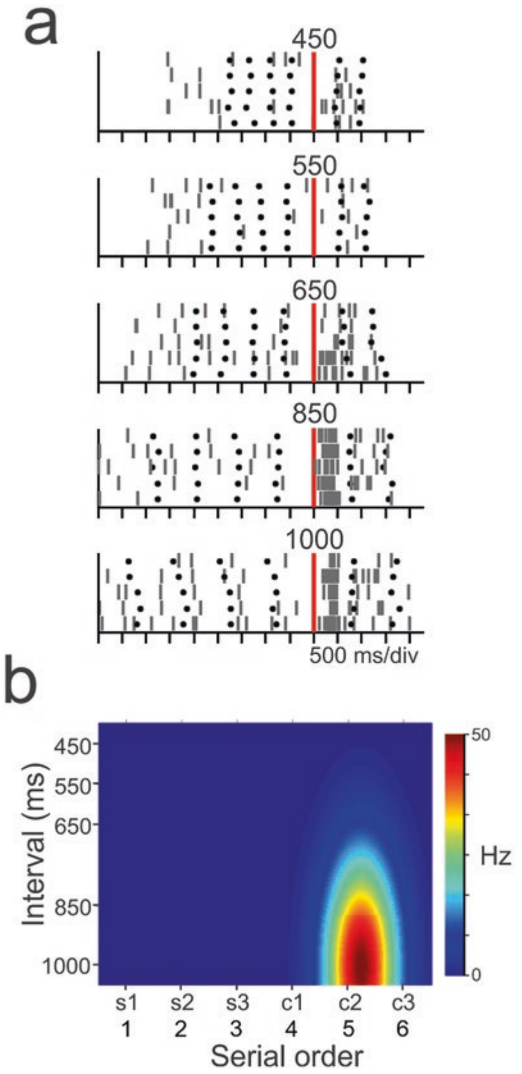


Fig. 4 (**a**) Responses of a sharply double-tuned MPC cell with a long-preferred interval and a preferred sequence order around the second continuation interval. Circles correspond to tap times. The raster is aligned (red line) to the second tap of the continuation phase. All target intervals are showed in a vertical arrangement. (**b**) Double-Gaussian tuning function for the cell responses depicted in a. The heatmap discharge scaler is on the right. (Modified from Merchant et al., 2013a)

monkey performing this task. The neuron shows larger activity for the longest durations, with a preferred interval of around 900 ms. A large population of MPC cells is tuned to different interval durations during the SCT, with a distribution of preferred intervals that covers all the tested durations, although there was a bias toward long preferred intervals (Merchant et al., 2013b). These observations suggest that the MPC represents interval duration, where different populations of interval-tuned cells are activated depending on the duration of the produced interval (Merchant & Bartolo, 2018). In addition, most of these cells also showed selectivity to the sequential organization of the task, as previously described in sequential motor tasks in MPC (Tanji, 2001). The cell in Fig. 4a, b also shows an increase in activity during the fifth produced interval in the SCT sequence. Again, at the cell population level, all the possible preferred ordinal sequences were covered (Merchant et al., 2013b). Hence, the temporal and sequential information is multiplexed in a cell population signal that defines the duration of the produced interval and its position in the learned SCT sequence (Bartolo et al., 2014; Merchant et al., 2013a). Overall, these findings support the notion that MPC uses mixed selectivity to represent the passage of time, the tempo duration, and the serial order during SCT (Merchant & Bartolo, 2018; Gámez et al., 2019; Merchant et al., 2013b).

Interval tuning during single interval- and beat-based timing has been reported in MPC (Mita et al., 2009; Merchant et al., 2013b), the prefrontal cortex (Henke et al., 2021), the putamen (Bartolo et al., 2014; Bartolo & Merchant, 2015), the caudate (Kameda et al., 2019; Kunimatsu et al., 2018), and the cerebellum (Okada et al., 2022; Ohmae et al., 2017). In addition, a chronomap in the medial premotor cortex has been described in humans using functional imaging. The interval-specific circuits show a topography with short preferred intervals in the anterior and long preferred intervals in the posterior portion of SMA/preSMA (Protopapa et al., 2019). Hence, timing not only depends on one population of cells that contracts or expands their activity patterns depending on a constant speed knob (Wang et al., 2018) but also on interval-specific neurons that build distinct timing circuits. It is well known that tuning and modularity are mechanisms for the division of labor that are used in cortical and subcortical circuits to represent sensory, cognitive and motor information

(Hubel & Wiesel, 1977; Mountcastle, 1998; Georgopoulos et al., 2007; Goldman-Rakic et al., 1984; Naselaris et al., 2006a, b). Consequently, interval tuning can provide large flexibility to mix time encoding and prediction with other task parameters, which are also represented in the premotor system but with different mapping frameworks (Merchant & Bartolo, 2018; Merchant & Yarrow, 2016; Zhou et al., 2022).

Neural Sequences

As a population, ramping and/or tuned MPC cells show activation profiles that are far from static. Indeed, these cells are recruited in rapid succession producing a progressive neural pattern of activation (called neural sequences or moving bumps) that flexibly fills the beat duration depending on the tapping tempo. Hence, neural sequences provide a relative representation of how far an interval has evolved between the taps (Merchant et al., 2015b; Crowe et al., 2014; Zhou et al., 2020). Notably, this periodic clock resets on every period cycle, encoding the interval pulse in this resetting. In fact, the neural chain progression starts with a group of cells, migrates to other cells during the timed interval, stops with the last group of cells, and simultaneously is initialized for the next produced interval with the previous initial set of cells (Lenc et al., 2021; Gámez et al., 2018; Merchant et al., 2015b). Three parameters of the moving bumps are directly linked with the representation of the tapping tempo duration: (1) The duration of the activation period for each cell within the moving bump, (2) the number of neurons comprising the neural sequence, and (3) neural recruitment lapse, which is the time between pairs of consecutively activated cells (see Fig. 5c). We have shown that the rhythmic neural clock in MPC uses a mixed encoding strategy between temporal scaling and absolute time encoding to represent the tempo. Under the temporal scaling scenario, the activation profile of a neuron in the moving bump is the same between target durations but shrinks for short and elongates for longer tempos. Under the absolute timing strategy, the activation periods are the same across durations, but additional neurons are recruited for longer durations so that the new neurons are active in the last portion of the interval (Zhou et al., 2022). Our results indicate that both the duration and the number of neurons increased as a function of the target tempo in the task, indicating the presence of both types of encoding strategies (Merchant et al., 2015b; Gámez et al., 2019).

Neural sequences have been reported in the striatum (Kim et al., 2013; Gouvêa et al., 2015; Zhou et al., 2020; Mello et al., 2015; Bakhurin et al., 2017; Jin et al., 2009), prefrontal cortex (Kim et al., 2013; Tiganj et al., 2017), and medial premotor areas (Merchant et al., 2015b; Gámez et al., 2019; Crowe et al., 2014; Murakami et al., 2014) in perception and production interval- and beat-based timing tasks across primate and rodent species. Hence, this is a population code that is widespread across the CTBGc core timing network and the prefrontal cortex during different timing paradigms and maybe the most reliable neural population clock. In fact, a recent empirical and computational study suggested that neural sequences are an effective representation of time from the downstream readout point of view (Zhou et al., 2020).

Neural Population Trajectories

The time-varying activity of cell populations can be projected into a low-dimensional state space using dimensional reduction techniques. These techniques capture the covariance of the activity of the cell population across time and can reveal emergent properties that are not present in the single-cell activity (Elsayed & Cunningham, 2017). Indeed, the neural trajectories that we obtained using principal component analysis (PCA) on more than a thousand MPC neurons recorded in the SCT show the following properties (Fig. 5a, d) (Mendoza et al., 2016). First, they have circular dynamics that form a regenerating loop for every produced interval. Notably, the population state during pulse-based predictive timing correlates with the traversed proportion of an interval (relative timing) instead of its absolute

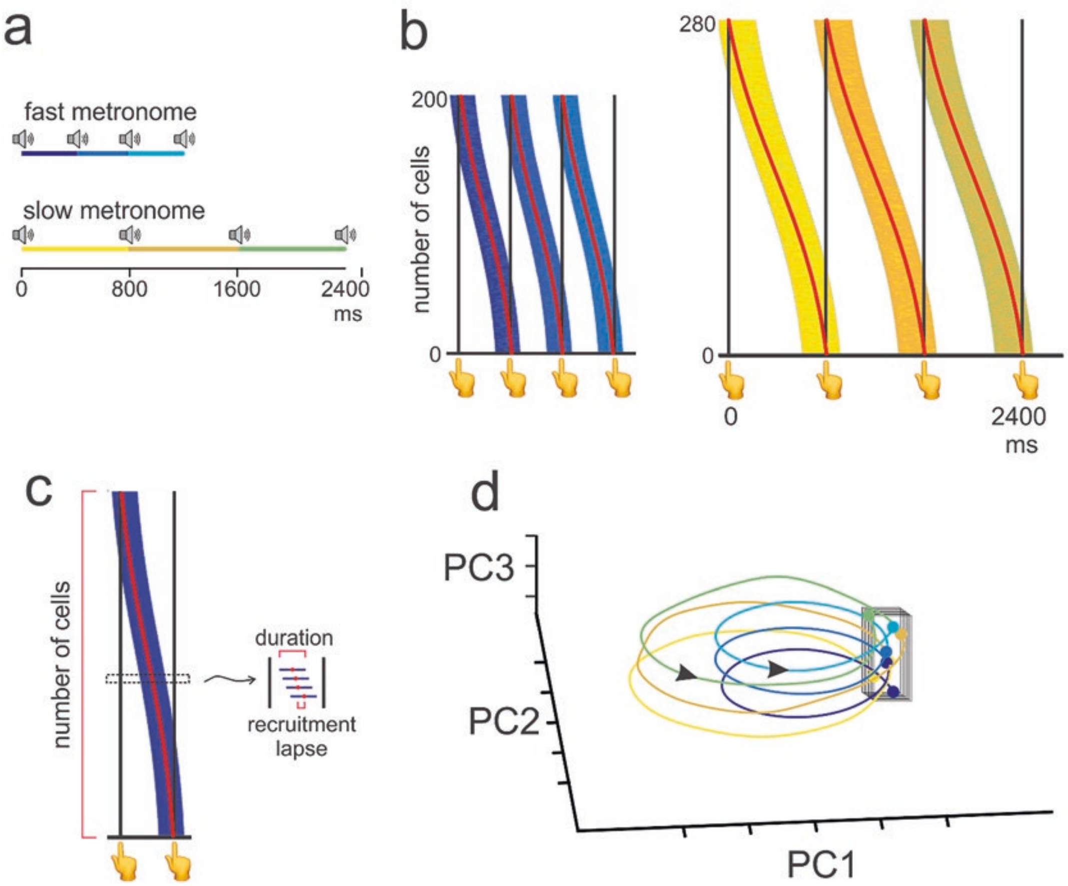


Fig. 5 (**a**) Simplified version of the SCT with only three produced intervals in the synchronization epoch. The small speakers correspond to the isochronous stimuli for the fast and slow tempos. The color code of the three produced intervals is used in b and c to signal the timing neural population codes in the task sequence. (**b**) Neural sequences for the fast (top) and slow (bottom) tempos producing three regenerating moving bumps for each produced interval. Individually colored stripes are constituted by multiple horizontal lines, where each line corresponds to the onset and duration of the activation period for one cell. For the fast tempo, we simulated 200 cells with a mean activation duration of 200 ms, while for the slow tempo we simulated 280 cells with a mean activation duration of 300 ms. All cells were sorted by their peak activation time. (**c**) The three key parameters of neural sequences are outlined in pink: the number of cells in the moving bump, the duration of the activation period of each neuron, and the recruitment lapse that corresponds to the time between the activation peaks of two consecutive neurons. (**d**) Neural trajectories during the synchronization task. The trajectory starts from a tapping separatrix (black cuboid), completes a cycle during every intertap interval, and returns to the tapping separatrix. The separatrix is invariant across durations and serial order elements of the task. The metronome's tempo modulates the amplitude of the trajectories and the serial order element as the third axes in the state population, generating for each produced interval an evolving population pattern of activation. The circles correspond to the tapping times and the arrows specify the direction of the trajectories' movement. Note that a similar population response profile is repeated cyclically for the three intervals (color coded) and that the resetting of each moving bump corresponds to a potential internal pulse representation

magnitude (Balasubramaniam et al., 2021; Gámez et al., 2019). Second, the periodic trajectories increase in amplitude as a function of the tapping interval. These period-dependent increments in the trajectory radius result from a larger number of neurons within a moving bump (Gámez et al., 2019). Finally, the neural population trajectories converge in similar state space at tapping times, resetting the pulse-based clock at this point (Gámez et al., 2019) (Fig. 5c). Hence, the convergence to this neural attractor state could be the internal representation of the pulse transmitted as a phasic top-down predictive signal to the auditory areas before each tap (Lenc et al., 2021).

The notion that beat-based timing depends on a population clock whose dynamics at each instant correspond to a unique pattern in state-space started in 2014 (Merchant et al., 2014). In that article, we suggested that on top of the single-cell encoding of time there was a population timer that represented the relative passage of time, the duration of the produced intervals, and the serial order elements of the tapping sequence during SCT. Since then, there have been many reports of neural population clocks during interval-based timing using neural trajectories in medial premotor areas (Wang et al., 2018; Sohn et al., 2019; Remington et al., 2018; Gámez et al., 2019), prefrontal cortex (Kim et al., 2013; Henke et al., 2021; Xu et al., 2014), and the striatum (Zhou et al., 2020). These neural clocks can compute elapsed time in the final position, time remaining for an action in the time scaling and, flexibly incorporate the task context into the kinematics of the neural trajectories (Henke et al., 2021; Remington et al., 2018; Egger et al., 2019; De Lafuente et al., 2022). Therefore, the neural trajectory clock is now the most accepted neurophysiological mechanism to quantify and predict events in time in both interval- and beat-based timing tasks (Tsao et al., 2022). The parallelisms between the geometric and kinematic properties of neural trajectories and the properties of neural sequences have also been thoroughly documented (Betancourt et al., 2023).

Integrating Layers of Neuronal Clocks in the Medial Premotor Cortex

Our neurophysiological recordings in behaving animals indicate that MPC, an area of the core timing mechanism (Merchant et al., 2013a), uses multiple encoding strategies and different organization levels to represent diverse aspects of the temporal and sequential structure of the SCT. The different types of ramping activity; the cells with mixed selectivity to tempo duration, serial order, and elapsed time; the progressive patterns of neuronal activation; and the neural trajectories must be interlinked to generate a coordinated population clock that flexibly processes temporal information across a wide variety of tasks. Consequently, it is crucial to generate metrics that define the rules of interaction between these temporal signals and to determine the bidirectional effects of lower to upper levels of the neural organization on timing. It is also fundamental to determine the interaction between the neural clock and other task components such as decision making, reinforcement learning, the calculation of value, and the impact of previous trials of the task-solving strategy. Important attempts to tackle the first issue come from our lab, where we have documented that different types of ramping activity are active at different stages of the neural sequences, that duration-tuned cells are mainly engaged in the intermediate part of the moving bumps, and that the geometry and kinematics of neural trajectories have a counterpart in the properties of the neural sequences (Betancourt et al., 2023; Gámez et al., 2019). Recently, the group of Thurley also linked the properties of mixed selectivity, ramping, neural sequences, and neural trajectories (Henke et al., 2021) during single interval measurement and reproduction in the prefrontal cortex of gerbils. Nevertheless, an integrated analysis on the timing mechanisms across tasks and species is lacking and urgently needed.

Structural Bases for Beat-Based Timing During SCT

Although rhythmic entrainment is prevalent across all human cultures (Nettle, 2000), there are wide individual differences in the precision, accuracy, and predictability of movement synchronization (García-Saldivar et al., 2022). A critical unanswered question is what the structural bases for these differences in SCT are. To address this issue, we first obtained diffusion-weighted images from human subjects who also performed the SCT with auditory or visual metronomes and five tempos ranging from 550 to 950 ms. Then, we developed a method to determine the fiber density of U-fibers running tangentially to the cortex (García-Saldivar et al., 2022), which are the core fiber system for corticocortical associations (Schüz & Braitenberg, 2002). Notably, the right audio-motor system (including MPC) showed individual differences in the density of U-fibers that were highly correlated with the degree of predictive entrainment across subjects, measured with the asynchronies during the SCT. These correlations were selective for the synchronization epoch and were specific for auditory metronomes with tempos around 700 ms. In addition, we found that the predictive rhythmic entrainment abilities of subjects were significantly associated with the density and bundle diameter of the corpus callosum (CC), forming a chronotopic map where behavioral correlations of short and long intervals were found with the anterior and posterior portions of the CC (García-Saldivar et al., 2022). Consequently, these findings support the notion that the structural properties of the superficial and deep white matter of the audio-motor system define the predictive abilities of subjects during rhythmic tapping.

A similar methodology can be applied to determine the structural changes in the white matter and the plastic modifications of cortical thickness due to intensive learning in nonhuman primates (Garcia-Saldivar et al., 2021; Messinger et al., 2021; Poirier et al., 2021). Indeed, a longitudinal approach can be used to collect and analyze image data at different stages of learning to determine the effects of training on rhythmic perception and entrainment tasks in macaques (Song et al., 2021; Milham et al., 2020). Preliminary observations from our lab suggest that the density of superficial white matter in the cortical areas of the core timing network (Merchant et al., 2013a) undergoes learning-induced changes. Crucially, this structural information can be used in conjunction with behavioral and neurophysiological information to generate a multimodal map that combines all these data in the same structural framework. This integrated construct can be used to understand more deeply the anatomofunctional correlates of rhythmic behavior for each animal and across animals (Garcia-Saldivar et al., 2021).

Some Considerations on the Brain Mechanisms Behind Interval- and Beat-Based Timing

The recent myriad of neurophysiological studies on the neural basis of timing have mainly focused on the timing signals of the striatum and medial prefrontal cortex in rodents performing single interval tasks (Gouvêa et al., 2015; Mello et al., 2015; Bakhurin et al., 2017; Jin et al., 2009; Tiganj et al., 2017; Milham et al., 2020; Emmons et al., 2017). These papers have documented ramping activity, neural sequences, and trajectories with time scaling properties. The main principle behind these studies is that time can be represented as stable, repeatable activity patterns in single cells so that the overall population state activity systematically changes as a function of time within these two connected brain areas. This rule is followed on tasks in the range of hundreds of milliseconds and seconds (Tsao et al., 2022). On the other hand, the human imaging and lesion literature support the hypothesis of functionally nonoverlapping mechanisms of interval- and beat-based timing, with the separable contributions of the cerebellum and basal ganglia to these two types of temporal processing (Grube et al., 2010a; Breska & Ivry, 2018; Teki et al., 2012). Obviously, the rodent neurophysiology on single interval tasks in the striatum contradicts this hypothesis. All these studies claim that the basal ganglia are a key component of single interval timing. This discrepancy could be due to several factors, including: (1) the limited generalization in the brain anatomy of the rodent and the human, especially regarding the frontal lobe and the basal ganglia (Mendoza & Merchant, 2014); (2) the limited range of timing behaviors than can be trained in rodents, compared with the immense flexibility and behavioral repertoires in humans for measuring time; (3) the methodological restrictions of measuring the neural activity in human using the slow BOLD signal (seconds resolution) or the inherent problems of using behavior in neurological patients to understand brain mechanisms (Merchant et al., 2008c). Needless to say, the neurophysiology in nonhuman primates can partially address these issues since the anatomy, physiology, and behavioral spectrum in monkeys is quite close to that of humans (Mendoza & Merchant, 2014; Merchant et al., 2003b, 2004; Honing & Merchant, 2014). In particular, invasive high-density electrodes can be placed in many areas simultaneously during task performance (Mendoza et al., 2016), obtaining neurophysiological signals that can reveal a potential neural clock generalizable to the *Homo sapiens*. Recordings in the primate putamen on single interval reproduction tasks have revealed neural signals with the same properties observed in rodents (Wang et al., 2018; Sohn et al., 2019); while experiments in monkeys performing a rhythmic event detection task have shown that the cerebellum provides more accurate time predicting information than the caudate (Kameda et al., 2019). Furthermore, in a rhythmic saccadic task where monkeys show predictive timing behavior (Takeya et al., 2017), the cerebellum shows a complex set of responses, including predictive rhythmic activity (Okada et al., 2022). Therefore, these studies clearly contradict the notion of a dissociation in the mechanisms for interval- and beat-based timing, and instead suggest that both the basal ganglia and the cerebellum play important roles in timing both single and periodic events in coordination with the medial premotor areas. Further experiments with simultaneous recordings across these circuits in macaques performing both single and rhythmic timing tasks are urgently needed to have a clear notion of how the basal ganglia and the cerebellum dynamically encode temporal information in conjunction with the premotor areas.

Acknowledgments We are very grateful for the wonderful and valuable comments that Vani Rajendran provided to our manuscript. We thank Raul Paulín for his technical assistance. This work was supported by Consejo Nacional de Ciencia y Tecnología (CONACYT) Grant CONACYT: A1-S-8430, UNAM-DGAPA-PAPIIT IN201721, and SECITI 2342.

References

Akdoğan, B., & Balcı, F. (2016). Stimulus probability effects on temporal bisection performance of mice (Mus musculus). *Animal Cognition, 19*, 15–30. https://doi.org/10.1007/s10071-015-0909-6

Allan, L. G. (2002a). The location and interpretation of the bisection point. *The Quarterly Journal of Experimental Psychology Section B, 55*, 43–60. https://doi.org/10.1080/02724990143000162

Allan, L. G. (2002b). Are the referents remembered in temporal bisection? *Learning and Motivation, 33*, 10–31. https://doi.org/10.1006/lmot.2001.1097

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

Assaneo, M. F., Rimmele, J. M., Sanz Perl, Y., & Poeppel, D. (2021). Speaking rhythmically can shape hearing. *Nature Human Behaviour, 5*, 71–82. https://doi.org/10.1038/s41562-020-00962-0

Ayala, Y. A., Lehmann, A., & Merchant, H. (2017). Monkeys share the neurophysiological basis for encoding sound periodicities captured by the frequency-following response with humans. *Scientific Reports, 7*, 16687. https://doi.org/10.1038/s41598-017-16774-8

Bakhurin, K. I., Goudar, V., Shobe, J. L., Claar, L. D., Buonomano, D. V., & Masmanidis, S. C. (2017). Differential encoding of time by prefrontal and striatal network dynamics. *The Journal of Neuroscience, 37*, 854–870. https://doi.org/10.1523/JNEUROSCI.1789-16.2016

Balasubramaniam, R., Haegens, S., Jazayeri, M., Merchant, H., Sternad, D., & Song, J. H. (2021). Neural encoding and representation of time for sensorimotor control and learning. *The Journal of Neuroscience, 41*, 866–872. https://doi.org/10.1523/JNEUROSCI.1652-20.2020

Bannier, D., Wearden, J., Le Dantec, C. C., & Rebaï, M. (2019). Differences in the temporal processing between identification and categorization of durations: A behavioral and ERP study. *Behavioural Brain Research, 356*, 197–203. https://doi.org/10.1016/j.bbr.2018.08.027

Bartolo, R., & Merchant, H. (2009). Learning and generalization of time production in humans: Rules of transfer across modalities and interval durations. *Experimental Brain Research, 197*, 91–100. https://doi.org/10.1007/s00221-009-1895-1

Bartolo, R., & Merchant, H. (2015). β oscillations are linked to the initiation of sensory-cued movement sequences and the internal guidance of regular tapping in the monkey. *The Journal of Neuroscience, 35*, 4635–4640. https://doi.org/10.1523/JNEUROSCI.4570-14.2015

Bartolo, R., Prado, L., & Merchant, H. (2014). Information processing in the primate basal ganglia during sensory-guided and internally driven rhythmic tapping. *The Journal of Neuroscience, 34*, 3910–3923. https://doi.org/10.1523/JNEUROSCI.2679-13.2014

Betancourt, A., Pérez, O., Gámez, J., Mendoza, G., & Merchant, H. (2023). Amodal population clock in the primate medial premotor system for rhythmic tapping. *Cell Reports, 42*, 113234. https://doi.org/10.1016/j.celrep.2023.113234

Bi, Z., & Zhou, C. (2020). Understanding the computation of time using neural network models. *Proceedings of the National Academy of Sciences, 117*, 10530–10540. https://doi.org/10.1073/pnas.1921609117

Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease. *Proceedings of the National Academy of Sciences, 115*, 12283–12288. https://doi.org/10.1073/pnas.1810596115

Brody, C. D. (2003). Timing and neural encoding of somatosensory parametric working memory in macaque prefrontal cortex. *Cerebral Cortex, 13*, 1196–1207. https://doi.org/10.1093/cercor/bhg100

Cadena-Valencia, J., García-Garibay, O., Merchant, H., Jazayeri, M., & De Lafuente, V. (2018). Entrainment and maintenance of an internal metronome in supplementary motor area. *eLife, 7*. https://doi.org/10.7554/eLife.38983

Cambraia, R., Vasconcelos, M., Jozefowiez, J., & Machado, A. (2019). Biasing performance through differential payoff in a temporal bisection task. *Journal of Experimental Psychology. Animal Learning and Cognition, 45*, 75–94. https://doi.org/10.1037/xan0000192

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Moving on time: Brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *Journal of Cognitive Neuroscience, 20*, 226–239. https://doi.org/10.1162/jocn.2008.20018

Church, R. M., & Deluty, M. Z. (1977). Bisection of temporal intervals. *Journal of Experimental Psychology. Animal Behavior Processes, 3*, 216–228. https://doi.org/10.1037/0097-7403.3.3.216

Collier, G. L., & Ogden, R. T. (2004). Adding drift to the decomposition of simple isochronous tapping: An extension of the Wing-Kristofferson model. *Journal of Experimental Psychology. Human Perception and Performance, 30*, 853–872. https://doi.org/10.1037/0096-1523.30.5.853

Criscuolo, A., Schwartze, M., Prado, L., Merchant, H., & Kotz, S. (2023). Macaque monkeys and humans sample temporal regularities in the acoustic environment. *Progress in Neurobiology, 229*(2023), 102502. https://doi.org/10.1016/j.pneurobio.2023.10250

Crowe, D. A., Zarco, W., Bartolo, R., & Merchant, H. (2014). Dynamic representation of the temporal and sequential structure of rhythmic movements in the primate medial premotor cortex. *The Journal of Neuroscience, 34*, 11972–11983. https://doi.org/10.1523/JNEUROSCI.2177-14.2014

De Lafuente, V., Jazayeri, M., Merchant, H., Gracía-Garibay, O., & Malagón, A. M. (2022). Keeping time and rhythm by replaying a sensory-motor engram. *bioRxiv*. https://doi.org/10.1101/2022.01.03.474812

Di Fabio, P. R., Merchant, H., Bartolo, R., & Tuite, P. (2011). Temporal discrimination learning for treatment of gait dysfunction in Parkinson's disease: A feasibility study using single subject design. *Journal of Parkinsonism and Restless Legs Syndrome, 1*, 8–11. https://doi.org/10.7157/jprls.2011v1n1pp8-11

Donnet, S., Bartolo, R., Fernandes, J. M., Paulo, J., Cunha, S., Prado, L., & Merchant, H. (2014). Monkeys time their pauses of movement and not their movement-kinematics during a synchronization-continuation rhythmic task. *Journal of Neurophysiology, 111*, 2138–2149. https://doi.org/10.1152/jn.00802.2013.-A

Egger, S. W., Remington, E. D., Chang, C. J., & Jazayeri, M. (2019). Internal models of sensorimotor integration regulate cortical dynamics. *Nature Neuroscience, 22*, 1871–1882. https://doi.org/10.1038/s41593-019-0500-6

Elsayed, G. F., & Cunningham, J. P. (2017). Structure in neural population recordings: An expected byproduct of simpler phenomena? *Nature Neuroscience, 20*, 1310–1318. https://doi.org/10.1038/nn.4617

Elsmore, T. F. (1972). Duration discrimination: Effects of probability of stimulus presentation. *Journal of the Experimental Analysis of Behavior, 18*, 465–469. https://doi.org/10.1901/jeab.1972.18-465

Emmons, E. B., De Corte, B. J., Kim, Y., Parker, K. L., Matell, M. S., & Narayanan, N. S. (2017). Rodent medial frontal control of temporal processing in the dorsomedial striatum. *The Journal of Neuroscience, 37*, 8718–8733. https://doi.org/10.1523/JNEUROSCI.1376-17.2017

Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science, 1979*(291), 312–316. https://doi.org/10.1126/science.291.5502.312

Gámez, J., Yc, K., Ayala, Y. A., Dotov, D., Prado, L., & Merchant, H. (2018). Predictive rhythmic tapping to isochronous and tempo changing metronomes in the nonhuman primate. *Annals of the New York Academy of Sciences, 1423*, 396–414. https://doi.org/10.1111/nyas.13671

Gámez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biology, 17*. https://doi.org/10.1371/journal.pbio.3000054

García-Garibay, O., Cadena-Valencia, J., Merchant, H., & de Lafuente, V. (2016). Monkeys share the human ability to internally maintain a temporal rhythm. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.01971

Garcia-Saldivar, P., Garimella, A., Garza-Villarreal, E. A., Mendez, F. A., Concha, L., & Merchant, H. (2021). PREEMACS: Pipeline for preprocessing and extraction of the macaque brain surface. *NeuroImage, 227*, 117671. https://doi.org/10.1016/j.neuroimage.2020.117671

García-Saldivar, P., De León, C., Concha, L., & Merchant, H. (2022). White matter structural bases for predictive tapping synchronization. *bioRxiv*. https://doi.org/10.1101/2022.09.05.506691

Genovesio, A., Tsujimoto, S., & Wise, S. P. (2009). Feature- and order-based timing representations in the frontal cortex. *Neuron, 63*, 254–266. https://doi.org/10.1016/j.neuron.2009.06.018

Georgopoulos, A. P., Merchant, H., Naselaris, T., & Amirikian, B. (2007). Mapping of the preferred direction in the motor cortex. *Proceedings of the National Academy of Sciences, 104*, 11068–11072. https://doi.org/10.1073/pnas.0611597104

Gescheider, G. A. (1997). *Psychophysics*. Psychology Press. https://doi.org/10.4324/9780203774458

Gibbon, J. (1981). On the form and location of the psychometric bisection function for time. *Journal of Mathematical Psychology, 24*, 58–87. https://doi.org/10.1016/0022-2496(81)90035-3

Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology, 7*, 170–184. https://doi.org/10.1016/S0959-4388(97)80005-0

Goldman-Rakic, P. S., Selemon, L. D., & Schwartz, M. L. (1984). Dual pathways connecting the dorsolateral prefrontal cortex with the hippocampal formation and parahippocampal cortex in the rhesus monkey. *Neuroscience, 12*, 719–743. https://doi.org/10.1016/0306-4522(84)90166-0

Gouvêa, T. S., Monteiro, T., Motiwala, A., Soares, S., Machens, C., & Paton, J. J. (2015). Striatal dynamics explain duration judgments. *eLife, 4*. https://doi.org/10.7554/eLife.11386

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *The Journal of Neuroscience, 29*, 7540–7548. https://doi.org/10.1523/JNEUROSCI.2018-08.2009

Grube, M., Cooper, F. E., Chinnery, P. F., & Griffiths, T. D. (2010a). Dissociation of duration-based and beat-based auditory timing in cerebellar degeneration. *Proceedings of the National Academy of Sciences, 107*, 11597–11601. https://doi.org/10.1073/pnas.0910473107

Grube, M., Lee, K.-H., Griffiths, T. D., Barker, A. T., & Woodruff, P. W. (2010b). Transcranial magnetic theta-burst stimulation of the human cerebellum distinguishes absolute, duration-based from relative, beat-based perception of subsecond time intervals. *Frontiers in Psychology, 1*. https://doi.org/10.3389/fpsyg.2010.00171

Henke, J., Bunk, D., von Werder, D., Häusler, S., Flanagin, V. L., & Thurley, K. (2021). Distributed coding of duration in rodent prefrontal cortex during time reproduction. *eLife, 10*. https://doi.org/10.7554/eLife.71612

Honing, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences, 1252*, 85–91. https://doi.org/10.1111/j.1749-6632.2011.06402.x

Honing, H., & Merchant, H. (2014). Differences in auditory timing between human and nonhuman primates. *The Behavioral and Brain Sciences, 37*, 557–558. https://doi.org/10.1017/S0140525X13004056

Honing, H., Merchant, H., Háden, G. P., Prado, L., & Bartolo, R. (2012). Rhesus monkeys (Macaca mulatta) detect rhythmic groups in music, but not the beat. *PLoS One, 7*, e51369. https://doi.org/10.1371/journal.pone.0051369

Honing, H., Bouwer, F.L., Prado, L., and Merchant, H. (2018). Rhesus monkeys (Macaca mulatta) sense isochrony in rhythm, but not the beat: Additional support for the gradual audiomotor evolution hypothesis. 12, 1–15. https://doi.org/10.3389/fnins.2018.00475

Hubel, D. H., & Wiesel, T. N. (1977). Ferrier lecture – Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London – Series B: Biological Sciences, 198*, 1–59. https://doi.org/10.1098/rspb.1977.0085

Iversen, J. R., Patel, A. D., Nicodemus, B., & Emmorey, K. (2015). Synchronization to auditory and visual rhythms in hearing and deaf individuals. *Cognition, 134*, 232–244. https://doi.org/10.1016/j.cognition.2014.10.018

Janata, P., & Grafton, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience, 6*, 682–687. https://doi.org/10.1038/nn1081

Jazayeri, M., & Shadlen, M. N. (2015). A neural mechanism for sensing and reproducing a time interval. *Current Biology, 25*, 2599–2609. https://doi.org/10.1016/j.cub.2015.08.038

Jin, D. Z., Fujii, N., & Graybiel, A. M. (2009). Neural representation of time in cortico-basal ganglia circuits. *Proceedings of the National Academy of Sciences, 106*, 19156–19161. https://doi.org/10.1073/pnas.0909881106

Kameda, M., Ohmae, S., & Tanaka, M. (2019). Entrained neuronal activity to periodic visual stimuli in the primate striatum compared with the cerebellum. *eLife, 8*. https://doi.org/10.7554/eLife.48702

Killeen, P. R., Gregor Fetterman, J., & Bizo, L. A. (1997). *Chapter 3: Time's causes* (pp. 79–131). https://doi.org/10.1016/S0166-4115(97)80055-6

Kim, J., Ghim, J.-W., Lee, J. H., & Jung, M. W. (2013). Neural correlates of interval timing in rodent prefrontal cortex. *The Journal of Neuroscience, 33*, 13834–13847. https://doi.org/10.1523/JNEUROSCI.1443-13.2013

Kononowicz, T. W., & van Rijn, H. (2014). Decoupling interval timing and climbing neural activity: A dissociation between CNV and N1P2 amplitudes. *The Journal of Neuroscience, 34*, 2931–2939. https://doi.org/10.1523/JNEUROSCI.2523-13.2014

Kopec, C. D., & Brody, C. D. (2010). Human performance on the temporal bisection task. *Brain and Cognition, 74*, 262–272. https://doi.org/10.1016/j.bandc.2010.08.006

Kunimatsu, J., Suzuki, T. W., Ohmae, S., & Tanaka, M. (2018). Different contributions of preparatory activity in the basal ganglia and cerebellum for self-timing. *eLife, 7*. https://doi.org/10.7554/eLife.35676

Lebedev, M. A., O'Doherty, J. E., & Nicolelis, M. A. L. (2008). Decoding of temporal intervals from cortical ensemble activity. *Journal of Neurophysiology, 99*, 166–186. https://doi.org/10.1152/jn.00734.2007

Lenc, T., Merchant, H., Keller, P. E., Honing, H., Varlet, M., & Nozaradan, S. (2021). *Mapping between sound, brain and behaviour: Four-level framework for understanding rhythm processing in humans and non-human primates*. Preprint at Royal Society Publishing. https://doi.org/10.1098/rstb.2020.0325

Leon, M. I., & Shadlen, M. N. (2003). Representation of time by neurons in the posterior parietal cortex of the macaque. *Neuron, 38*, 317–327. https://doi.org/10.1016/S0896-6273(03)00185-5

Lindbergh, C. A., & Kieffaber, P. D. (2013). The neural correlates of temporal judgments in the duration bisection task. *Neuropsychologia, 51*, 191–196. https://doi.org/10.1016/j.neuropsychologia.2012.09.001

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics, 53*, 49–70. https://doi.org/10.3758/BF03211715

Madison, G. (2001). Variability in isochronous tapping: Higher order dependencies as a function of intertap interval. *Journal of Experimental Psychology. Human Perception and Performance, 27*, 411–422. https://doi.org/10.1037/0096-1523.27.2.411

Maimon, G., & Assad, J. A. (2006). A cognitive signal for the proactive timing of action in macaque LIP. *Nature Neuroscience, 9*, 948–955. https://doi.org/10.1038/nn1716

McAuley, J. D., & Jones, M. R. (2003). Modeling effects of rhythmic context on perceived duration: A comparison of interval and entrainment approaches to short-interval timing. *Journal of Experimental Psychology. Human Perception and Performance, 29*, 1102–1125. https://doi.org/10.1037/0096-1523.29.6.1102

Meegan, D. V., Aslin, R. N., & Jacobs, R. A. (2000). Motor timing learned without motor training. *Nature Neuroscience, 3*, 860–862. https://doi.org/10.1038/78757

Mello, G. B. M., Soares, S., & Paton, J. J. (2015). A scalable population code for time in the striatum. *Current Biology, 25*, 1113–1122. https://doi.org/10.1016/j.cub.2015.02.036

Mendez, J. C., Prado, L., Mendoza, G., & Merchant, H. (2011). Temporal and spatial categorization in human and non-human primates. *Frontiers in Integrative Neuroscience, 5*. https://doi.org/10.3389/fnint.2011.00050

Méndez, J. C., Pérez, O., Prado, L., & Merchant, H. (2014). Linking perception, cognition, and action: Psychophysical observations and neural network modelling. *PLoS One, 9*, e102553. https://doi.org/10.1371/journal.pone.0102553

Méndez, J. C., Rocchi, L., Jahanshahi, M., Rothwell, J., & Merchant, H. (2017). Probing the timing network: A continuous theta burst stimulation study of temporal categorization. *Neuroscience, 356*, 167–175. https://doi.org/10.1016/j.neuroscience.2017.05.023

Mendoza, G., & Merchant, H. (2014). Motor system evolution and the emergence of high cognitive functions. *Progress in Neurobiology, 122*, 73–93. https://doi.org/10.1016/j.pneurobio.2014.09.001

Mendoza, G., Peyrache, A., Gámez, J., Prado, L., Buzsáki, G., & Merchant, H. (2016). Recording extracellular neural activity in the behaving monkey using a semichronic and high-density electrode system. *Journal of Neurophysiology, 116*, 563–574. https://doi.org/10.1152/jn.00116.2016

Mendoza, G., Méndez, J. C., Pérez, O., Prado, L., & Merchant, H. (2018). Neural basis for categorical boundaries in the primate pre-SMA during relative categorization of time intervals. *Nature Communications, 9*, 1098. https://doi.org/10.1038/s41467-018-03482-8

Merchant, H., & Averbeck, B. B. (2017). The computational and neural basis of rhythmic timing in medial premotor cortex. *The Journal of Neuroscience, 37*, 4552–4564. https://doi.org/10.1523/JNEUROSCI.0367-17.2017

Merchant, H., & Bartolo, R. (2018). *Primate beta oscillations and rhythmic behaviors*. Preprint at Springer-Verlag Wien. https://doi.org/10.1007/s00702-017-1716-9

Merchant, H., & Georgopoulos, A. P. (2006). Neurophysiology of perceptual and motor aspects of interception. *Journal of Neurophysiology, 95*, 1–13. https://doi.org/10.1152/jn.00422.2005

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience*. https://doi.org/10.3389/fnins.2013.00274

Merchant, H., & Pérez, O. (2009). Neurophysiology of interceptive behavior in the primate: Encoding and decoding target parameters in the parietofrontal system. In *Coherent behavior in neuronal networks* (pp. 191–206). Springer. https://doi.org/10.1007/978-1-4419-0389-1_10

Merchant, H., & Pérez, O. (2020). *Estimating time with neural networks*. Preprint at Nature Research. https://doi.org/10.1038/s42256-020-00229-3

Merchant, H., & Yarrow, K. (2016). *How the motor system both encodes and influences our sense of time*. Preprint at Elsevier Ltd. https://doi.org/10.1016/j.cobeha.2016.01.006

Merchant, H., Zainos, A., Hernández, A., Salinas, E., & Romo, R. (1997). Functional properties of primate putamen neurons during the categorization of tactile stimuli. *Journal of Neurophysiology, 77*, 1132–1154. https://doi.org/10.1152/jn.1997.77.3.1132

Merchant, H., Battaglia-Mayer, A., & Georgopoulos, A. P. (2001). Effects of optic flow in motor cortex and area 7a. *Journal of Neurophysiology, 86*, 1937–1954. https://doi.org/10.1152/jn.2001.86.4.1937

Merchant, H., Battaglia-Mayer, A., & Georgopoulos, A. P. (2003a). Functional Organization of Parietal Neuronal Responses to optic-flow stimuli. *Journal* *of Neurophysiology, 90*, 675–682. https://doi.org/10.1152/jn.00331.2003

Merchant, H., Battaglia-Mayer, A., & Georgopoulos, A. P. (2003b). Interception of real and apparent motion targets: Psychophysics in humans and monkeys. *Experimental Brain Research, 152*, 106–112. https://doi.org/10.1007/s00221-003-1514-5

Merchant, H., Battaglia-Mayer, A., & Georgopoulos, A. P. (2004a). Neural responses during interception of real and apparent circularly moving stimuli in motor cortex and area 7a. *Cerebral Cortex, 14*, 314–331. https://doi.org/10.1093/cercor/bhg130

Merchant, H., Battaglia-Mayer, A., & Georgopoulos, A. P. (2004b). Neural responses in motor cortex and area 7a to real and apparent motion. *Experimental Brain Research, 154*, 291–307. https://doi.org/10.1007/s00221-003-1664-5

Merchant, H., Fortes, A., & Georgopoulos, A. P. (2004). Short-term memory effects on the representation of two-dimensional space in the rhesus monkey. *Animal Cognition, 7*. https://doi.org/10.1007/s10071-003-0201-z

Merchant, H., Zarco, W., & Prado, L. (2008a). Do we have a common mechanism for measuring time in the hundreds of millisecond range? Evidence from multiple-interval timing tasks. *Journal of Neurophysiology, 99*, 939–949. https://doi.org/10.1152/jn.01225.2007

Merchant, H., Zarco, W., Bartolo, R., & Prado, L. (2008b). The context of temporal processing is represented in the multidimensional relationships between timing tasks. *PLoS One, 3*, e3169. https://doi.org/10.1371/journal.pone.0003169

Merchant, H., Naselaris, T., & Georgopoulos, A. P. (2008c). Dynamic sculpting of directional tuning in the primate motor cortex during three-dimensional reaching. *The Journal of Neuroscience, 28*, 9164–9172. https://doi.org/10.1523/JNEUROSCI.1898-08.2008

Merchant, H., Crowe, D. A., Robertson, M. S., Fortes, A. F., & Georgopoulos, A. P. (2011a). Top-down spatial categorization signal from prefrontal to posterior parietal cortex in the primate. *Frontiers in Systems Neuroscience, 5*. https://doi.org/10.3389/fnsys.2011.00069

Merchant, H., Zarco, W., Pérez, O., Prado, L., & Bartolo, R. (2011b). Measuring time with different neural chronometers during a synchronization-continuation task. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 19784–19789. https://doi.org/10.1073/pnas.1112933108

Merchant, H., Harrington, D. L., & Meck, W. H. (2013a). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience, 36*, 313–336. https://doi.org/10.1146/annurev-neuro-062012-170349

Merchant, H., Perez, O., Zarco, W., & Gamez, J. (2013b). Interval tuning in the primate medial premotor cortex as a general timing mechanism. *The Journal of Neuroscience, 33*, 9082–9096. https://doi.org/10.1523/JNEUROSCI.5513-12.2013

Merchant, H., Bartolo, R., Pérez, O., Méndez, J. C., Mendoza, G., Gámez, J., Yc, K., & Prado, L. (2014). *Neurophysiology of timing in the hundreds of milliseconds: Multiple layers of neuronal clocks in the medial premotor areas* (pp. 143–154). https://doi.org/10.1007/978-1-4939-1782-2_8

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015a). Finding the beat: A neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 370*, 20140093. https://doi.org/10.1098/rstb.2014.0093

Merchant, H., Pérez, O., Bartolo, R., Méndez, J. C., Mendoza, G., Gámez, J., Yc, K., & Prado, L. (2015b). Sensorimotor neural dynamics during isochronous tapping in the medial premotor cortex of the macaque. *The European Journal of Neuroscience, 41*, 586–602. https://doi.org/10.1111/ejn.12811

Merritt, D. J., Casasanto, D., & Brannon, E. M. (2010). Do monkeys think in metaphors? Representations of space and time in monkeys and humans. *Cognition, 117*, 191–202. https://doi.org/10.1016/j.cognition.2010.08.011

Messinger, A., Sirmpilatze, N., Heuer, K., Loh, K. K., Mars, R. B., Sein, J., Xu, T., Glen, D., Jung, B., Seidlitz, J., et al. (2021). A collaborative resource platform for non-human primate neuroimaging. *NeuroImage, 226*, 117519. https://doi.org/10.1016/j.neuroimage.2020.117519

Milham, M., Petkov, C. I., Margulies, D. S., Schroeder, C. E., Basso, M. A., Belin, P., Fair, D. A., Fox, A., Kastner, S., Mars, R. B., et al. (2020). Accelerating the evolution of nonhuman primate neuroimaging. *Neuron, 105*, 600–603. https://doi.org/10.1016/j.neuron.2019.12.023

Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., & Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nature Neuroscience, 12*, 502–507. https://doi.org/10.1038/nn.2272

Mountcastle, V. B. (1998). *Perceptual neuroscience: The cerebral cortex*. Harvard University Press.

Murakami, M., Vicente, M. I., Costa, G. M., & Mainen, Z. F. (2014). Neural antecedents of self-initiated actions in secondary motor cortex. *Nature Neuroscience, 17*, 1574–1582. https://doi.org/10.1038/nn.3826

Nagarajan, S. S., Blake, D. T., Wright, B. A., Byl, N., & Merzenich, M. M. (1998). Practice-related improvements in somatosensory interval discrimination are temporally specific but generalize across skin location, hemisphere, and modality. *The Journal of Neuroscience, 18*, 1559–1570. https://doi.org/10.1523/JNEUROSCI.18-04-01559.1998

Naselaris, T., Merchant, H., Amirikian, B., & Georgopoulos, A. P. (2006a). Large-scale organization of preferred directions in the motor cortex. I. Motor cortical Hyperacuity for forward reaching. *Journal of Neurophysiology, 96*, 3231–3236. https://doi.org/10.1152/jn.00487.2006

Naselaris, T., Merchant, H., Amirikian, B., & Georgopoulos, A. P. (2006b). Large-scale organization of preferred directions in the motor cortex. II. Analysis of local distributions. *Journal of Neurophysiology, 96*, 3237–3247. https://doi.org/10.1152/jn.00488.2006

Nettle, D. (2000). Linguistic diversity, population spread and time depth. In *Time depth in historical linguistics* (pp. 665–677).

Ng, K. K., Tobin, S., & Penney, T. B. (2011). Temporal accumulation and decision processes in the duration bisection task revealed by contingent negative variation. *Frontiers in Integrative Neuroscience, 5*. https://doi.org/10.3389/fnint.2011.00077

Nozaradan, S. (2014). *Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging*. Preprint at Royal Society of London. https://doi.org/10.1098/rstb.2013.0393

Nozaradan, S., Schönwiesner, M., Caron-Desrochers, L., & Lehmann, A. (2016). Enhanced brainstem and cortical encoding of sound during synchronized movement. *NeuroImage, 142*, 231–240. https://doi.org/10.1016/j.neuroimage.2016.07.015

Nozaradan, S., Mouraux, A., Jonas, J., Colnat-Coulbois, S., Rossion, B., & Maillard, L. (2017). Intracerebral evidence of rhythm transform in the human auditory cortex. *Brain Structure & Function, 222*, 2389–2404. https://doi.org/10.1007/s00429-016-1348-0

Ohmae, S., Kunimatsu, J., & Tanaka, M. (2017). Cerebellar roles in self-timing for sub- and supra-second intervals. *The Journal of Neuroscience, 37*, 3511–3522. https://doi.org/10.1523/JNEUROSCI.2221-16.2017

Okada, K., Takeya, R., & Tanaka, M. (2022). Neural signals regulating motor synchronization in the primate deep cerebellar nuclei. *Nature Communications, 13*, 2504. https://doi.org/10.1038/s41467-022-30246-2

Oshio, K., Chiba, A., & Inase, M. (2008). Temporal filtering by prefrontal neurons in duration discrimination. *The European Journal of Neuroscience, 28*, 2333–2343. https://doi.org/10.1111/j.1460-9568.2008.06509.x

Patel, A. D. (2018). Music as a transformative technology of the mind: An update. In *The origins of musicality*. The MIT Press. https://doi.org/10.7551/mitpress/10636.003.0009

Pérez, O., & Merchant, H. (2018). The synaptic properties of cells define the hallmarks of interval timing in a recurrent neural network. *The Journal of Neuroscience, 38*, 4186–4199. https://doi.org/10.1523/JNEUROSCI.2651-17.2018

Perez, O., Kass, R. E., & Merchant, H. (2013). Trial time warping to discriminate stimulus-related from movement-related neural activity. *Journal of Neuroscience Methods, 212*, 203–210. https://doi.org/10.1016/j.jneumeth.2012.10.019

Pérez, O., Delle Monache, S., Lacquaniti, F., Bosco, G., & Merchant, H. (2023). Rhythmic tapping to a moving beat motion kinematics overrules natural gravity. *iScience, 26*, 107543. https://doi.org/10.1016/j.isci.2023.107543

Perrett, S. P. (1998). Temporal discrimination in the cerebellar cortex during conditioned eyelid responses. *Experimental Brain Research, 121*, 115–124. https://doi.org/10.1007/s002210050443

Poirier, C., Hamed, S. B., Garcia-Saldivar, P., Kwok, S. C., Meguerditchian, A., Merchant, H., Rogers, J., Wells, S., & Fox, A. S. (2021). Beyond MRI: On the scientific value of combining non-human primate neuroimaging with metadata. *NeuroImage, 228*, 117679. https://doi.org/10.1016/j.neuroimage.2020.117679

Protopapa, F., Hayashi, M. J., Kulashekhar, S., van der Zwaag, W., Battistella, G., Murray, M. M., Kanai, R., & Bueti, D. (2019). Chronotopic maps in human supplementary motor area. *PLoS Biology, 17*, e3000026. https://doi.org/10.1371/journal.pbio.3000026

Rajendran, V. G., Teki, S., & Schnupp, J. W. H. (2018). Temporal processing in audition: Insights from music. *Neuroscience, 389*, 4–18. https://doi.org/10.1016/j.neuroscience.2017.10.041

Rao, S. M., Harrington, D. L., Haaland, K. Y., Bobholz, J. A., Cox, R. W., & Binder, J. R. (1997). Distributed neural systems underlying the timing of movements. *The Journal of Neuroscience, 17*, 5528–5535.

Remington, E. D., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible sensorimotor computations through rapid reconfiguration of cortical dynamics. *Neuron, 98*, 1005–1019.e5. https://doi.org/10.1016/j.neuron.2018.05.020

Renoult, L., Roux, S., & Riehle, A. (2006). Time is a rubberband: Neuronal activity in monkey motor cortex in relation to time estimation. *The European Journal of Neuroscience, 23*, 3098–3108. https://doi.org/10.1111/j.1460-9568.2006.04824.x

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review, 12*, 969–992. https://doi.org/10.3758/BF03206433

Romo, R., Merchant, H., Ruiz, S., Crespo, P., & Zainos, A. (1995). Neuronal activity of primate putamen during categorical perception of somaesthetic stimuli. *Neuroreport, 6*, 1013–1017. https://doi.org/10.1097/00001756-199505090-00016

Romo, R., Merchant, H., Zainos, A., & Hernández, A. (1996). Categorization of somaesthetic stimuli: Sensorimotor performance and neuronal activity in primary somatic sensory cortex of awake monkeys. *Neuroreport, 7*, 1273–1279.

Sánchez-Moncada, I., Concha, L., & Merchant, H. (2020). The timing network is engaged in the practice of internally driven tapping independently of the learning transfer from perceptual to motor timing. https://doi.org/10.1101/2020.12.17.423301

Schüz, A., & Braitenberg, V. (2002). The human cortical white matter: Quantitative aspects of cortico-cortical long-range connectivity. In *Cortical areas* (pp. 389–398). CRC Press.

Schwartze, M., Rothermich, K., & Kotz, S. A. (2012). Functional dissociation of pre-SMA and SMA-proper in temporal processing. *NeuroImage, 60*, 290–298. https://doi.org/10.1016/j.neuroimage.2011.11.089

Sohn, H., Narain, D., Meirhaeghe, N., & Jazayeri, M. (2019). Bayesian computation through cortical latent dynamics. *Neuron, 103*, 934–947.e5. https://doi.org/10.1016/j.neuron.2019.06.012

Song, X., García-Saldivar, P., Kindred, N., Wang, Y., Merchant, H., Meguerditchian, A., Yang, Y., Stein, E. A., Bradberry, C. W., Ben Hamed, S., et al. (2021). Strengths and challenges of longitudinal non-human primate neuroimaging. *NeuroImage, 236*, 118009. https://doi.org/10.1016/j.neuroimage.2021.118009

Stubbs, D. A. (1976). Response bias and the discrimination of stimulus duration. *Journal of the Experimental Analysis of Behavior, 25*, 243–250. https://doi.org/10.1901/jeab.1976.25-243

Takeya, R., Kameda, M., Patel, A. D., & Tanaka, M. (2017). Predictive and tempo-flexible synchronization to a visual metronome in monkeys. *Scientific Reports, 7*, 6127. https://doi.org/10.1038/s41598-017-06417-3

Tanaka, M. (2007). Cognitive signals in the primate motor thalamus predict saccade timing. *The Journal of Neuroscience, 27*, 12109–12118. https://doi.org/10.1523/JNEUROSCI.1873-07.2007

Tanji, J. (2001). Sequential organization of multiple movements: Involvement of cortical motor areas. *Annual Review of Neuroscience, 24*, 631–651. https://doi.org/10.1146/annurev.neuro.24.1.631

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience, 31*, 3805–3812. https://doi.org/10.1523/JNEUROSCI.5561-10.2011

Teki, S., Grube, M., & Griffiths, T. (2012). A unified model of time perception accounts for duration-based and beat-based timing mechanisms. *Frontiers in Integrative Neuroscience, 5*, 90. https://doi.org/10.3389/fnint.2011.00090

Tiganj, Z., Jung, M. W., Kim, J., & Howard, M. W. (2017). Sequential firing codes for time in rodent medial prefrontal cortex. *Cerebral Cortex, 27*, 5663–5671. https://doi.org/10.1093/cercor/bhw336

Tsao, A., Yousefzadeh, S. A., Meck, W. H., Moser, M.-B., & Moser, E. I. (2022). The neural bases for timing of durations. *Nature Reviews. Neuroscience, 23*, 646–665. https://doi.org/10.1038/s41583-022-00623-3

Wang, J., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible timing by temporal scaling of cortical responses. *Nature Neuroscience, 21*, 102–112. https://doi.org/10.1038/s41593-017-0028-6

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

Wearden, J. H. (1992). Temporal generalization in humans. *Journal of Experimental Psychology. Animal Behavior Processes, 18*, 134–144. https://doi.org/10.1037/0097-7403.18.2.134

Wehrman, J. J., Wearden, J., & Sowman, P. (2020). Decisional carryover effects in interval timing: Evidence of a generalized response bias. *Attention, Perception, & Psychophysics, 82*, 2147–2164. https://doi.org/10.3758/s13414-019-01922-1

Wing, A. M. (2002). Voluntary timing and brain function: An information processing approach. *Brain and Cognition, 48*, 7–30. https://doi.org/10.1006/brcg.2001.1301

Woodrow, H. (1930). The reproduction of temporal intervals. *Journal of Experimental Psychology, 13*, 473–499. https://doi.org/10.1037/h0070462

Woodrow, H. (1934). The temporal indifference interval determined by the method of mean error. *Journal of Experimental Psychology, 17*, 167–188. https://doi.org/10.1037/h0070235

Xu, M., Zhang, S. Y., Dan, Y., & Poo, M. M. (2014). Representation of interval timing by temporally scalable firing patterns in rat prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America, 111*, 480–485. https://doi.org/10.1073/pnas.1321314111

Yc, K., Prado, L., & Merchant, H. (2019). The scalar property during isochronous tapping is disrupted by a D2-like agonist in the nonhuman primate. *Journal of Neurophysiology, 121*, 940–949. https://doi.org/10.1152/jn.00804.2018

Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology, 102*, 3191–3202. https://doi.org/10.1152/jn.00066.2009

Zhou, S., Masmanidis, S. C., & Buonomano, D. V. (2020). Neural sequences as an optimal dynamical regime for the readout of time. *Neuron, 108*, 651–658.e5. https://doi.org/10.1016/j.neuron.2020.08.020

Zhou, S., Masmanidis, S. C., & Buonomano, D. V. (2022). Encoding time in neural dynamic regimes with distinct computational tradeoffs. *PLoS Computational Biology, 18*, e1009271. https://doi.org/10.1371/journal.pcbi.1009271

The Interactions of Temporal and Sensory Representations in the Basal Ganglia

Pavel E. Rueda-Orozco, Ana E. Hidalgo-Balbuena,
Perla González-Pereyra, Mario G. Martinez-Montalvo,
and Ana S. Báez-Cordero

Abstract

In rodents and primates, interval estimation has been associated with a complex network of cortical and subcortical structures where the dorsal striatum plays a paramount role. Diverse evidence ranging from individual neurons to population activity has demonstrated that this area hosts temporal-related neural representations that may be instrumental for the perception and production of time intervals. However, little is known about how temporal representations interact with other well-known striatal representations, such as kinematic parameters of movements or somatosensory representations. An attractive hypothesis suggests that somatosensory representations may serve as the scaffold for complex representations such as elapsed time. Alternatively, these representations may coexist as independent streams of information that could be integrated into downstream nuclei, such as the substantia nigra or the globus pallidus. In this review, we will revise the available information suggesting an instrumental role of sensory representations in the construction of temporal representations at population and single-neuron levels throughout the basal ganglia.

Keywords

Basal ganglia · Timing · Action · Sensory processing

Introduction

The basal ganglia (BG) are a group of subcortical nuclei implicated in motor function. While their specific function is still debated, multiple streams of evidence in human and nonhuman primates, felines, and rodents indicate their involvement in action selection, postural control, and modulation of movement parameters, such as vigor, timing, and kinematics (Graybiel, 2008; Redgrave et al., 2010; Turner & Desmurget, 2010; Dudman & Krakauer, 2016; Yin, 2017; Robbe, 2018). In rodents, the main nucleus of the BG is the striatum, which in turn can be subdivided into three functional regions: the ventral striatum, receiving information mainly from limbic regions including the medial and orbital prefrontal cortex; the dorsomedial striatum, receiving information mainly from dorsolateral prefrontal and premotor cortices; and the dorsolateral striatum (DLS), receiving information mainly from sensorimotor

---

P. E. Rueda-Orozco (✉) · A. E. Hidalgo-Balbuena
P. González-Pereyra · M. G. Martinez-Montalvo
A. S. Báez-Cordero
Institute of Neurobiology, National Autonomous University of México, Querétaro, Mexico
e-mail: ruedap@unam.mx

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_8

regions of the cortex and thalamus (Redgrave et al., 2010). The DLS, also known as sensorimotor striatum, has been more commonly implicated in motor control and time-related processes (Buhusi & Meck, 2005; Redgrave et al., 2010; Merchant et al., 2013). In humans, BG lesions produce chronometric dysfunctions; for example, patients with bilateral striatal lesions present problems with interval production and, to a lesser extent, with interval perception (Coslett et al., 2010; Schwartze et al., 2015; Nozaradan et al., 2017). Similar results have been found in patients with Parkinson's disease and Huntington's disease who were tested in different interval timing protocols. Both types of patients tended to present strong impairments not only in interval perception but also in interval production (Rowe et al., 2010; Beudel et al., 2013; Honma et al., 2016, 2018). In rodents, bilateral striatal lesions completely abolished timing behavior in a peak-interval timing procedure (Meck, 2006). In a protocol where rats were trained to execute a timed sequence of movements while running on a motorized treadmill, bilateral striatal lesions affected speed control and behavioral timing (Jurado-Parras et al., 2020). While the BG have been associated with various behaviors, little is known about the potential interactions between the well-characterized sensorimotor functions and other relevant functions, such as the perception or production of time intervals. With the objective of creating a conceptual bridge between these apparently different functions, in the following sections, we will analyze the available literature where, in our opinion, timing and sensorimotor functions intersect and arise from the neural dynamics of the same neural populations.

The Dorsolateral Striatum: A Sensorimotor Hub Implicated in Interval Perception and Production

Cortical activity has been typically associated with temporal processing from ramping spiking patterns at the single-cell level in the medial premotor cortex (MPC) (Merchant et al., 2011) to whole cortical regions containing temporal information embedded in the population activity, commonly known as "population clocks" (Buonomano & Laje, 2010). In these population dynamics, temporal information can be extracted from the sequential activation of neurons within the population (Zhou et al., 2020), that is, from the "temporal position" of each cell within an activation sequence triggered by a particular stimulus or behavioral transition (Jin et al., 2009). Furthermore, temporal information can also be extracted from the geometrical characteristics of the cortical population's neural trajectories (Gámez et al., 2019). Altogether, various lines of evidence in primate literature suggest that the MPC carries time-related information at the single-cell and population level. The evidence also supports the notion that in primates, this region is implicated in the perception and production of time intervals. However, sensorimotor cortices massively innervate the striatum, where single-neuron and population temporal-related dynamics have also been reported, making it difficult to understand if the temporal representations in the striatum are partially or completely inherited from cortical regions, or whether partially or completely produced locally (Wang et al., 2018). For example, in primates, striatal spiking activity has been associated with timing perception. In an interval discrimination procedure, the striatal bursting activity of individual neurons was associated with the presentation of different stimulus durations, delay periods, and final choices, suggesting a potential involvement in interval categorization (Chiba et al., 2015). Conversely, when animals were required to produce a particular time interval, individual striatal neurons tracked the passage of time in an adjustable ramping fashion during three different instructed intervals (Kunimatsu et al., 2018). Moreover, important examples of the role of the corticostriatal networks at the single-cell and population levels have recently emerged from the rodent literature. At the single-cell level, rodent cortical dynamics have been shown to present both ramping activity and sequential activation in relation to specific time intervals (Kim et al., 2013; Xu et al., 2014; Henke et al., 2021).

Likewise, subpopulations of DLS neurons have been described to discharge at specific time intervals associated with rewards, suggesting that this brain region could represent behaviorally relevant intervals (Matell et al., 2003). Similar to what has been found in primates, Gouvêau et al. (2015) demonstrated that in rats trained to discriminate between long- and short-time intervals on the scale of a few hundreds of milliseconds, striatal neural population dynamics are organized in sequences of activation, and population trajectories drawn from dimensionally reduced dynamics from the same activity adjust their speed depending on the duration of the interval (Gouvêa et al., 2015). Along the same line, in rats trained to lever press at different fixed-interval ratios ranging from 12 to 60 s, striatal population dynamics covered the interval to be estimated and the sequential activation order of the cells scaled with the target interval (Mello et al., 2015). In another example, Bakhurin and colleagues (Bakhurin et al., 2017) trained mice to lick from a water spout starting 2.5 s after the presentation of a particular odorant. In this work, striatal and cortical population dynamics were analyzed and compared, demonstrating that striatal activity was significantly more useful to decode the passage of time in the 2.5-s interval between the odor presentation and motor response.

The previous examples illustrate the striatal involvement in the perception of time intervals in rodents, but a few other examples have also addressed its participation in time interval production. Rueda-Orozco and Robbe trained rats to execute a timed sequence of accelerations and deaccelerations while running on a motorized treadmill at a fixed speed (Rueda-Orozco & Robbe, 2015). The authors found robust DLS speed and spatial representations of the sequence of movements, but interestingly and in contrast with the previous examples, they only found a few individual cells that presented correlations with elapsed time, suggesting that this variable may be better represented in the ensemble activity. In the same work, the authors were able to compare the sequential activation of neurons associated with execution sequences of movements in expert and naïve animals. While the kinematic representations were specific to the expert rats, the sequential activation elicited by movement execution was present in both groups. This observation suggested a strong learning-independent sequential organization evolving in time and triggered by relevant behavioral events, as a potential temporal structure to contextualize motor commands (Paton & Lau, 2015; Rueda-Orozco & Robbe, 2015). Recently, a similar idea was proposed and tested by Toso et al. (2021). In this work, rats were trained to estimate the duration of patterns of somatosensory stimulation applied to the animal's whiskers. These patterns changed in duration, frequency, and amplitude. As in all the previous examples, the authors found clear DLS sequential activations that spanned for the duration of trials, but inherently, these were decoupled from the animals' judgments of duration. The authors concluded that trial-related temporal representations in the form of sequential activation are inherent to the DLS but have no impact on the judgment of specific intervals presented during the same trial (Toso et al., 2021).

The previous examples support the notion that striatal sequential activations are a robust phenomenon that can be triggered by specific behavioral circumstances, such as odors or cutaneous stimulations, or the beginning of trained (or untrained) movement sequences. In fact, in a recent article, it has been demonstrated that these DLS sequences can also be evoked in anesthetized conditions by applying specific patterns of somatosensory stimulation (Hidalgo-Balbuena et al., 2019). While it is difficult to question the presence of this organized neural unfolding, its specific contribution to timed behavior, if any, is still yet to be fully determined. In this context, two future lines of research may help to clarify this point: first, determining the exact source of these sequential activations and whether they are linked to sensory feedback or can be self-sustained after being triggered by sensory or motor commands from cortical or thalamic inputs to the striatum, and second, manipulating these sequences during different behavioral contexts, for example, during behavioral protocols specifically designed to produce or sense time intervals.

In the following sections, we will address the latest evidence in both directions.

Sensory Processing in the DLS: Beyond Sensory Feedback for Sensorimotor Transformations

The somatotopic sensory representation in the striatum has been described for decades in primates and rodents (West et al., 1990; Carelli & West, 1991; West, 1998; Reiner et al., 2003; Coffey et al., 2016; Hintiryan et al., 2016). These representations have been accepted as the intrinsic anatomical organization of corticostriatal communication, and recently, it has been demonstrated that cortical activity permeates the striatum in a topographically organized fashion (Peters et al., 2021). However, the exact role of these representations is not entirely clear. In the context of motor control, a straightforward possibility is that these representations may be necessary to contextualize action maps and provide sensory feedback for movement control. In humans, somatosensory functions, such as somatosensory discrimination or tactile object shape recognition, have been reported to be impaired in Parkinsonian patients (Weder et al., 1999, 2000), and basic sensory processing and integration appear to be disrupted in rodent models of Parkinson's disease (Ketzef et al., 2017; Peña-Rangel et al., 2021).

How sensory information is processed in the striatum? In this section, we will focus on somatosensory processing, mostly from the whisker and forelimb systems. First, in the rodent, different streams of sensory information reach the DLS from the primary somatosensory cortex (Hoover et al., 2003) and sensory regions of the thalamus (Erro et al., 2001, 2002; Díaz-Hernández et al., 2018; Hidalgo-Balbuena et al., 2019). Second, it has been demonstrated that striatal medium spiny neurons (MSNs) from both direct (dMSNs) and indirect (iMSNs) pathways are able to integrate information from different sensory modalities, such as somatosensory and visual modalities (Reig & Silberberg, 2014; Coffey et al., 2017), and from ipsilateral and contralateral sensorimotor cortices (Reig & Silberberg, 2016). Third, it has been suggested that information is greatly "filtered" in its transition from the cortex to the striatum. In a rodent study, cortical layer 5 neurons from the barrel field projecting to the DLS were antidromically identified by stimulating the DLS. While most of the cortical neurons were responsive to whisker stimulation, only about a third of the neurons recorded in the DLS responded to the same stimulation (Pidoux et al., 2011). Fourth, striatal sensory representations are used to guide learning and reward associations. In an elegant work on the barrel system of mice, whisker stimulation was associated with water rewards, and conditioned responses were defined as "licks" to a waterspout. Whisker-related depolarizations associated with rewards (but not when animals made mistakes) were detected in the membrane potential of MSNs from the DLS (Sippy et al., 2015). In the same work, it was demonstrated that both dMSNs and iMSNs exhibited significant membrane potential depolarizations, but dMSNs presented the shortest response latencies and associated action potentials, while the iMSNs' depolarizations were not sufficient to trigger action potentials. The authors also reported that substituting whisker stimulation (in catch trials) with optogenetic activation of dMSNs, but not iMSNs, efficiently evoked the conditioned response (Sippy et al., 2015). These results are consistent with the role of the BG in sensorimotor transformation, which can either bias the behavior toward a reward or simply initiate a sensory-guided action, in this case "licking." Additional work has shown that the primary motor (M1) and sensory (S1) cortex provide differential inputs to the subpopulations of neurons in the DLS. Lee et al. have proposed that S1 (but not M1) provides stronger inputs to parvalbumin interneurons, typically known as fast-spiking interneurons, than to MSNs (Lee et al., 2019; Johansson & Silberberg, 2020). Interestingly, optogenetic S1-DLS terminal activation during a sensory discrimination protocol induced a response inhibition, while M1-DLS terminal activation induced a facilitated behavioral outcome.

Until recently, striatal sensory information has been studied in the context of stimulus-reward associations, providing a potential basis for movement timing; that is, the moment when movements may be initiated or repressed. However, movement initiation constitutes only a piece of the puzzle. Sensory information throughout the cortico-BG-thalamic circuits is a constant flow that accompanies motor execution, for example, when rodents freely explore the world with whiskers and forelimbs, usually in a rhythmic fashion (West et al., 1990; Shi et al., 2005; Rueda-Orozco & Robbe, 2015; Hidalgo-Balbuena et al., 2019). Hence, the study of sensorimotor processing in the BG has recently benefitted from sensory information analyses in the context of continuous motor execution or stimulation patterns that mimic motor execution. In this context, repetitive sensory stimulation of the whiskers or forelimbs induces different patterns of response adaptation in cortical and DLS neurons (Smith et al., 2012; Peña-Rangel et al., 2021). It has been proposed that different patterns of adaptation (short-term facilitation or inhibition) are related to the propensity of recurrent networks, such as the cortex or striatum, to generate sequential patterns of activation (Carrillo-Reid et al., 2015b; Goel & Buonomano, 2016). These striatal sequential patterns that can be observed in isolated brain slices (Carrillo-Reid et al., 2008) or in rodent behavior (Barbera et al., 2016) may be a mechanism through which a constant sensory flow to the BG helps to sense or to produce time intervals.

Sensory Information in the DLS: A Dynamic Pacemaker to Contextualize Motor Timing

One intrinsic characteristic of sensory systems is that, under particular circumstances, they are entrained to behavioral dynamics. For example, in the case of the olfactory system, the sensation is rhythmically organized and tightly linked to the respiratory rhythm (Cury & Uchida, 2010; Junek et al., 2010). The encoding consequences of this imposed rhythm are still under debate, but one possibility is that breathing may act as an oscillatory pacemaker creating windows for the integration of information at brain-wide scales that may be important for cognitive functions, such as memory consolidation (Karalis & Sirota, 2022). Similarly, rhythmic auditory and visual information have been probed to serve as external pacemakers that entrain cortical and subcortical activity into circular population dynamics to encode beats or complex rhythms which may ultimately guide rhythmic behavior (Cadena-Valencia et al., 2018; Gámez et al., 2019). Another important characteristic is that sensory-evoked responses at the single-cell and population levels appear to share temporal dynamics independently of the sensory modality, suggesting a common scaffold for information processing that can be used for cognitive and motor functions (Luczak et al., 2015). In the somatosensory system, at the single-cell level, these responses can be observed as complex patterns consisting of short-latency activations frequently followed by a transient inactivation and a rebound (Fig. 1a) (Chapin et al., 1981; Hidalgo-Balbuena et al., 2019; Peña-Rangel et al., 2021). These patterns present a variety of response latencies across the population, giving the impression of sequential activation covering hundreds of milliseconds following the stimulation onset (Fig. 1b, c). Furthermore, the temporal architecture of these population dynamics (Fig. 1c) appears to be spontaneous and recruited by the corresponding sensory stream (Luczak et al., 2007, 2009; Carrillo-Reid et al., 2015a). In this context, DLS activity is strongly permeated by topographically organized somatosensory representations (West et al., 1990; West, 1998; Reiner et al., 2003; Hintiryan et al., 2016). From these, forelimb, hindlimb, mouth, and whisker representations would be naturally prone to be rhythmically entrained during locomotion, mastication, or exploration. The whisking cycle, for example, oscillates at around 10 Hz (Deschênes et al., 2016; Sreenivasan & Petersen, 2016). These signals can be recorded in the DLS of unrestrained or anesthetized rodents (Alloway et al., 2009; Smith et al., 2012, 2014). The relevance of these rhythmic signals in the striatum has not been

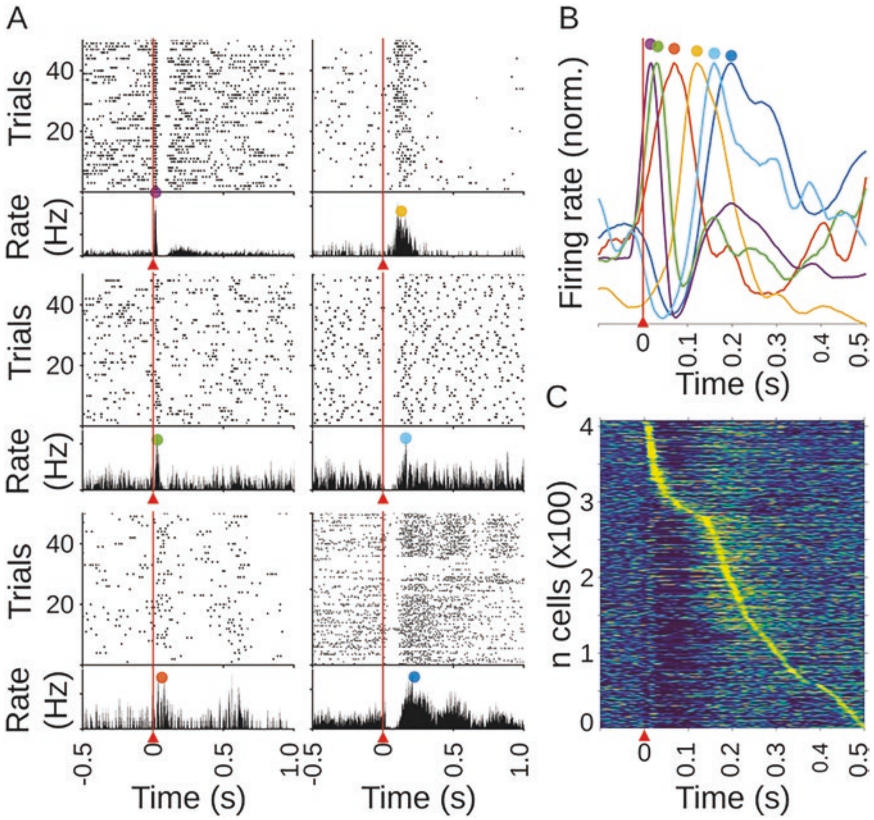


Fig. 1 Sensory-evoked responses organized in sequences covering hundreds of milliseconds. (**a**) Somatosensory-evoked activity in six representative neurons recorded in the dorsolateral striatum (color-coded). Spike rasters and peri-histograms are aligned to the onset of somatosensory stimulation (red lines and arrowheads) of the contralateral forepaw palm to the recording site. For each neuron, the moment of the highest firing rate after stimulus onset is indicated by a colored dot. (**b**) The average activity of neurons displayed in A was normalized to the highest peak, revealing a variety of latencies spanning hundreds of milliseconds. (**c**) Average firing rate (z-scored) for more than 400 neurons recorded in the striatum, sorted by the time of the highest activity after stimulus onset (red arrowhead). Cells displayed in this figure were collected in experiments reported in (Hidalgo-Balbuena et al., 2019; Peña-Rangel et al., 2021)

determined, but one possibility is that it may serve as a pacemaker to facilitate inter-region or interstructure communication. Along the same line of thinking, in rats, the forelimb cycle during locomotion typically oscillates between 2 and 3 Hz depending on movement speed (West et al., 1990; Carelli & West, 1991; Shi et al., 2004). In this case, in previous work from our group, we explored the relevance of these striatal rhythmic signals (Hidalgo-Balbuena et al., 2019). First, by using an anesthetized model where sensory signals can be reliably recorded in sensory regions of cortical, striatal, and nigral regions (Mochol et al., 2015; Peña-Rangel et al., 2021), we induced striatal population sensory responses evoked by 3.3 Hz trains of cutaneous stimuli mimicking the forelimb cycle when animals are running at speeds of ~30 cm/s. This stimulation protocol evoked cortical and striatal population dynamics organized in neural sequences that covered the 300-ms interval between stimuli (Hidalgo-Balbuena et al., 2019; Peña-Rangel et al., 2021). Interestingly, when activity was analyzed throughout the train of 5 stimuli, sensory-evoked striatal population dynamics were more useful to decode elapsed time as compared to the decoding from cortical activity (Hidalgo-Balbuena et al., 2019). These observations suggest a prewired striatal network configuration that can be recalled by sensory stimulation even in artificial experimental conditions and that is independent of motor commands. Such network dynamics have been observed in in vitro conditions in cortical culture cells (Goel & Buonomano, 2016) and similar dynamics have been reported in striatal slices (Carrillo-Reid et al., 2008), supporting the idea that an intrinsic functional property of neural networks is the organization in sequential dynamics from which time can be decoded (Goel & Buonomano, 2014). While this is an attractive alternative, there are still some important questions to answer. For example, what type of timing would these dynamics represent? Would these signals be useful to exclusively perceive time, or would they also be useful to guide timed actions? To address these possibilities, in a previous report, we optogenetically activated/inactivated sensory pathways to the DLS in rats performing a stereotyped sequence of movements in a temporal interval of 7 s. Inhibiting and exciting sensory pathways to the DLS induced the execution of shorter and longer sequences of movements, respectively. Further behavioral analysis revealed that changes in sequence duration were independent of movement speed or motor confounds, confirming that these changes were related to under and overestimation of the temporal interval, respectively (Hidalgo-Balbuena et al., 2019). These data indicate that rhythmic sensory inputs to the DLS may constitute a temporal reference for movement production. In this study, striatal neural dynamics were not assessed during movement execution; hence, it was not possible to determine if the sensory manipulations that induced behavioral changes also impacted DLS population dynamics. In this sense, a more recent work suggests that striatal neural sequences can guide temporal judgments (Monteiro et al., 2023). By manipulating striatal temperature in rodents, authors were able to speed up or slow down both striatal sequential dynamics and temporal judgments, strongly suggesting that scaling striatal temporal population dynamics is the mechanism used to estimate elapsed time (Monteiro et al., 2021). Further experiments would be necessary to elucidate the exact neural mechanism underlying the production of temporal intervals, but in our opinion, the previous evidence strongly suggests that rhythmic entrainment induced by different sensory streams to the BG may play a central role in this function.

Sensory- and Time-Related Signals Throughout Cortico-Basal Ganglia-Thalamic Circuits

The striatum belongs to a macrocircuit that includes BG and extra-BG structures; hence, an important question arises about the diversity of the temporal representations that reach the circuits of the BG. Are striatal sensory and sequential representations distributed homogeneously throughout the different subpopulations of striatal neurons and downstream BG nuclei? MSNs, the main striatal population, are further divided into two subpopulations expressing distinct subclasses: dopaminergic metabotropic receptors D1 and D2 (Albin et al., 1989; Gerfen et al., 1990). MSNs expressing the D1 receptor are the origin of the direct pathway (dMSNs), projecting directly to the output nuclei of the BG, the globus pallidus internal segment (GPi) and the substantia nigra pars reticulata (SNr). MSNs expressing the D2 receptor are the origin of the indirect pathway (iMSNs), projecting to the external segment of the globus pallidus (GPe). Both pathways exert opposing effects over movement control, with dMSNs and iMSNs facilitating and inhibiting movement, respectively (Albin et al., 1989; Gerfen et al., 1990; Kravitz et al., 2010; Cruz et al., 2022). The main source of dopamine to all BG nuclei is the substantia nigra pars compacta (SNc), and it has been reported that SNc neurons encode interval duration and that its activation or inactivation is sufficient to slow down or speed up time interval perception, respectively (Soares et al., 2016). These data suggest that both pathways are permeated by time-related signals, but whether each pathway would represent time or not is still unknown. In this context, a recent work in mice demonstrated that both subpopulations of neurons develop strong and reliable sequential activations associated with forelimb movements (Sheng et al., 2019). Importantly, the experimental design included fixed intertrial intervals, and it was during this period when more iMSNs were implicated in sequential organization. These data suggest that interval representation may not be homogeneously represented in the striatum. Whether these iMSN/dMSN sequential activations are exclusively training-dependent or also spontaneous or triggered by sensory streams is yet to be defined. However, the latter possibility would not be surprising since; for example, it has been shown that whole-body somatosensory representations (Coffey et al., 2017) and rhythmic signals related to licking (Chen et al., 2021) are present in both pathways, and both subpopulations display different patterns of recurrent connectivity with other elements of the striatal microcircuit, such as neurons of the opposite pathway or interneurons (Taverna et al., 2008; Chuhma et al., 2011). In this context, systemic, and intrastriatal administration of D2 but not D1 antagonist produced an overestimation of time intervals in a "peak-interval" procedure (Drew et al., 2003; De Corte et al., 2019; Kamada & Hata, 2021). However, earlier studies suggest that both D1 and D2 receptor agonists and antagonists may produce similar time overestimations and underestimations, respectively (Frederick, 1996). However, Cheung et al. found that systemic administration of quinpirole, a D2 receptor agonist, caused an overestimation of time, while the administration of a D2 antagonist produced no impact on timing (Cheung et al., 2007). More recently, it has been reported that the striatal optogenetic manipulation of the BG's direct pathway produced a temporal representation reset of a fixed interval (Bakhurin et al., 2020). Indirect pathway activation induced a pause on the interval that proportionally recovered as optogenetic stimulation ceased (Bakhurin et al., 2020). There are still many questions on whether the activity of a single pathway would be better to encode elapsed time; however, the available evidence arising from pharmacological and optogenetic research and experiments suggests that this may be the case.

The External Segment of the Globus Pallidus While there is a consensus on the participation of the dopaminergic system in temporal processing, the specific anatomical and functional mechanisms are far from being fully understood. An alternative possibility to explore the role of the different pathways in temporal processing would be to focus on different elements of the BG beyond the striatum. Additionally, time representation may be considered an emergent property of neural networks (striatal or extra-striatal), where the temporal dynamics could be mediated by their connectivity and the intrinsic properties of their elements (Goel & Buonomano, 2014). In this context, the GPe and the subthalamic nucleus (SthN) are anatomically independent entities typically associated with the indirect pathway. The GPe is a GABAergic nucleus considered the first relay of the indirect pathway. It targets the output nuclei of the BG (the internal segment of the GP and the SNr) with reciprocal connections with the striatum and SthN (Kita & Jaeger, 2016). Given that the GPe is centrally located within the BG, studying this area may be an interesting approach to explore the indirect pathway's contribution to movement control and other components of behavior, including timing. In primates and rodents, GPe neurons display firing rates between 30 and 80 Hz and adjust their activity patterns to passive and active movements of the arm and orofacial regions (DeLong, 1971; Alexander & DeLong, 1985). In addition, these neurons can also couple their spiking activity to different temporal phases of movement or movement sequences, firing before, during, or at the end of the sequence (Hegeman et al., 2016). Previous reports on non-human primates also demonstrate somatotopic representations of the different parts of the contralateral hemibody, inheriting the topographical organization from the sensorimotor cortices and the DLS (Bevan, 2002; Jaeger & Kita, 2011; Nambu, 2011; Iwamuro et al., 2017). In terms of intrinsic properties, GPe shares important features with time-related structures, such as the cortex of the striatum. For example, cortical sensory/motor stimulations or even task-relevant events produce complex response patterns in the firing rate of individual GPe neurons, characterized by an initial high-frequency discharge followed by a transient inhibition and a final burst rebound (Alexander et al., 1985; Bevan, 2002; Jaeger & Kita, 2011; Kita & Jaeger, 2016; Iwamuro et al., 2017). These complex response patterns have been described in cortical and striatal regions (Chapin et al., 1981; Hidalgo-Balbuena et al., 2019; Peña-Rangel et al., 2021) and have been proposed as "building blocks" of population spontaneous or evoked packets of information from which temporal representations can be extracted (Goel & Buonomano, 2014; Luczak et al., 2015; Hidalgo-Balbuena et al., 2019) (Fig. 1). In summary, while there are no network analysis studies directly exploring the contribution of the GPe to time interval estimation, the basic network elements observed in other brain regions would make this indirect pathway structure a suitable candidate to explore this possibility.

The Substantia Nigra Pars Reticulata The SNr is one of the two major output nuclei of the BG, receiving mostly inhibitory afferents from the striatum (direct pathway) and excitatory afferents from the SthN (indirect pathway) (Yoshida & Precht, 1971; Dray et al., 1976; Parent et al., 1984). SNr neurons synthesize and release GABA tonically to its main targets, the motor thalamus (ventral medial/lateral; VM/VL) and the dorsal midbrain (Carpenter et al., 1976; Parent et al., 1983; Hikosaka, 2007). In addition, SNr neurons display multiple collateral contacts with other SNr neurons, providing intrinsic feedback to regulate the output of the BG (Brown et al., 2014). SNr neurons show complex associations with sensory events (Nagy et al., 2005; Brown et al., 2014; Báez-Cordero et al., 2020) and different phases of movement, such as preparation and execution (Hikosaka & Wurtz, 1983; Turner & Desmurget, 2010; Schmidt et al., 2013). An interesting feature of SNr neurons is that changes in their firing rate are rarely related to a single event. For example, the same cell might have multisensory responses (Nagy et al., 2005) and be modulated by saccades triggered by visual stimulation. Regarding the potential contribution to temporal processing, Shultz (1986) studied the electrophysiological properties of SNr neurons with extracellular recordings from single neurons in monkeys performing a behavioral GO/NO-GO paradigm, employing an initial preparatory tone, visual stimuli, and forelimb reaching movements. SNr neurons showed consistent changes in activity during performance, and these changes were not simply related to sensory stimuli or forelimb movements. Some neurons increase or decrease their activity during the period between an initial sensory stimulus and the availability to move for reward. Similar complex responses have been seen in SNr neurons in relation to oculomotor mechanisms (Wurtz & Hikosaka, 1986). These complex responses have been typically described as neural mechanisms related to motor planning or movement onset; however, some of their temporal dynamics may also be related to temporal processing. For example, in mice trained to lick a waterspout for reward in a fixed-interval schedule, stimulation of the SNr to the superior colliculus pathway not only canceled licking (motor effect) but also delayed the initiation of anticipatory licking for the next interval, suggesting that SNr may play a role in initiating actions and in adjusting central timing mechanisms (Toda et al., 2017). However, in a cue-guided decision-making licking task in mice (Catanese & Jaeger, 2021), the authors recorded single-unit activity in the VM/VL thalamus; these neurons showed a prominent ramping of activity during a delay epoch in anticipation of a Go cue signaling the onset of movement, suggesting that ramping activity was related to the timing of action initiation. Next, the authors explored whether inhibition from the SNr during the delay epoch could affect behavioral outcomes through an alteration of ramping activity in VM/VL. To test this possibility, they optogenetically stimulated the SNr GABAergic terminals in the VM/VL thalamus during the delay epoch and found a decrease in impulsive licks and an increase in movement omissions. These behavioral changes were accompanied by a decrease in the ramping activity of multiple VM/VL neurons, further supporting the notion that SNr contributes to the appropriate timing of action initiation.

The Motor Thalamus So far, we have revised BG neural dynamics that can be linked to temporal processing. However, the concerted activity of these motor nuclei would be meaningless without a reader. The main target of the output of the BG is the motor thalamus (MTh), a group of nuclei linking subcortical and cortical areas. The MTh can be divided into three regions: ventral anterior (VA), VL, and VM. This complex receives the main projections from the cerebral cortex, the cerebellum, and the BG (Pare et al., 1987; Sommer, 2003). Even when its participation in behavioral outcomes is not completely understood, the evidence suggests that this complex is involved in motor functions such as timing, preparation, initiation, and vigor of movements (Guo et al., 2017; Gaidica et al., 2018; Catanese & Jaeger, 2021; Inagaki et al., 2022). In this context, imaging studies in humans suggest the MTh is part of a network that activates during perceptual timing tasks (Rao et al., 1997; Stevens et al., 2007; Teki et al., 2011). For example, in a synchronization-continuation tapping task for humans, functional magnetic resonance imaging revealed that both conditions produced equivalent activation on the sensorimotor cortex, cerebellum (dentate nucleus), and the right superior temporal gyrus, but only the “continuation” condition produced an activation of the premotor medial system: supplementary motor area, putamen, and thalamus (VL nuclei), suggesting that the internal generation of the motor interval also implicates MTh (Rao et al., 1997). This has been further confirmed in stroke patients whose MTh is affected, producing more variable motor intervals than healthy subjects or brain-injured controls (stroke in another brain region) (Mole et al., 2018). More specific studies from the group of Masaki Tanaka indicate that the MTh may be implicated in motor timing. For example, pharmacological inactivation of the VL thalamus delayed saccades in primates (Tanaka, 2006), and many of the thalamic neurons displayed ramping activity in anticipation of external stimuli that trigger the saccade or self-timed saccades, suggesting that MTh neurons carry preparatory signals that keep track of elapsed time until movement onset (Tanaka, 2007). More recently, the same group demonstrated that subpopulations of MTh neurons can extract temporal predictions from periodic sensory events (Matsuyama & Tanaka, 2021). The mechanisms behind these temporal computations are yet to be determined, but the previous evidence suggests that, as in the BG, time may be embedded as an emergent property of thalamic network dynamics.

Conclusions

In this work, we summarized relevant observations linking sensory processing with timed behavior throughout cortical, BG, and thalamic circuits. The first important observation is that sensory inputs to the sensorimotor striatum produce various robust complex patterns of activation at the single-cell level. These patterns are characterized by brief periods of bursting activity intermingled with silences with variable latencies after stimulation onset. These sensory-evoked responses can be observed under anesthetized conditions, suggesting a prewired configuration, and are not exclusive to the sensorimotor region of the striatum but can also be observed throughout the different relays of the BG, such as the GPe or the SNr, and extra-BG regions such as the MTh and the primary somatosensory and motor cortices (Fig. 2). When these type of responses are plotted together and sorted according to the moment of highest activity, ensemble activity appears to organize as neural sequences covering hundreds of milliseconds after the presentation of a triggering stimulus (even in anesthetized conditions) or a relevant behavioral event (Luczak et al., 2015; Rueda-Orozco & Robbe, 2015; Hidalgo-Balbuena et al., 2019; Monteiro et al., 2021; Peña-Rangel et al., 2021). This sequential dynamic appears to be a robust organization from which time can be decoded, but whose definitive function is still debated. A possible explanation is that these sequences can be a parallel and independent organization to the typical sensorimotor organization that includes visual, auditory, and somatosensory representations, or the representation of task-relevant events or motor-related parameters, such as rewards or kinematic

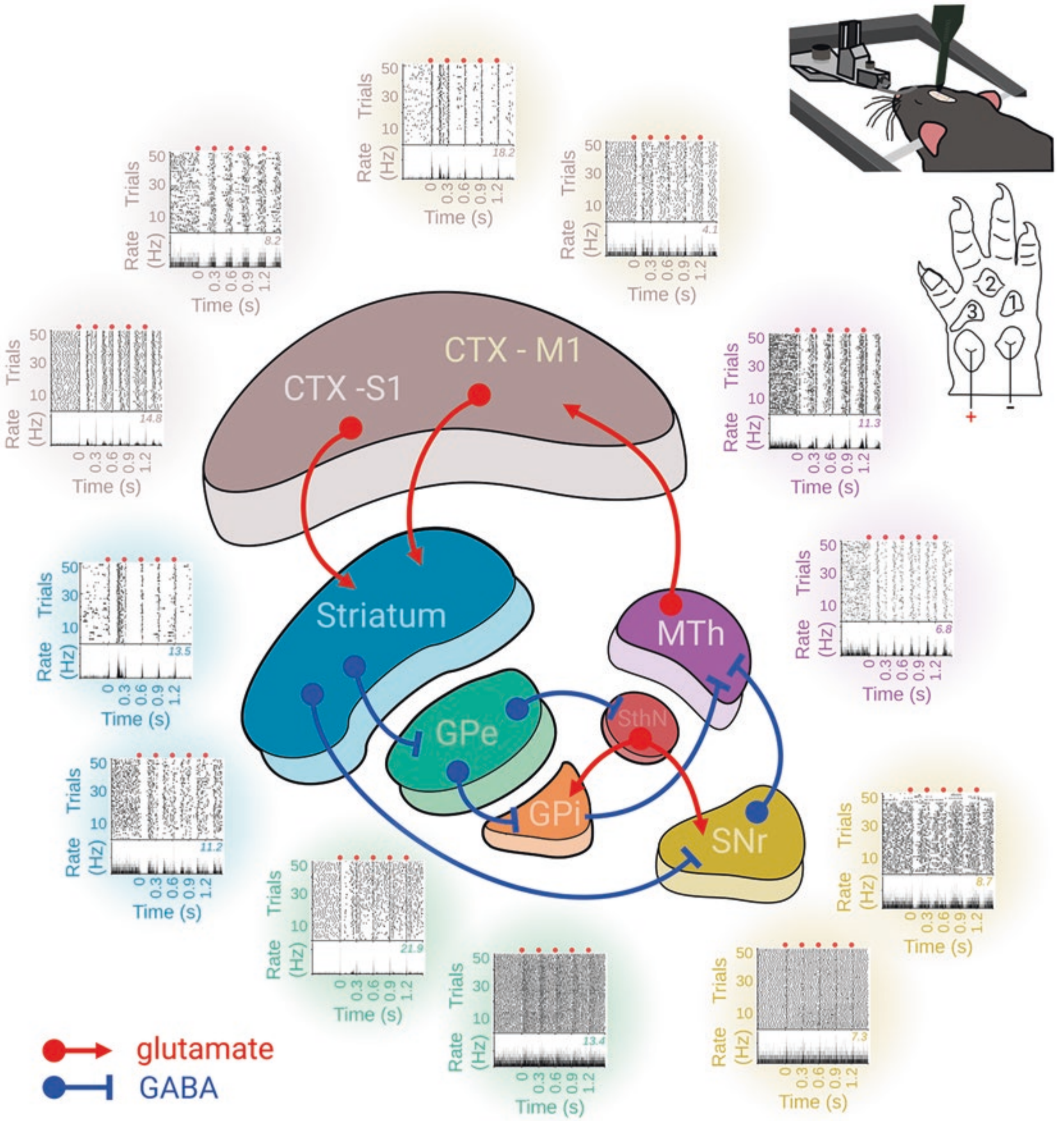


Fig. 2 Sensory representations throughout cortico-basal ganglia-thalamic circuits. Somatosensory-evoked activity of example neurons recorded at different relays within the cortico-BG-thalamic circuit (color-coded). Activity in spike rasters and peri-stimulus histograms is aligned to the first five consecutive stimuli delivered at 3.3 Hz to the forepaw contralateral to the recording sites (upper right corner). All recordings were performed in urethane-anesthetized animals. Representative neurons from primary motor (M1) and sensory (S1) cortices (brown), dorsolateral striatum (blue), and substantia nigra pars reticulata (SNr; golden) were extracted from experiments performed in the following publications (Hidalgo-Balbuena et al., 2019; Báez-Cordero et al., 2020; Peña-Rangel et al., 2021; Pimentel-Farfan et al., 2022). Representative neurons for the external segment of the globus pallidus (green) and the motor thalamus (MTh; purple) were collected in unpublished experiments from our group under identical conditions. No available data for the internal segment of the globus pallidus (GPi; orange) or the subthalamic nucleus (SthN, red). Maximum firing rates are indicated for each neuron on top of the histogram

representations, respectively. In support of this notion, researchers have recorded striatal sequential activation in behaving rats performing sequences of movements (Rueda-Orozco & Robbe, 2015) or sensory discriminations (Toso et al., 2021) and have shown that this activation is independent of the learning stage, kinematic encoding, or sensory discriminative properties of the striatal network. These sequences may provide a general sense of the passage of time, which might be useful to interval perception or to the temporal constraint of motor commands during interval production. These sequences may be started or entrained by sensory events, such as the rhythmic representation of the forelimb during locomotion (Hidalgo-Balbuena et al., 2019) or the clicking sound during lever pressing tasks (Cook et al., 2022).

The fact that the ensemble dynamics recruited by sensory inputs throughout the different relays of the cortico-BG-thalamic circuits share important features (e.g., sequential activation) suggests that to fully understand the role of sensory-evoked activity in perception and movement timing, it is necessary to determine the relationship between the sequential activation observed in different structures, for example, in the striatum and its anatomical targets—the GPe, GPi, and SNr. To this aim, a key challenge is developing experimental approaches to manipulate sequential activation at different stages of these circuits and then evaluating the impact of these manipulations on time perception and time production. This has been partially achieved by specifically manipulating somatosensory pathways to the DLS in animals performing behavioral protocols designed to produce time intervals (Hidalgo-Balbuena et al., 2019). However, it is still necessary to combine this manipulative approach with simultaneous high-density neural recording methods and more complete behavioral protocols where production and perception of time intervals can be compared.

References

Albin, R. L., Young, A. B., & Penney, J. B. (1989). The functional anatomy of basal ganglia disorders. *Trends in Neurosciences, 12*, 366–375. Available at: https://linkinghub.elsevier.com/retrieve/pii/016622368990074X

Alexander, G. E., & DeLong, M. R. (1985). Microstimulation of the primate neostriatum. II. Somatotopic organization of striatal microexcitable zones and their relation to neuronal response properties. *Journal of Neurophysiology, 53*, 1417–1430. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=4009227

Alexander, G. E., DeLong, M. R., & Delong, R. (1985). Microstimulation of the primate neostriatum. I. Physiological properties of striatal microexcitable zones. *Journal of Neurophysiology, 53*, 1401–1416. Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=4009226

Alloway, K. D., Smith, J. B., Beauchemin, K. J., & Olson, M. L. (2009). Bilateral projections from rat MI whisker cortex to the neostriatum, thalamus, and claustrum: Forebrain circuits for modulating whisking behavior. *The Journal of Comparative Neurology, 515*, 548–564. Available at: https://onlinelibrary.wiley.com/doi/10.1002/cne.22073

Báez-Cordero, A. S., Pimentel-Farfan, A. K., Peña-Rangel, T., & Rueda-Orozco, P. E. (2020). Unbalanced inhibitory/excitatory responses in the substantia Nigra Pars Reticulata underlie cannabinoid-related slowness of movements. *The Journal of Neuroscience, 40*, 5769–5784. Available at: http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.0045-20.2020

Bakhurin, K. I., Goudar, V., Shobe, J. L., Claar, L. D., Buonomano, D. V., & Masmanidis, S. C. (2017). Differential encoding of time by prefrontal and striatal network dynamics. *The Journal of Neuroscience, 37*, 854–870. Available at: http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.1789-16.2017

Bakhurin, K. I., Li, X., Friedman, A. D., Lusk, N. A., Watson, G. D. R., Kim, N., & Yin, H. H. (2020). Opponent regulation of action performance and timing by striatonigral and striatopallidal pathways. *eLife, 9*, 1–25. Available at: https://elifesciences.org/articles/54831

Barbera, G., Liang, B., Zhang, L., Gerfen, C. R., Culurciello, E., Chen, R., Li, Y., & Lin, D. T. (2016). Spatially compact neural clusters in the dorsal striatum encode locomotion relevant information. *Neuron, 92*, 202–213. Available at: https://doi.org/10.1016/j.neuron.2016.08.037

Beudel, M., de Geus, C. M., Leenders, K. L., & de Jong, B. M. (2013). Acceleration bias in visually perceived velocity change and effects of Parkinson's bradykinesia. *Neuroreport, 24*, 773–778. Available at: https://journals.lww.com/00001756-201310020-00003

Bevan, M. (2002). Move to the rhythm: Oscillations in the subthalamic nucleus–external globus pallidus network. *Trends in Neurosciences, 25*, 525–531. Available at: https://linkinghub.elsevier.com/retrieve/pii/S016622360202235X

Brown, J., Pan, W.-X., & Dudman, J. T. (2014). The inhibitory microcircuit of the substantia nigra provides feedback gain control of the basal ganglia output. *eLife, 3*, 1–25.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews. Neuroscience, 6*, 755–765. Available at: http://www.nature.com/articles/nrn1764

Buonomano, D. V., & Laje, R. (2010). Population clocks: Motor timing with neural dynamics. *Trends in Cognitive Sciences, 14*, 520–527. Available at: https://doi.org/10.1016/j.tics.2010.09.002

Cadena-Valencia, J., García-Garibay, O., Merchant, H., Jazayeri, M., & de Lafuente, V. (2018). Entrainment and maintenance of an internal metronome in supplementary motor area. *eLife, 7*, 1–23. Available at: https://elifesciences.org/articles/38983

Carelli, R. M., & West, M. O. (1991). Representation of the body by single neurons in the dorsolateral striatum of the awake, unrestrained rat. *The Journal of Comparative Neurology, 309*, 231–249.

Carpenter, M. B., Nakano, K., & Kim, R. (1976). Nigrothalamic projections in the monkey demonstrated by autoradiographic technics. *The Journal of Comparative Neurology, 165*, 401–415. Available at: https://onlinelibrary.wiley.com/doi/10.1002/cne.901650402

Carrillo-Reid, L., Tecuapetla, F., Tapia, D., Hernandez-Cruz, A., Galarraga, E., Drucker-Colin, R., & Bargas, J. (2008). Encoding network states by striatal cell assemblies. *Journal of Neurophysiology, 99*, 1435–1450. Available at: http://jn.physiology.org/cgi/doi/10.1152/jn.01131.2007

Carrillo-Reid, L., Kang Miller, J., Hamm, J. P., Jackson, J., & Yuste, R. (2015a). Endogenous sequential cortical activity evoked by visual stimuli. *The Journal of Neuroscience, 35*, 8813–8828. Available at: http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.5214-14.2015

Carrillo-Reid, L., Lopez-Huerta, V. G., Garcia-Munoz, M., Theiss, S., & Arbuthnott, G. W. (2015b). Cell assembly signatures defined by short-term synaptic plasticity in cortical networks. *International Journal of Neural Systems, 25*, 1550026. Available at: https://www.worldscientific.com/doi/abs/10.1142/S0129065715500264

Catanese, J., & Jaeger, D. (2021). Premotor ramping of thalamic neuronal activity is modulated by nigral inputs and contributes to control the timing of action release. *The Journal of Neuroscience, 41*, 1878–1891. Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.1204-20.2020

Chapin, J. K., Waterhouse, B. D., & Woodward, D. J. (1981). Differences in cutaneous sensory response properties of single somatosensory cortical neurons in awake and halothane anesthetized rats. *Brain Research Bulletin, 6*, 63–70. Available at: https://linkinghub.elsevier.com/retrieve/pii/S036192308180069X

Chen, Z., Zhang, Z.-Y., Zhang, W., Xie, T., Li, Y., Xu, X.-H., & Yao, H. (2021). Direct and indirect pathway neurons in ventrolateral striatum differentially regulate licking movement and nigral responses. *Cell Reports, 37*, 109847. Available at: https://doi.org/10.1016/j.celrep.2021.109847

Cheung, T. H. C., Bezzina, G., Hampson, C. L., Body, S., Fone, K. C. F., Bradshaw, C. M., & Szabadi, E. (2007). Effect of quinpirole on timing behaviour in the free-operant psychophysical procedure: Evidence for the involvement of D2 dopamine receptors. *Psychopharmacology, 193*, 423–436. Available at: https://link.springer.com/10.1007/s00213-007-0798-8

Chiba, A., Oshio, K., & Inase, M. (2015). Neuronal representation of duration discrimination in the monkey striatum. *Physiological Reports, 3*, e12283. Available at: https://doi.wiley.com/10.14814/phy2.12283

Chuhma, N., Tanaka, K. F., Hen, R., & Rayport, S. (2011). Functional connectome of the striatal medium spiny neuron. *The Journal of Neuroscience, 31*, 1183–1192. Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3833-10.2011

Coffey, K. R., Nader, M., & West, M. O. (2016). Single body parts are processed by individual neurons in the mouse dorsolateral striatum. *Brain Research, 1636*, 200–207. Available at: https://doi.org/10.1016/j.brainres.2016.01.031

Coffey, K. R., Nader, M., Bawa, J., & West, M. O. (2017). Homogeneous processing in the striatal direct and indirect pathways: Single body part sensitive type IIb neurons may express either dopamine receptor D1 or D2. *The European Journal of Neuroscience, 46*, 2380–2391.

Cook, J. R., Li, H., Nguyen, B., Huang, H., Mahdavian, P., Kirchgessner, M. A., Strassmann, P., Engelhardt, M., Callaway, E. M., & Jin, X. (2022). Secondary auditory cortex mediates a sensorimotor mechanism for action timing. *Nature Neuroscience, 25*, 330–344. Available at: https://www.nature.com/articles/s41593-022-01025-5

Coslett, H. B., Wiener, M., & Chatterjee, A. (2010). Dissociable neural systems for timing: Evidence from subjects with basal ganglia lesions Tell F, ed. *PLoS One, 5*, e10324. Available at: https://dx.plos.org/10.1371/journal.pone.0010324

Cruz, B. F., Guiomar, G., Soares, S., Motiwala, A., Machens, C. K., & Paton, J. J. (2022). Action suppression reveals opponent parallel control via striatal circuits. *Nature, 607*, 521–526. Available at: https://www.nature.com/articles/s41586-022-04894-9

Cury, K. M., & Uchida, N. (2010). Robust odor coding via inhalation-coupled transient activity in the mammalian olfactory bulb. *Neuron, 68*, 570–585. Available at: https://doi.org/10.1016/j.neuron.2010.09.040

De Corte, B. J., Wagner, L. M., Matell, M. S., & Narayanan, N. S. (2019). Striatal dopamine and the temporal control of behavior. *Behavioural Brain Research, 356*, 375–379. Available at: https://doi.org/10.1016/j.bbr.2018.08.030

DeLong, M. R. (1971). Of pallidal during movement. *The Journal of Physiology, 34*, 414–427.

Deschênes, M., Takatoh, J., Kurnikova, A., Moore, J. D., Demers, M., Elbaz, M., Furuta, T., Wang, F., & Kleinfeld, D. (2016). Inhibition, not excitation, drives rhythmic whisking. *Neuron, 90*, 374–387. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0896627316001860

Díaz-Hernández, E., Contreras-López, R., Sánchez-Fuentes, A., Rodríguez-Sibrían, L., Ramírez-Jarquín, J. O., & Tecuapetla, F. (2018). The thalamostriatal projections contribute to the initiation and execution of a sequence of movements. *Neuron, 100*, 739–752.e5. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0896627318308547

Dray, A., Gonye, T. J., & Oakley, N. R. (1976). Caudate stimulation and substantia nigra activity in the rat. *The Journal of Physiology, 259*, 825–849. Available at: http://doi.wiley.com/10.1113/jphysiol.1976.sp011497

Drew, M. R., Fairhurst, S., Malapani, C., Horvitz, J. C., & Balsam, P. D. (2003). Effects of dopamine antagonists on the timing of two intervals. *Pharmacology, Biochemistry, and Behavior, 75*, 9–15. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0091305703000364

Dudman, J. T., & Krakauer, J. W. (2016). The basal ganglia: From motor commands to the control of vigor. *Current Opinion in Neurobiology, 37*, 158–166. Available at: https://doi.org/10.1016/j.conb.2016.02.005

Erro, M. E., Lanciego, J. L., Arribas, J., & Gimenez-Amaya, J. M. (2001). Striatal input from the ventrobasal complex of the rat thalamus. *Histochemistry and Cell Biology, 115*, 447–454.

Erro, M. E., Lanciego, J. L., & Giménez-Amaya, J. M. (2002). Re-examination of the thalamostriatal projections in the rat with retrograde tracers. *Neuroscience Research, 42*, 45–55. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0168010201003029

Frederick, D. (1996). Effects of selective dopamine D1- and D2-agonists and antagonists on timing performance in rats. *Pharmacology, Biochemistry, and Behavior, 53*, 759–764. Available at: https://linkinghub.elsevier.com/retrieve/pii/0091305795021035

Gaidica, M., Hurst, A., Cyr, C., & Leventhal, D. K. (2018). Distinct populations of motor thalamic neurons encode action initiation, action selection, and movement vigor. *The Journal of Neuroscience, 38*, 6563–6573. Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.0463-18.2018

Gámez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping Zatorre R, ed. *PLoS Biology, 17*, e3000054. Available at: https://dx.plos.org/10.1371/journal.pbio.3000054

Gerfen, C. R., Engber, T. M., Mahan, L. C., Susel, Z., Chase, T. N., Monsma, F. J., & Sibley, D. R. (1990). D1 and D2 dopamine receptor-regulated gene expression of striatonigral and striatopallidal neurons. *Science (80-), 250*, 1429–1432. Available at: https://www.science.org/doi/10.1126/science.2147780

Goel, A., & Buonomano, D. V. (2014). Timing as an intrinsic property of neural networks: Evidence from in vivo and in vitro experiments. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 369*, 20120460. Available at: https://royalsocietypublishing.org/doi/10.1098/rstb.2012.0460

Goel, A., & Buonomano, D. V. (2016). Temporal interval learning in cortical cultures is encoded in intrinsic network dynamics. *Neuron, 91*, 320–327. Available at: https://doi.org/10.1016/j.neuron.2016.05.042

Gouvêa, T. S., Monteiro, T., Motiwala, A., Soares, S., Machens, C., & Paton, J. J. (2015). Striatal dynamics explain duration judgments. *eLife, 4*, 1–14. Available at: https://elifesciences.org/articles/11386

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience, 31*, 359–387. Available at: http://www.annualreviews.org/doi/10.1146/annurev.neuro.29.051605.112851

Guo, Z. V., Inagaki, H. K., Daie, K., Druckmann, S., Gerfen, C. R., & Svoboda, K. (2017). Maintenance of persistent activity in a frontal thalamocortical loop. *Nature, 545*, 181–186. Available at: http://www.nature.com/articles/nature22324

Hegeman, D. J., Hong, E. S., Hernández, V. M., & Chan, C. S. (2016). The external globus pallidus: Progress and perspectives Bolam P, ed. *The European Journal of Neuroscience, 43*, 1239–1265. Available at: https://onlinelibrary.wiley.com/doi/10.1111/ejn.13196

Henke, J., Bunk, D., von Werder, D., Häusler, S., Flanagin, V. L., & Thurley, K. (2021). Distributed coding of duration in rodent prefrontal cortex during time reproduction. *eLife, 10*, 1–24. Available at: https://elifesciences.org/articles/71612

Hidalgo-Balbuena, A. E., Luma, A. Y., Pimentel-Farfan, A. K., Peña-Rangel, T., & Rueda-Orozco, P. E. (2019). Sensory representations in the striatum provide a temporal reference for learning and executing motor habits. *Nature Communications, 10*, 4074. Available at: https://doi.org/10.1038/s41467-019-12075-y

Hikosaka, O. (2007). GABAergic output of the basal ganglia. *Progress in Brain Research*, 209–226. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0079612306600125

Hikosaka, O., & Wurtz, R. H. (1983). Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent visual and saccade responses. *Journal of Neurophysiology, 49*, 1268–1284. Available at: http://www.physiology.org/doi/10.1152/jn.1983.49.5.1268

Hintiryan, H., Foster, N. N., Bowman, I., Bay, M., Song, M. Y., Gou, L., Yamashita, S., Bienkowski, M. S., Zingg, B., Zhu, M., Yang, X. W., Shih, J. C., Toga, A. W., & Dong, H. W. (2016). The mouse cortico-striatal projectome. *Nature Neuroscience, 19*, 1100–1114.

Honma, M., Kuroda, T., Futamura, A., Shiromaru, A., & Kawamura, M. (2016). Dysfunctional counting of mental time in Parkinson's disease. *Scientific Reports, 6*, 25421. Available at: http://www.nature.com/articles/srep25421

Honma, M., Masaoka, Y., Koyama, S., Kuroda, T., Futamura, A., Shiromaru, A., Terao, Y., Ono, K., & Kawamura, M. (2018). Impaired cognitive modification for estimating time duration in Parkinson's disease Ginsberg SD, ed. *PLoS One, 13*, e0208956. Available at: https://dx.plos.org/10.1371/journal.pone.0208956

Hoover, J. E., Hoffer, Z. S., & Alloway, K. D. (2003). Projections from primary somatosensory cortex to the neostriatum: The role of somatotopic continuity in corticostriatal convergence. *Journal of Neurophysiology, 89*, 1576–1587. Available at: http://jn.physiology.org/cgi/doi/10.1152/jn.01009.2002

Inagaki, H. K., Chen, S., Ridder, M. C., Sah, P., Li, N., Yang, Z., Hasanbegovic, H., Gao, Z., Gerfen, C. R., & Svoboda, K. (2022). A midbrain-thalamus-cortex circuit reorganizes cortical dynamics to initiate movement. *Cell, 185*, 1065–1081.e23. Available at: https://doi.org/10.1016/j.cell.2022.02.006

Iwamuro, H., Tachibana, Y., Ugawa, Y., Saito, N., & Nambu, A. (2017). Information processing from the motor cortices to the subthalamic nucleus and globus pallidus and their somatotopic organizations revealed electrophysiologically in monkeys. *The European Journal of Neuroscience, 46*, 2684–2701. Available at: https://onlinelibrary.wiley.com/doi/10.1111/ejn.13738

Jaeger, D., & Kita, H. (2011). Functional connectivity and integrative properties of globus pallidus neurons. *Neuroscience, 198*, 44–53. Available at: https://doi.org/10.1016/j.neuroscience.2011.07.050

Jin, D. Z., Fujii, N., & Graybiel, A. M. (2009). Neural representation of time in cortico-basal ganglia circuits. *Proceedings of the National Academy of Sciences, 106*, 19156–19161. Available at: https://pnas.org/doi/full/10.1073/pnas.0909881106

Johansson, Y., & Silberberg, G. (2020). The functional organization of cortical and thalamic inputs onto five types of striatal neurons is determined by source and target cell identities. *Cell Reports, 30*, 1178–1194.e3. Available at: https://doi.org/10.1016/j.celrep.2019.12.095

Junek, S., Kludt, E., Wolf, F., & Schild, D. (2010). Olfactory coding with patterns of response latencies. *Neuron, 67*, 872–884. Available at: https://doi.org/10.1016/j.neuron.2010.08.005

Jurado-Parras, M.-T., Safaie, M., Sarno, S., Louis, J., Karoutchi, C., Berret, B., & Robbe, D. (2020). The dorsal striatum energizes motor routines. *Current Biology, 30*, 4362–4372.e6. Available at: https://doi.org/10.1016/j.cub.2020.08.049

Kamada, T., & Hata, T. (2021). Striatal dopamine D1 receptors control motivation to respond, but not interval timing, during the timing task. *Learning & Memory, 28*, 24–29. Available at: http://learnmem.cshlp.org/lookup/doi/10.1101/lm.052266.120

Karalis, N., & Sirota, A. (2022). Breathing coordinates cortico-hippocampal dynamics in mice during offline states. *Nature Communications, 13*, 467. Available at: https://www.nature.com/articles/s41467-022-28090-5

Ketzef, M., Spigolon, G., Johansson, Y., Bonito-Oliva, A., Fisone, G., & Silberberg, G. (2017). Dopamine depletion impairs bilateral sensory processing in the striatum in a pathway-dependent manner. *Neuron, 94*, 855–865.e5. Available at: https://doi.org/10.1016/j.neuron.2017.05.004

Kim, J., Ghim, J.-W., Lee, J. H., & Jung, M. W. (2013). Neural correlates of interval timing in rodent prefrontal cortex. *The Journal of Neuroscience, 33*, 13834–13847. Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.1443-13.2013

Kita, H., & Jaeger, D. (2016). Organization of the globus pallidus. In *Handbook of behavioral neuroscience* (pp. 259–276). Available at: https://linkinghub.elsevier.com/retrieve/pii/B9780128022061000131

Kravitz, A. V., Freeze, B. S., Parker, P. R. L., Kay, K., Thwin, M. T., Deisseroth, K., & Kreitzer, A. C. (2010). Regulation of parkinsonian motor behaviours by optogenetic control of basal ganglia circuitry. *Nature, 466*, 622–626. Available at: http://www.nature.com/articles/nature09159

Kunimatsu, J., Suzuki, T. W., Ohmae, S., & Tanaka, M. (2018). Different contributions of preparatory activity in the basal ganglia and cerebellum for self-timing. *eLife, 7*, 1–19. Available at: https://elifesciences.org/articles/35676

Lee, C. R., Yonk, A. J., Wiskerke, J., Paradiso, K. G., Tepper, J. M., & Margolis, D. J. (2019). Opposing influence of sensory and motor cortical input on striatal circuitry and choice behavior. *Current Biology, 29*, 1313–1323.e5. Available at: https://doi.org/10.1016/j.cub.2019.03.028

Luczak, A., Barthó, P., Marguet, S. L., Buzsáki, G., & Harris, K. D. (2007). Sequential structure of neocortical spontaneous activity in vivo. *Proceedings of the National Academy of Sciences, 104*, 347–352. Available at: https://pnas.org/doi/full/10.1073/pnas.0605643104

Luczak, A., Barthó, P., & Harris, K. D. (2009). Spontaneous events outline the realm of possible sensory responses in neocortical populations. *Neuron, 62*, 413–425. Available at: https://doi.org/10.1016/j.neuron.2009.03.014

Luczak, A., McNaughton, B. L., & Harris, K. D. (2015). Packet-based communication in the cortex. *Nature Reviews. Neuroscience, 16*, 745–755. Available at: https://doi.org/10.1038/nrn4026

Matell, M. S., Meck, W. H., & Nicolelis, M. A. L. (2003). Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behavioral Neuroscience, 117*, 760–773. Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/0735-7044.117.4.760

Matsuyama, K., & Tanaka, M. (2021). Temporal prediction signals for periodic sensory events in the primate central thalamus. *The Journal of Neuroscience, 41*, 1917–1927. Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.2151-20.2021

Meck, W. H. (2006). Neuroanatomical localization of an internal clock: A functional link between mesolimbic, nigrostriatal, and mesocortical dopaminergic systems. *Brain Research, 1109*, 93–107. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0006899306017203

Mello, G. B. M., Soares, S., & Paton, J. J. (2015). A scalable population code for time in the striatum. *Current Biology, 25*, 1113–1122. Available at: https://doi.org/10.1016/j.cub.2015.02.036

Merchant, H., Zarco, W., Bartolo, R., Perez, O., & Prado, L. (2011). Measuring time with different neural chronometers during a synchronization-continuation task. *Proceedings of the National Academy of Sciences, 108*, 19784–19789.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience, 36*, 313–336. Available at: https://www.annualreviews.org/doi/10.1146/annurev-neuro-062012-170349

Mochol, G., Hermoso-Mendizabal, A., Sakata, S., Harris, K. D., & de la Rocha, J. (2015). Stochastic transitions into silence cause noise correlations in cortical circuits. *Proceedings of the National Academy of Sciences, 112*, 3529–3534. Available at: http://www.pnas.org/lookup/doi/10.1073/pnas.1410509112

Mole, J., Winegardner, J., Malley, D., & Fish, J. (2018). Time perception impairment following thalamic stroke: A case study. *Neuropsychological Rehabilitation, 28*, 208–222. Available at: https://doi.org/10.1080/09602011.2017.1383273

Monteiro, T., Rodrigues, F. S., Pexirra, M., Cruz, B. F., Rueda-Orozco, P. E., & Paton, J. J. (2021). Using temperature to analyse the neural basis of a latent temporal decision. *bioRxiv*. Available at: https://doi.org/10.1101/2020.08.24.251827

Monteiro, T., Rodrigues, F. S., Pexirra, M., Cruz, B. F., Gonçalves, A. I., Rueda-Orozco, P. E., & Paton, J. J. (2023). Using temperature to analyze the neural basis of a time-based decision. *Nature Neuroscience*. Available at: https://doi.org/10.1038/s41593-023-01378-5

Nagy, A., Paróczy, Z., Norita, M., & Benedek, G. (2005). Multisensory responses and receptive field properties of neurons in the substantia nigra and in the caudate nucleus. *The European Journal of Neuroscience, 22*, 419–424. Available at: http://doi.wiley.com/10.1111/j.1460-9568.2005.04211.x

Nambu, A. (2011). Somatotopic organization of the primate basal ganglia. *Frontiers in Neuroanatomy, 5*, 1–9. Available at: http://journal.frontiersin.org/article/10.3389/fnana.2011.00026/abstract

Nozaradan, S., Schwartze, M., Obermeier, C., & Kotz, S. A. (2017). Specific contributions of basal ganglia and cerebellum to the neural tracking of rhythm. *Cortex, 95*, 156–168. Available at: https://doi.org/10.1016/j.cortex.2017.08.015

Pare, D., Steriade, M., Deschenes, M., & Oakson, G. (1987). Physiological characteristics of anterior thalamic nuclei, a group devoid of inputs from reticular thalamic nucleus. *Journal of Neurophysiology, 57*, 1669–1685. Available at: https://www.physiology.org/doi/10.1152/jn.1987.57.6.1669

Parent, A., Mackey, A., Smith, Y., & Boucher, R. (1983). The output organization of the substantia nigra in primate as revealed by a retrograde double labeling method. *Brain Research Bulletin, 10*, 529–537. Available at: https://linkinghub.elsevier.com/retrieve/pii/036192308390151X

Parent, A., Bouchard, C., & Smith, Y. (1984). The striatopallidal and striatonigral projections: Two distinct fiber systems in primate. *Brain Research, 303*, 385–390. Available at: https://linkinghub.elsevier.com/retrieve/pii/0006899384912241

Paton, J. J., & Lau, B. (2015). Tread softly and carry a clock's tick. *Nature Neuroscience, 18*, 329–330. Available at: https://doi.org/10.1038/nn.3959

Peña-Rangel, T. M., Lugo-Picos, P. I., Báez-Cordero, A. S., Hidalgo-Balbuena, A. E., Luma, A. Y., Pimentel-Farfan, A. K., & Rueda-Orozco, P. E. (2021). Altered sensory representations in parkinsonian cortical and basal ganglia networks. *Neuroscience, 466*, 10–25. Available at: http://www.ncbi.nlm.nih.gov/pubmed/33965505

Peters, A. J., Fabre, J. M. J., Steinmetz, N. A., Harris, K. D., & Carandini, M. (2021). Striatal activity topographically reflects cortical activity. *Nature, 591*, 420–425. Available at: https://doi.org/10.1038/s41586-020-03166-8

Pidoux, M., Mahon, S., Deniau, J. M., & Charpier, S. (2011). Integration and propagation of somatosensory responses in the corticostriatal pathway: An intracellular study in vivo. *The Journal of Physiology, 589*, 263–281.

Pimentel-Farfan, A. K., Báez-Cordero, A. S., Peña-Rangel, T. M., & Rueda-Orozco, P. E. (2022). Cortico-striatal circuits for bilaterally coordinated movements. *Science Advances, 8*, 1–19. Available at: https://www.science.org/doi/10.1126/sciadv.abk2241

Rao, S. M., Harrington, D. L., Haaland, K. Y., Bobholz, J. A., Cox, R. W., & Binder, J. R. (1997). Distributed neural systems underlying the timing of movements. *The Journal of Neuroscience, 17*, 5528–5535. Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.17-14-05528.1997

Redgrave, P., Rodriguez, M., Smith, Y., Rodriguez-Oroz, M. C., Lehericy, S., Bergman, H., Agid, Y., Delong, M. R., & Obeso, J. A. (2010). Goal-directed and habitual control in the basal ganglia: Implications for Parkinson's disease. *Nature Reviews. Neuroscience, 11*, 760–772. Available at: https://doi.org/10.1038/nrn2915

Reig, R., & Silberberg, G. (2014). Multisensory integration in the mouse striatum. *Neuron, 83*, 1200–1212. Available at: https://doi.org/10.1016/j.neuron.2014.07.033

Reig, R., & Silberberg, G. (2016). Distinct corticostriatal and intracortical pathways mediate bilateral sensory responses in the striatum. *Cerebral Cortex, 26*, 4405–4415. Available at: https://academic.oup.com/cercor/article-lookup/doi/10.1093/cercor/bhw268

Reiner, A., Jiao, Y., Del Mar, N., Laverghetta, A. V., & Lei, W. L. (2003). Differential morphology of pyramidal tract-type and intratelencephalically projecting-type corticostriatal neurons and their intrastriatal terminals in rats. *The Journal of Comparative Neurology, 457*, 420–440.

Robbe, D. (2018). To move or to sense? Incorporating somatosensory representation into striatal functions. *Current Opinion in Neurobiology, 52*, 123–130. Available at: https://doi.org/10.1016/j.conb.2018.04.009

Rowe, K. C., Paulsen, J. S., Langbehn, D. R., Duff, K., Beglinger, L. J., Wang, C., O'Rourke, J. J. F., Stout, J. C., & Moser, D. J. (2010). Self-paced timing detects and tracks change in prodromal Huntington disease. *Neuropsychology, 24*, 435–442. Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/a0018905

Rueda-Orozco, P. E., & Robbe, D. (2015). The striatum multiplexes contextual and kinematic information to constrain motor habits execution. *Nature Neuroscience, 18*, 435–460.

Schmidt, R., Leventhal, D. K., Mallet, N., Chen, F., & Berke, J. D. (2013). Canceling actions involves a race between basal ganglia pathways. *Nature Neuroscience, 16*, 1118–1124. Available at: http://www.nature.com/articles/nn.3456

Schultz, W. (1986). Activity of pars reticulata neurons of monkey substantia nigra in relation to motor, sensory, and complex events. *Journal of Neurophysiology, 55*, 660–677. Available at: https://www.physiology.org/doi/10.1152/jn.1986.55.4.660

Schwartze, M., Stockert, A., & Kotz, S. A. (2015). Striatal contributions to sensory timing: Voxel-based lesion mapping of electrophysiological markers. *Cortex, 71*, 332–340. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0010945215002610

Sheng, M., Lu, D., Shen, Z., & Poo, M. (2019). Emergence of stable striatal D1R and D2R neuronal ensembles with distinct firing sequence during motor learning. *Proceedings of the National Academy of Sciences, 116*, 11038–11047. Available at: https://pnas.org/doi/full/10.1073/pnas.1901712116

Shi, L. H., Luo, F., Woodward, D. J., & Chang, J. Y. (2004). Neural responses in multiple basal ganglia regions during spontaneous and treadmill locomotion tasks in rats. *Experimental Brain Research, 157*, 303–314. Available at: http://link.springer.com/10.1007/s00221-004-1844-y

Shi, L. H., Luo, F., Woodward, D. J., & Chang, J. Y. (2005). Dose and behavioral context dependent inhibition of movement and basal ganglia neural activity by delta-9-tetrahydrocannabinol during spontaneous and treadmill locomotion tasks in rats. *Synapse, 55*, 1–16. Available at: http://doi.wiley.com/10.1002/syn.20088

Sippy, T., Lapray, D., Crochet, S., & Petersen, C. C. H. (2015). Cell-type-specific sensorimotor processing in striatal projection neurons during goal-directed behavior. *Neuron, 88*, 298–305. Available at: https://doi.org/10.1016/j.neuron.2015.08.039

Smith, J. B., Mowery, T. M., & Alloway, K. D. (2012). Thalamic POm projections to the dorsolateral striatum of rats: Potential pathway for mediating stimulus-response associations for sensorimotor habits. *Journal of Neurophysiology, 108*, 160–174. Available at: http://jn.physiology.org/cgi/doi/10.1152/jn.00142.2012

Smith, Y., Galvan, A., Ellender, T. J., Doig, N., Villalba, R. M., Huerta-Ocampo, I., Wichmann, T., & Bolam, J. P. (2014). The thalamostriatal system in normal and diseased states. *Frontiers in Systems Neuroscience, 8*, 1–18. Available at: http://journal.frontiersin.org/article/10.3389/fnsys.2014.00005/abstract

Soares, S., Atallah, B. V., & Paton, J. J. (2016). Midbrain dopamine neurons control judgment of time. *Science (80-), 354*, 1273–1277. Available at: https://www.science.org/doi/10.1126/science.aah5234

Sommer, M. A. (2003). The role of the thalamus in motor control. *Current Opinion in Neurobiology, 13*, 663–670. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0959438803001697

Sreenivasan, V., & Petersen, C. C. H. (2016). Inhibition patterns the whisking rhythm. *Neuron, 90*, 211–213. Available at: https://doi.org/10.1016/j.neuron.2016.04.012

Stevens, M. C., Kiehl, K. A., Pearlson, G., & Calhoun, V. D. (2007). Functional neural circuits for mental timekeeping. *Human Brain Mapping, 28*, 394–408. Available at: https://onlinelibrary.wiley.com/doi/10.1002/hbm.20285

Tanaka, M. (2006). Inactivation of the central thalamus delays self-timed saccades. *Nature Neuroscience, 9*, 20–22. Available at: http://www.nature.com/articles/nn1617

Tanaka, M. (2007). Cognitive signals in the primate motor thalamus predict saccade timing. *The Journal of Neuroscience, 27*, 12109–12118. Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.1873-07.2007

Taverna, S., Ilijic, E., & Surmeier, D. J. (2008). Recurrent collateral connections of striatal medium spiny neurons are disrupted in models of Parkinson's disease. *The Journal of Neuroscience, 28*, 5504–5512. Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.5493-07.2008

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience, 31*, 3805–3812. Available at: https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.5561-10.2011

Toda, K., Lusk, N. A., Watson, G. D. R., Kim, N., Lu, D., Li, H. E., Meck, W. H., & Yin, H. H. (2017). Nigrotectal stimulation stops interval timing in mice. *Current Biology, 27*, 3763–3770.e3. Available at: https://doi.org/10.1016/j.cub.2017.11.003

Toso, A., Reinartz, S., Pulecchi, F., & Diamond, M. E. (2021). Time coding in rat dorsolateral striatum. *Neuron, 109*, 3663–3673.e6. Available at: https://doi.org/10.1016/j.neuron.2021.08.020

Turner, R. S., & Desmurget, M. (2010). Basal ganglia contributions to motor control: A vigorous tutor. *Current Opinion in Neurobiology, 20*, 704–716. Available at: https://doi.org/10.1016/j.conb.2010.08.022

Wang, J., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible timing by temporal scaling of cortical responses. *Nature Neuroscience, 21*, 102–110. Available at: https://doi.org/10.1038/s41593-017-0028-6

Weder, B. J., Leenders, K. L., Vontobel, P., Nienhusmeier, M., Keel, A., Zaunbauer, W., Vonesch, T., & Ludin, H.-P. (1999). Impaired somatosensory discrimination of shape in Parkinson's disease: Association with caudate nucleus dopaminergic function. *Human*

Brain Mapping, 8, 1–12. Available at: https://onlinelibrary.wiley.com/doi/10.1002/(SICI)1097-0193(1999)8:1%3C1::AID-HBM1%3E3.0.CO;2-E

Weder, B., Azari, N. P., Knorr, U., Seitz, R. J., Keel, A., Nienhusmeier, M., Maguire, R. P., Leenders, K. L., & Ludin, H.-P. (2000). Disturbed functional brain interactions underlying deficient tactile object discrimination in Parkinson's disease. *Human Brain Mapping, 11*, 131–145. Available at: https://onlinelibrary.wiley.com/doi/10.1002/1097-0193(200011)11:3%3C131::AID-HBM10%3E3.0.CO;2-M

West, M. O. (1998). Anesthetics eliminate somatosensory-evoked discharges of neurons in the somatotopically organized sensorimotor striatum of the rat. *The Journal of Neuroscience, 18*, 9055–9068.

West, M. O., Carelli, R. M., Pomerantz, M., Cohen, S. M., Gardner, J. P., Chapin, J. K., & Woodward, D. J. (1990). A region in the dorsolateral striatum of the rat exhibiting single-unit correlations with specific locomotor limb movements. *Journal of Neurophysiology, 64*, 1233–1246. Available at: http://www.physiology.org/doi/10.1152/jn.1990.64.4.1233

Wurtz, R. H., & Hikosaka, O. (1986). Role of the basal ganglia in the initiation of saccadic eye movements. *Progress in Brain Research, 64*, 175–190. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0079612308634123

Xu, M., Zhang, S., Dan, Y., & Poo, M. (2014). Representation of interval timing by temporally scalable firing patterns in rat prefrontal cortex. *Proceedings of the National Academy of Sciences, 111*, 480–485. Available at: https://pnas.org/doi/full/10.1073/pnas.1321314111

Yin, H. H. (2017). The basal ganglia in action. *Neuroscience, 23*, 299–313. Available at: http://journals.sagepub.com/doi/10.1177/1073858416654115

Yoshida, M., & Precht, W. (1971). Monosynaptic inhibition of neurons of the substantia nigra by caudatonigral fibers. *Brain Research, 32*, 225–228. Available at: https://linkinghub.elsevier.com/retrieve/pii/0006899371901703

Zhou, S., Masmanidis, S. C., & Buonomano, D. V. (2020). Neural sequences as an optimal dynamical regime for the readout of time. *Neuron, 108*, 651–658.e5. Available at: https://doi.org/10.1016/j.neuron.2020.08.020

Estimating Time and Rhythm by Predicting External Stimuli

Mildred Salgado-Ménez, Marisol Espinoza-Monroy, Ana M. Malagón, Karla Mercado, and Victor de Lafuente

Abstract

In this chapter, we present recent findings from our group showing that elapsed time, interval timing, and rhythm maintenance might be achieved by the well-known ability of the brain to predict the future states of the world. The difference between predictions and actual sensory evidence is used to generate perceptual and behavioral adjustments that help subjects achieve desired behavioral goals. Concretely, we show that (1) accumulating prediction errors is a plausible strategy humans could use to determine whether a train of consecutive stimuli arrives at regular or irregular intervals. By analyzing the behavior of human and non-human primate subjects performing rhythm perception tasks, we demonstrate that (2) the ability to estimate elapsed time and internally maintain rhythms is shared across primates and humans. Neurophysiological recordings show that (3) the medial premotor cortex engages in rhythm entrainment and maintains oscillatory activity that reveals an internal metronome's spatial and temporal characteristics. Finally, we demonstrate that (4) the amplitude of gamma oscillations within this cortex increases proportionally to the total elapsed time. In conjunction with our most recent experiments, our results suggest that timing might be achieved by an internal simulation of the sensory stimuli and the motor commands that define the timing task that needs to be performed.

Keywords

Model of timing · Decision making · Interval timing rhythm · Time

Introduction

Predicting the occurrence of future events and timing our behavior is a fundamental ability of nervous systems. When events have a constant interval separating them, we can easily predict the time of the next occurrence and plan our behavioral actions accordingly. Playing a musical instrument or dancing to music is a remarkable example of our ability to estimate time and rhythm and closely synchronize our actions to external stimuli. We propose that the ability to predict, more specifically, our ability to simulate the unfolding of events into the future, is key to understanding the cognitive process that we encompass within the *timing* label.

M. Salgado-Ménez · M. Espinoza-Monroy ·
A. M. Malagón · K. Mercado · V. d. Lafuente (✉)
Institute of Neurobiology, National Autonomous University of Mexico, Querétaro, México
e-mail: lafuente@unam.mx

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_9

To accurately predict future events, we have developed several cognitive, sensory, and motor proxies that help us track the passage of time. To estimate elapsed time, for example, we commonly use overt counting mechanisms that keep track of the number of seconds elapsed from a past event (one—Mississippi, two—Mississippi, etc.). Similarly, we can use backward counting to estimate the time-to-go that remains before a relevant sensory or motor event (3, 2, 1, etc.). A fundamental question then arises—what mechanism does the brain implement to generate and maintain such chronometers?

Numerous specialized mechanisms have been proposed that could plausibly account for our ability to keep track of time (Bangert et al., 2011; Becker & Rasmussen, 2007; Grondin, 2001; Ivry & Spencer, 2004; Matell & Meck, 2004; Simen et al., 2011). A pacemaker-accumulator was among the first of such specialized mechanisms. Upon starting such an internal chronometer, the regular ticks from a hypothesized pacemaker are accumulated so that the total number of ticks is proportional to elapsed time. A threshold on this accumulated quantity could trigger an action at a particular time.

Mechanisms that use hypothetical oscillators have also been proposed in which cortical neurons could generate regular volleys of activity directed toward the neurons of the basal ganglia. If cortical neurons span a wide range of intervals (oscillators of different frequencies), basal ganglia neurons could be trained to decode time by detecting peak activities among combinations of such oscillators (Allman et al., 2014; Matell et al., 2007; Matell & Meck, 2000; Matell & Meck, 2004; McAuley, 1995; Teki et al., 2011; Treisman et al., 1990; Wiener & Kanai, 2016).

More recently, the study of the dynamical systems that arise from the activity of recurrently connected neurons has been used to demonstrate that time can be decoded from the ever-changing neuronal states defined by the numerous neurons' joint activity (Remington et al., 2018a). Importantly, this new approach to timing eliminates the need to use explicit clocks or regularly firing neurons. It has demonstrated that any pattern of activity that changes with time can be used to decode time as long as the activity of the population changes reproducibly across resets of the system.

A dynamical systems approach that is grounded in electrophysiological data recorded from behaving monkeys has refined this *timing-through-dynamics* approach to show that the combined activity of neurons in premotor cortices generates low-dimensional manifolds (surfaces) that can change their geometry and speed of change to measure and reproduce different time intervals reliably (Remington et al., 2018b).

In parallel with this theoretically inclined work, more empirical studies have consistently demonstrated that a wide array of brain structures are engaged by tasks that require estimating time and generating time-constrained behavioral actions. Motor and premotor areas of the cortex in humans and primates show ramping or cycling patterns of neuronal activity that are closely related to the ability of subjects to measure time and generate behavioral decisions based on time and rhythm estimation (Merchant et al., 2011, 2013, 2015; Merchant & Averbeck, 2017).

Sub-cortical areas such as the basal ganglia, the cerebellum, and the hippocampus contain neurons that increase their activity and reach precisely timed peaks that indicate when a particular period of time has elapsed. They might also indicate that the time to execute an action has arrived (Breska & Ivry, 2018; Coull et al., 2011; Ferrandez et al., 2003; Gibbon et al., 1997; Grahn & Brett, 2007; Jin et al., 2009; Schubotz & Von Cramon, 2004).

Notably, the fundamental studies discussed so far have in common the view that there is a dedicated mechanism that the brain uses to estimate time. This mechanism might be distributed among different brain areas and structures and can plausibly be implemented by the variety of algorithms just described. It is crucial to note that these studies propose a group of algorithms and processes with the specialized objective of implementing the timing abilities of the nervous systems (Finnerty et al., 2015; Karmarkar & Buonomano, 2007; Mauk & Buonomano, 2004; Paton & Buonomano, 2018).

In this chapter, we describe recent findings from our group that were aimed at characterizing the ability of human and non-human primates to estimate and maintain rhythms of different frequencies. We also show electrophysiological recordings from the medial premotor cortex (MPC; also called the supplementary motor area, SMA), demonstrating that the motor system might be using simulated movements as a proxy to estimate and maintain metronomes with different tempos.

Finally, we briefly discuss recent findings from our laboratory that strongly support the hypothesis that the brain uses its well-known ability to simulate past, present, and future states internally. We propose that this capacity to simulate possible states of the world is used to keep track of time, predict the onset of future events, and plan behavioral actions accordingly.

Results

1. *The accumulation of prediction errors can be used to distinguish regular from irregular trains of sensory stimuli*

When a train of sensory or motor events occurs at regular intervals, the brain can easily predict the onset of future occurrences of these events (Zacks & Tversky, 2001). This capacity allows us to dance, appreciate, and generate music, and, in general, allows us to plan and execute our behavior in close timing to predictable future events. However, how do we decide whether a train of sensory events occurs at regular intervals?

Suppose we approach timing as the ability to internally model current and future states of the world. In that case, a mechanism that utilizes the differences between predictions and actual events might be useful to distinguish regular from irregular rhythms (Cooper, 2021; Schulze, 1978, 1989). In one of our recent investigations (Espinoza-Monroy & de Lafuente, 2021), we made use of a regularity discrimination task and a model that accumulated the time difference between the predicted onset and the actual onset of sensory stimuli.

In our regularity discrimination task, human participants perceived a train of consecutive sensory stimuli and had to decide whether they occurred at regular or irregular time intervals (Figs. 1 and 2). As expected, subjects rapidly and accurately communicated their "irregular" decisions when the intervals between stimuli were highly variable. On the contrary, when the variability of the intervals between consecutive stimuli was highly regular, subjects took more time. They generated a significant proportion of errors (false alarms), thus revealing the difficulty of perceiving slight differences between predicted and actual onset times.

Making use of this *regularity detection task*, we compared the resulting psychometric (accuracy; Fig. 3b) and chronometric (response time; Fig. 3c) curves across auditory, visual, and somatosensory modalities (Fig. 3a). Consistent with previous research (Patel et al., 2005; Repp, 2003), we found that human subjects are better at estimating the timing of auditory stimuli. Tactile and visual modalities demonstrated similar accuracy and decision times (Fig. 3d).

Importantly, by making use of the decision-making model, we provided support to the hypothesis that human subjects do not wait to detect large deviations of predicted and actual onset times but instead accumulate the successive differences between predictions and evidence and then commit to a "regular" or an "irregular" decision once the accumulation of many of these differences reaches a decision bound.

Our behavioral and modeling efforts supported the hypothesis that the human brain perceives a train of sensory pulses as irregular when observed onset times deviate from predicted onset times.

2. *The ability to internally maintain rhythms is shared across human and non-human primates*

Ours and previous research firmly established that humans can internally maintain rhythms of different tempos (Hary & Moore, 1987; Mates, 1994; Repp, 2005; Zarco et al., 2009). Are monkeys able to do the same, more so, without overt movements?

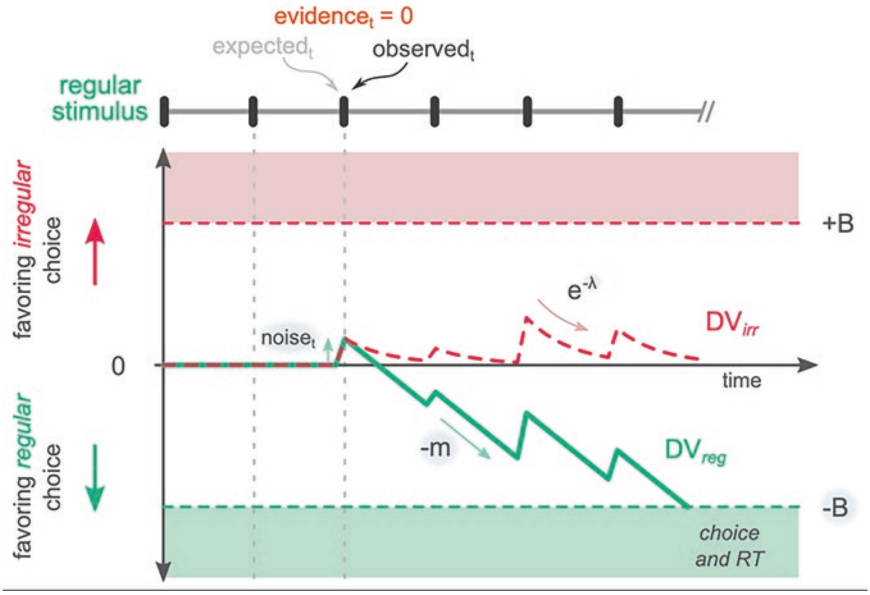


Fig. 1 A simple model to decide whether a train of stimuli comes at regular or irregular intervals. The top traces show an example of sensory pulses presented at regular intervals. When the time of stimulus presentation closely coincides with the brain's expected onset time, then no error signal is generated, and a decision variable moving toward the "regular" choice bound wins the race (DV_{reg}), resulting in a "regular stimulus" decision by the participant. A few model parameters, like the slope and which the "regular" decision variable approaches the lower bound, or an alpha parameter determining how much of the "irregular" evidence takes to decay, allow the model to match not only the resulting accuracy of the decision closely, but also the time it takes the subjects to reach that decision

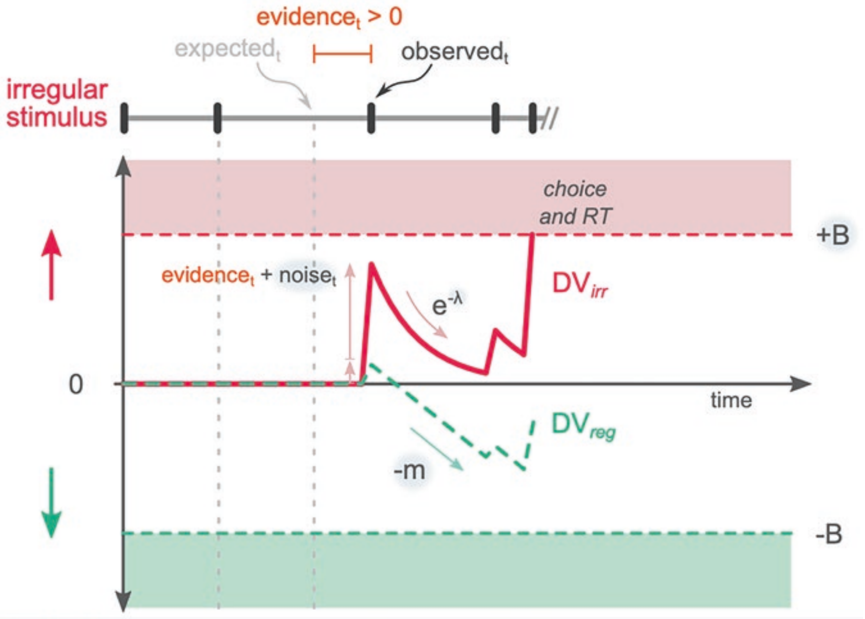


Fig. 2 When sensory stimuli arrive at unexpected times, our model proposes that these time differences accumulate, increasing the amount of "irregular" evidence. When such accumulation (DV_{irr}) reaches the decision bound, the subject makes an "irregular" decision

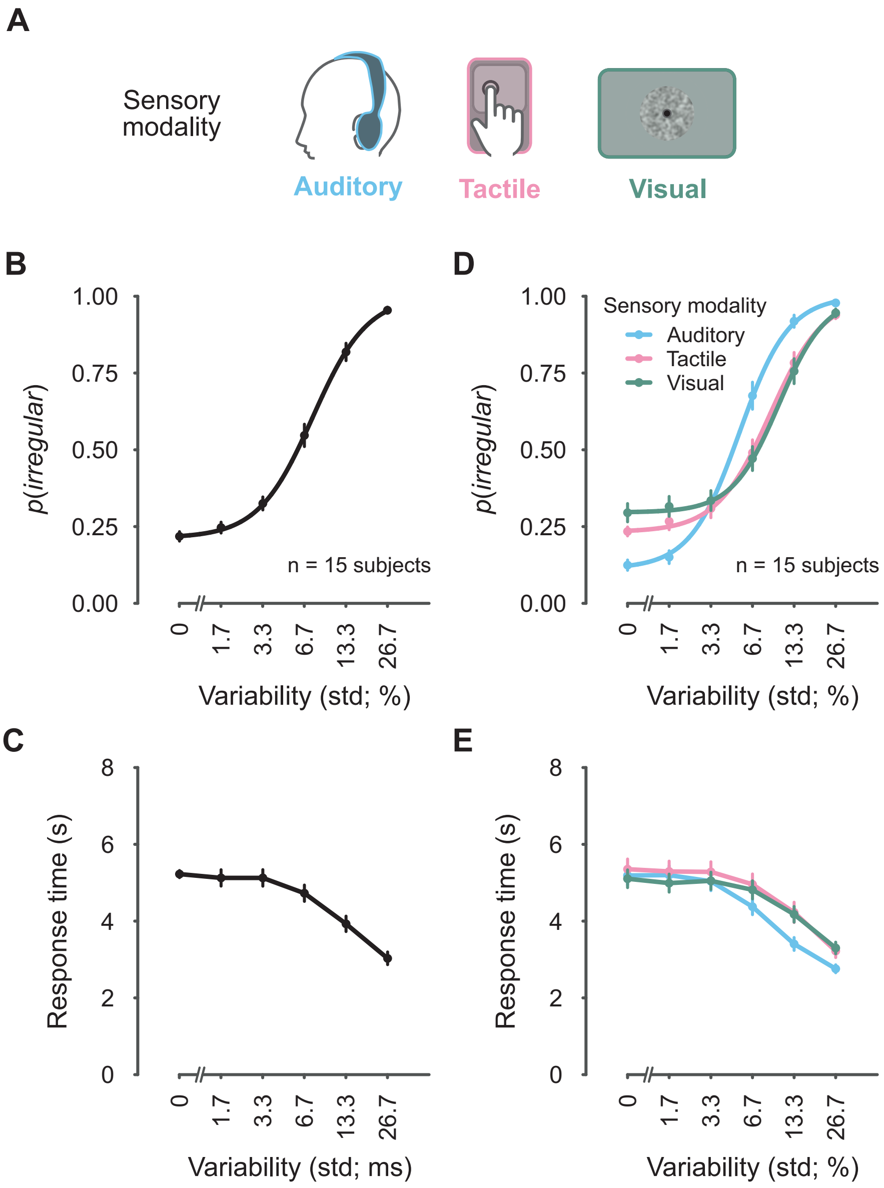


Fig. 3 Behavioral results of a *regular or irregular* decision-making task. (**a**) Brief pulses (50 ms) of auditory, tactile, and visual stimuli were used to generate a train of pulses that determined a rhythm that human subjects had to determine whether they presented at regular or irregular intervals. (**b**) As expected, when sensory pulses arrived at highly irregular intervals (high variability, std. % of interval length), participants accurately identified them as "irregular" p(*irregular*), and this ability decreased as the inter-pulse intervals became more regular (low variability, x-axis). (**c**) As expected, the time subjects took to communicate their "irregular" decisions reduced as the inter-pulse intervals became highly irregular. (**d**) Consistent with previous results, we found that the auditory system was more sensitive and generated faster decisions (**e**) when distinguishing regular from irregular rhythms

In recent work (García-Garibay et al., 2016), we addressed this question by making use of a left-right visual metronome task in which human and non-human primates had to observe and then internally maintain the tempo at which a visual metronome alternated between the left and right sides of a computer screen (Fig. 4). In this task, the subjects observed three intervals of the visually defined metronome, and when the metronome disappeared, they were required to internally estimate its left-right position as a function of elapsed time. Notably, at a random time after the metronome was extinguished, a visual cue instructed the subjects to touch the side of the screen at which they estimated the metronome should be.

As expected, the behavioral results demonstrated that the ability to identify the metronome's position correctly gradually decreased as a function of elapsed time, reflecting the fact that the pace of the internal metronome (no longer visible on the screen) gradually fell out of sync from the correct tempo (Fig. 5). This behavioral pattern was well captured by a model that made use of Weber's law—called the scalar property in the context of timing—stating that the variance of time estimates grows linearly with elapsed time (Fig. 5).

Importantly, by making use of our model, we were able to provide support to the hypothesis that subjects time individual intervals (i.e., they reset their clock after each interval is completed), as opposed to making use of a chronometer that estimates total elapsed time. In addition, our behavioral results demonstrated that subjects establish a mean prior tempo, which causes them to get ahead of trials with slow tempos and fall behind in trials with faster ones.

Overall, our behavioral and modeling work using the visual metronome task demonstrated that monkeys share the human ability to maintain a rhythm internally and, crucially, can do it in the absence of movements.

3. *The medial premotor cortex engages in rhythm entrainment, and it helps maintain oscillatory activity that reveals the timing characteristics of an internal metronome*

Next, we investigated the neuronal correlates of this internal metronome (Cadena-Valencia et al., 2018). What are the neuronal mechanisms allowing human and non-human primates to maintain internal rhythms in the absence of movements? The supplementary motor area (SMA) has been thoroughly studied in the context of timing, motor preparation, and motor execution (Bengtsson et al., 2009; Grahn & Brett, 2007; Grahn & Rowe, 2009; Mita et al., 2009; Rao et al., 1997; Schubotz & von Cramon, 2002). Thus, we decided to investigate what role this area plays in maintaining an internal metronome that does not require motor actions. Of course, we used our recently developed visual metronome task in which monkeys had to maintain an internal rhythm established by a briefly presented metronome (Fig. 6).

Our results demonstrated that circuit activity recorded by local field potentials (LPFs) dis-

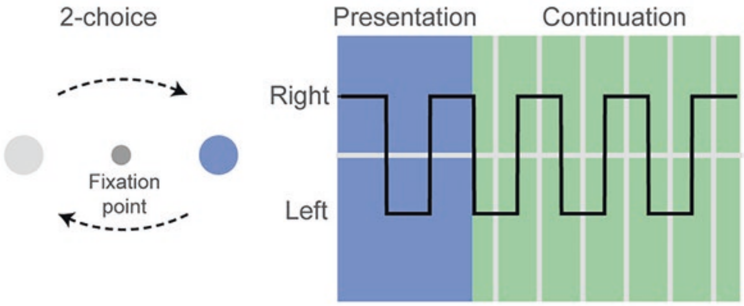


Fig. 4 The visual metronome task. A visual stimulus switches from left to right from an eye fixation point at the center of a computer screen. After three visible stimuli (*presentation* phase), the stimulus disappears, and the participant's task is to keep track of the no longer visible stimulus as a function of time (*continuation* phase)

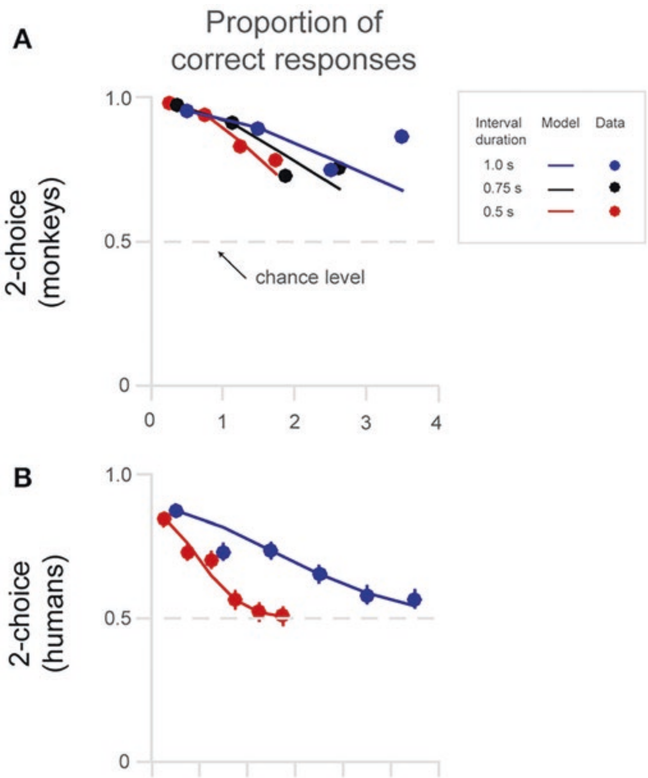


Fig. 5 The ability to correctly identify the position of a no longer visual metronome p(*correct*) decreases as the internal metronome gradually falls out of sync from the true pace of the metronome, for both Rhesus monkeys (**a**) and humans (**b**). This reduction in accuracy is well captured by a model that uses the well-known “scalar property of timing” (colored lines). Please note that, overall, the performance of the Rhesus monkeys is well above human performance in this rhythm task, most likely due to the extensive training that monkeys received

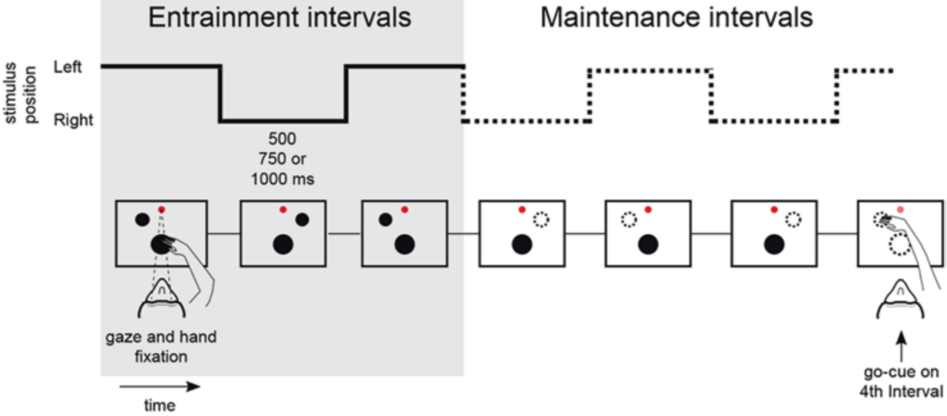


Fig. 6 Detailed illustration of the left-right metronome task. At the center of the screen, eye (red) and hand (black) fixation areas instructed the monkeys to maintain their hands and eyes over the center of the screen throughout the trial. A visual stimulus appeared left or right (randomly determined) and disappeared after three *entrainment* intervals. When the metronome was no longer visible, the *maintenance* intervals could span 1–6 randomly selected intervals. At the go-cue (disappearance of the hand fixation area), monkeys knew they had to communicate their estimated position of the metronome

played rhythmic bursts of high gamma activity (30–40 HZ) that reflected the position and tempo of the internal metronome (Fig. 7a). Remarkably, monkey subjects were able to switch between slow, fast, and rapid tempos, on a trial-by-trial basis. The neuronal activity reflected this slowing

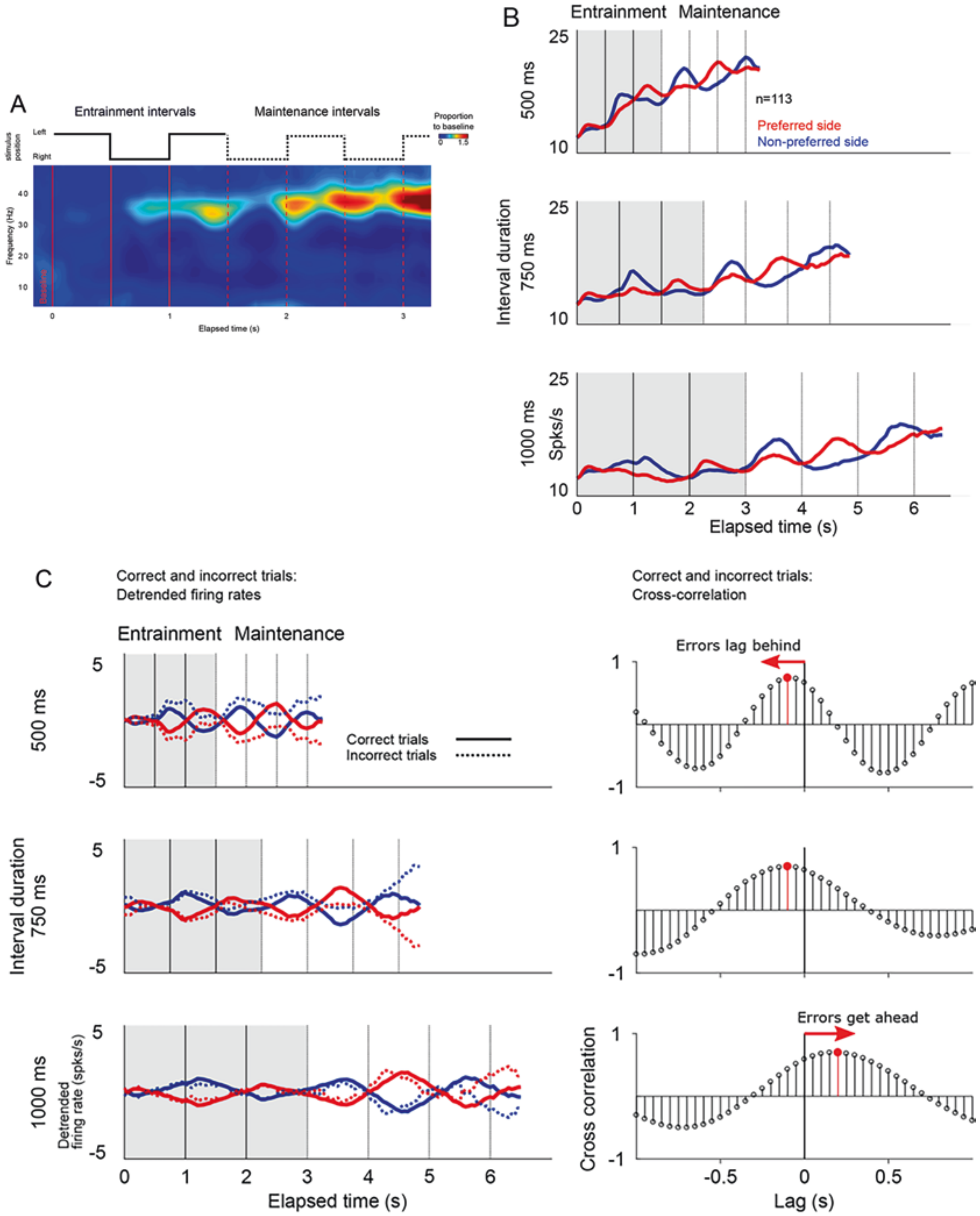


Fig. 7 Single neuron activity and local field potential activity were recorded at the supplementary motor area in monkeys performing the left-right metronome task. (**a**) The high gamma oscillations in LFP (spectrogram) demonstrate clear oscillations during the maintenance epoch of the task, in which the metronome is no longer visible. Please note that these are internal modulations of activity, driven exclusively by the estimates that subjects have of the invisible metronome. (**b**) The mean firing rate activity of supplementary motor neurons oscillates at the metronome's rhythm, and they can stretch or compress this oscillation to match the tempo of the estimated metronome. (**c**) Moreover, by analyzing the dynamics of error trials, compared to correct trials, we were able to determine that monkeys lag behind fast tempos and get ahead of slow ones, importantly determining that monkeys rarely guess the positions of the metronome. Instead of guessing, they tried their best to maintain the correct tempo even on incorrect trials

and stretching of the internal chronometer (Fig. 7b). We decoded the monkey's internal decisions based on single trials of a single LFP, which demonstrated a close relationship between gamma bust activity, firing rate activity, and the pace of the internal metronome. Moreover, the activity from single neurons allowed us to predict whether this internal metronome was falling ahead or behind the true tempo (Fig. 7c).

The fact that a premotor area, such as the supplementary motor area, shows such strong and highly behavioral-predictive activity supports to the theory that the brain uses the motor system as a proxy to estimate the passage of time. Indeed, this interpretation of our results is consistent with the idea that the motor system simulates alternating reaching movements to the left and right as a proxy to maintain an internal rhythm (Fig. 7b) (Egger et al., 2019; Robbe, 2023; Schubotz & von Cramon, 2002; Schubotz & Von Cramon, 2004). Moreover, given that the amplitude of both oscillatory single neuron activity and the amplitude of the gamma bursts gradually increase over time, we propose that the supplementary motor areas not only keep track of individual intervals but, importantly, are able to estimate the total elapsed time spanned since the beginning of the given trial (Fig. 7a, b).

4. *Timing as the ability to simulate future states of the world*

In our most recent investigations (de Lafuente et al., 2024) we aimed to support the idea that the brain does not use a specialized mechanism to estimate rhythm or time. Instead, we tested the hypothesis that the brain internally simulates future motor plans and also stimulates the possible sensory stimuli that might appear in the near future that could be behaviorally relevant.

To this end, we again made use of the visual metronome task. However, we additionally recorded from visual area 4, which is related to encoding visual information and participates in attention-driven processes (Ghose & Maunsell, 2002; Leinweber et al., 2017). Interestingly, we discovered that this visual area not only activates in response to the visual presentation of the metronome but also, crucially, maintains its left-right oscillatory activity once the visual stimulus is no longer visible. We were able to demonstrate that the internal metronome was periodically activating single neurons and also the larger circuit recorded in the LFPs.

In summary, our research has led us to conclude that the brain might not use a specialized mechanism to track time or rhythm. Instead, we propose that the brain utilizes its well-known machinery to simulate the present and future states of the world. Concretely, our experimental findings support the idea that monkeys recreate the visual stimuli and the motor plans that will be needed to correctly identify the position of the metronome once it is no longer visible. In other words, the brain internally recreates the sensory and motor aspects of the metronome once they are no longer visible.

References

Allman, M. J., Teki, S., & Griffiths, T. D. (2014). Properties of the internal clock: First-and second-order principles of subjective time auditory motion-specific mechanisms in the primate brain view project. *Annual Review of Psychology, 65*, 743–771. https://doi.org/10.1146/annurev-psych-010213-115117

Bangert, A., Reuter-Lorenz, P., & Seidler, R. (2011). Dissecting the clock: Understanding the mechanisms of timing across tasks and temporal intervals. *Acta Psychologica, 136*(1), 20–34. https://www.sciencedirect.com/science/article/pii/S0001691810001897

Becker, M. W., & Rasmussen, I. P. (2007). The rhythm aftereffect: Support for time sensitive neurons with broad overlapping tuning curves. *Brain and Cognition, 64*(3), 274–281. https://doi.org/10.1016/J.BANDC.2007.03.009

Bengtsson, S. L., Ullén, F., Henrik Ehrsson, H., Hashimoto, T., Kito, T., Naito, E., Forssberg, H., & Sadato, N. (2009). Listening to rhythms activates motor and premotor cortices. *Cortex, 45*(1), 62–71. https://doi.org/10.1016/J.CORTEX.2008.07.002

Breska, A., & Ivry, R. B. (2018). Double dissociation of single-interval and rhythmic temporal prediction in cerebellar degeneration and Parkinson's disease. *Proceedings of the National Academy of Sciences of the United States of America, 115*(48), 12283–12288. https://doi.org/10.1073/PNAS.1810596115

Cadena-Valencia, J., Garcia-Garibay, O., Merchant, H., Jazayeri, M., & DeLaFuente, V. (2018). Entrainment and maintenance of an internal metronome in supplementary motor area. *eLife, 7*, e38983.

Cooper, R. P. (2021). Action production and event perception as routine sequential behaviors. *Topics in Cognitive Science, 13*(1), 63–78. https://doi.org/10.1111/TOPS.12462

Coull, J., Cheng, R., & Meck, W. (2011). Neuroanatomical and neurochemical substrates of timing. *Neuropsychopharmacology, 36*(1), 3–25. https://www.nature.com/articles/npp2010113

de Lafuente, V., Jazayeri, M., Merchant, H., García-Garibay, O., Cadena-Valencia, J., & Malagón, A. M. (2024). Keeping time and rhythm by internal simulation of sensory stimuli and behavioral actions. *Science Advances, 10*(2). https://doi.org/10.1126/SCIADV.ADH8185/SUPPL_FILE/SCIADV.ADH8185_SM.PDF

Egger, S. W., Remington, E. D., Chang, C. J., & Jazayeri, M. (2019). Internal models of sensorimotor integration regulate cortical dynamics. *Nature Neuroscience, 22*(11), 1871–1882. https://doi.org/10.1038/s41593-019-0500-6

Espinoza-Monroy, M., & de Lafuente, V. (2021). Discrimination of regular and irregular rhythms explained by a time difference accumulation model. *Neuroscience, 459*, 16–26. https://doi.org/10.1016/J.NEUROSCIENCE.2021.01.035

Ferrandez, A. M., Hugueville, L., Lehéricy, S., & Poline, J. B. (2003). Basal ganglia and supplementary motor area subtend duration perception: An fMRI study. *NeuroImage, 19*(4), 1532–1544. https://www.sciencedirect.com/science/article/pii/S1053811903001599

Finnerty, G. T., Shadlen, M. N., Jazayeri, M., Nobre, A. C., & Buonomano, D. V. (2015). Time in cortical circuits. *Journal of Neuroscience, 35*(41), 13912–13916. https://doi.org/10.1523/JNEUROSCI.2654-15.2015

García-Garibay, O., Cadena-Valencia, J., Merchant, H., & de Lafuente, V. (2016). Monkeys share the human ability to internally maintain a temporal rhythm. *Frontiers in Psychology, 7*(DEC), 1–12. https://doi.org/10.3389/fpsyg.2016.01971

Ghose, G., & Maunsell, J. (2002). Attentional modulation in visual cortex depends on task timing. *Nature, 419*(6907), 616–620. https://www.nature.com/articles/nature01057

Gibbon, J., Malapani, C., Dale, C., & Gallistel, C. R. (1997). Toward a neurobiology of temporal cognition: Advances and challenges. *Current Opinion in Neurobiology, 7*(2), 170–184. https://www.sciencedirect.com/science/article/pii/S0959438897800050

Grahn, J., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience, 19*(5), 893–906. https://direct.mit.edu/jocn/article-abstract/19/5/893/4337

Grahn, J., & Rowe, J. (2009). Feeling the beat: Premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience, 29*(23), 7540–7548. https://doi.org/10.1523/JNEUROSCI.2018-08.2009

Grondin, S. (2001). From physical time to the first and second moments of psychological time. *Psychological Bulletin, 127*(1), 22. https://psycnet.apa.org/record/2001-16276-002

Hary, D., & Moore, G. P. (1987). Synchronizing human movement with an external clock source. *Biological Cybernetics, 56*(5–6), 305–311. https://doi.org/10.1007/BF00319511

Ivry, R., & Spencer, R. (2004). The neural representation of time. *Current Opinion in Neurobiology, 14*(2), 225–232. https://www.sciencedirect.com/science/article/pii/S0959438804000455

Jin, D. Z., Fujii, N., & Graybiel, A. M. (2009). Neural representation of time in cortico-basal ganglia circuits. *Proceedings of the National Academy of Sciences of the United States of America, 106*(45), 19156–19161. https://doi.org/10.1073/PNAS.0909881106

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: Encoding time in neural network states. *Neuron, 53*(3), 427–438. https://doi.org/10.1016/j.neuron.2007.01.006

Leinweber, M., Ward, D. R., Sobczak, J. M., Attinger, A., & Keller, G. B. (2017). A sensorimotor circuit in mouse cortex for visual flow predictions. *Neuron, 95*(6), 1420–1432.e5. https://doi.org/10.1016/J.NEURON.2017.08.036

Matell, M., & Meck, W. (2000). *Neuropsychological mechanisms of interval timing behavior*. https://doi.org/10.1002/(SICI)1521-1878(200001)22:1

Matell, M., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cognitive Brain Research, 21*(2), 139–170. https://www.sciencedirect.com/science/article/pii/S0926641004001697

Matell, M., Meck, W., & Lustig, C. (2007). Not "just" a coincidence: Frontal-striatal interactions in working memory and interval timing. *Taylor & Francis, 13*(3–4), 441–448. https://doi.org/10.1080/09658210344000404

Mates, J. (1994). A model of synchronization of motor acts to a stimulus sequence—I. Timing and error corrections. *Biological Cybernetics, 70*(5), 463–473. https://doi.org/10.1007/BF00203239

Mauk, M. D., & Buonomano, D. V. (2004). The neural basis of temporal processing. *Annual Review of Neuroscience, 27*(1), 307–340. https://doi.org/10.1146/annurev.neuro.27.070203.144247

McAuley, J. (1995). *Perception of time as phase: Toward an adaptive-oscillator model of rhythmic pattern processing*. https://search.proquest.com/openview/9c3e298761c5ca4d84a8a07fe6b9a0e1/1?pq-origsite=gscholar&cbl=18750&diss=y

Merchant, H., & Averbeck, B. (2017). The computational and neural basis of rhythmic timing in medial premotor cortex. *Journal of Neuroscience, 37*(17), 4552–4564. https://doi.org/10.1523/JNEUROSCI.0367-17.2017

Merchant, H., Zarco, W., Pérez, O., Prado, L., & Bartolo, R. (2011). Measuring time with different neural chronometers during a synchronization-continuation task. *Proceedings of the National Academy of Sciences of the United States of America, 108*(49), 19784–19789. https://doi.org/10.1073/pnas.1112933108

Merchant, H., Pérez, O., Zarco, W., & Gámez, J. (2013). Interval tuning in the primate medial premotor cortex as a general timing mechanism. *Journal of Neuroscience, 33*(21), 9082–9096. https://doi.org/10.1523/JNEUROSCI.5513-12.2013

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: A neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society, B: Biological Sciences, 370*(1664). https://doi.org/10.1098/RSTB.2014.0093

Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., & Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nature Neuroscience, 12*(4), 502–507. https://www.nature.com/articles/nn.2272

Patel, A. D., Iversen, J. R., Chen, Y., & Repp, B. H. (2005). The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research, 163*(2), 226–238. https://doi.org/10.1007/S00221-004-2159-8

Paton, J., & Buonomano, D. V. (2018). The neural basis of timing: Distributed mechanisms for diverse functions. *Neuron, 98*(4), 687–705. https://www.sciencedirect.com/science/article/pii/S0896627318302514

Rao, S. M., Harrington, D. L., Haaland, K. Y., Bobholz, J. A., Cox, R. W., & Binder, J. R. (1997). Distributed neural systems underlying the timing of movements. *Journal of Neuroscience, 17*(14), 5528–5535. https://www.jneurosci.org/content/17/14/5528.short

Remington, E. D., Egger, S. W., Narain, D., Wang, J., & Jazayeri, M. (2018a). A dynamical systems perspective on flexible motor timing. *Trends in Cognitive Sciences, 22*(10), 938–952. https://doi.org/10.1016/J.TICS.2018.07.010

Remington, E. D., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018b). Flexible sensorimotor computations through rapid reconfiguration of cortical dynamics. *Neuron, 98*(5), 1005–1019.e5. https://doi.org/10.1016/J.NEURON.2018.05.020

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. *Journal of Motor Behavior, 35*(4), 355–370. https://doi.org/10.1080/00222890309603156

Repp, B. H. (2005). Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin and Review, 12*(6), 969–992. https://doi.org/10.3758/BF03206433

Robbe, D. (2023). Lost in time: Relocating the perception of duration outside the brain. *Neuroscience & Biobehavioral Reviews,* 153105312. https://doi.org/10.1016/j.neubiorev.2023.105312.

Schubotz, R., & von Cramon, D. (2002). Predicting perceptual events activates corresponding motor schemes in lateral premotor cortex: An fMRI study. *NeuroImage, 15*(4), 787–796. https://doi.org/10.1006/nimg.2001.1043

Schubotz, R., & Von Cramon, D. (2004). Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery. *Journal of Neuroscience, 24*(24), 5467–5474. https://doi.org/10.1523/JNEUROSCI.1169-04.2004

Schulze, H. H. (1978). The detectability of local and global displacements in regular rhythmic patterns. *Psychological Research, 40*(2), 173–181. https://doi.org/10.1007/BF00308412

Schulze, H. H. (1989). The perception of temporal deviations in isochronic patterns. *Perception & Psychophysics, 45*(4), 291–296. https://doi.org/10.3758/BF03204943

Simen, P., Balci, F., Desouza, L., Cohen, J. D., & Holmes, P. (2011). A model of interval timing by neural integration. *The Journal of Neuroscience*. https://doi.org/10.1523/JNEUROSCI.3121-10.2011

Teki, S., Grube, M., Kumar, S., & Griffiths, T. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *Journal of Neuroscience, 31*(10), 3805–3812. https://doi.org/10.1523/JNEUROSCI.5561-10.2011

Treisman, M., Faulkner, A., Naish, P. L., & Brogan, D. (1990). The internal clock: Evidence for a temporal oscillator underlying time perception with some estimates of its characteristic frequency. *Perception, 19*(6), 705–743. https://doi.org/10.1068/P190705

Wiener, M., & Kanai, R. (2016). Frequency tuning for temporal perception and prediction. *Current Opinion in Behavioral Sciences, 8*, 1–6. https://doi.org/10.1016/J.COBEHA.2016.01.001

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127*(1), 3–21. https://doi.org/10.1037//0033-2909.127.1.3

Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology, 102*(6), 3191–3202. https://doi.org/10.1152/JN.00066.2009

Cognition of Time and Thinking Beyond

Zedong Bi

Abstract

A common research protocol in cognitive neuroscience is to train subjects to perform deliberately designed experiments while recording brain activity, with the aim of understanding the brain mechanisms underlying cognition. However, how the results of this protocol of research can be applied in technology is seldom discussed. Here, I review the studies on time processing of the brain as examples of this research protocol, as well as two main application areas of neuroscience (neuroengineering and brain-inspired artificial intelligence). Time processing is a fundamental dimension of cognition, and time is also an indispensable dimension of any real-world signal to be processed in technology. Therefore, one may expect that the studies of time processing in cognition profoundly influence brain-related technology. Surprisingly, I found that the results from cognitive studies on timing processing are hardly helpful in solving practical problems. This awkward situation may be due to the lack of generalizability of the results of cognitive studies, which are under well-controlled laboratory conditions, to real-life situations. This lack of generalizability may be rooted in the fundamental unknowability of the world (including cognition). Overall, this paper questions and criticizes the usefulness and prospect of the abovementioned research protocol of cognitive neuroscience. I then give three suggestions for future research. First, to improve the generalizability of research, it is better to study brain activity under real-life conditions instead of in well-controlled laboratory experiments. Second, to overcome the unknowability of the world, we can engineer an easily accessible surrogate of the object under investigation, so that we can predict the behavior of the object under investigation by experimenting on the surrogate. Third, the paper calls for technology-oriented research, with the aim of technology creation instead of knowledge discovery.

Keywords

Metascience · Unknowability · Cognitive neuroscience · Generalizability · Technology application

Z. Bi (✉)
Lingang Laboratory, Shanghai, China

Institute for Future, Qingdao University,
Qingdao, China

School of Automation, Shandong Key Laboratory of
Industrial Control Technology, Qingdao University,
Qingdao, China
e-mail: bizedong@lglab.ac.cn

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental
Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_10

Prologue

> Humans are great through reasoning, but are matured by recognizing the limitations of reasoning. ---Prof. Qing Liu, School of Politics and International Relations, East China Normal University, Shanghai

"So interesting! Why is it?" This is perhaps the question that inspires your curiosity about the brain and marks the beginning of a neuroscience research journey. However, before you embark on such an investigation, I advise to think twice about whether the mechanism underlying the phenomenon is worth investigating. In most cases, such investigation is useless in solving practical problems.

"Too short-sighted!" You may criticize me. You may believe that even though your results cannot lead to practical breakthroughs directly, they belong to the ongoing accumulation of knowledge about the brain. As the accumulation continues, people will eventually have a very good understanding of the brain and develop advanced brain technology to solve practical problems.

However, your criticism neglects a possibility: Some aspects of the brain may be unknowable. If such is the case, we may never be able to fully understand the brain, regardless of how much knowledge we accumulate. This unknowability reflects the fundamental limitations of human reasoning capabilities.

If you ever doubt the limitations of human reasoning, take a trip to an art museum. As you peruse the galleries, you may ask yourself: Is it possible to develop a logical system that, through a series of if-then reasoning, could lead to the creation of a masterpiece? If you doubt the existence of such a logical system for the creation of art, then why do you believe that a logical system for the workings of the brain exists? After all, the brain is believed to be much more complex than any artwork created by humans.

In this paper, I will review literature that highlights the limitations of mechanism-investigating research in solving practical problems. I will then explore the concept of the unknowability of the brain through the lenses of neuroscience, philosophy, physics, and AI. Finally, I will provide suggestions for conducting meaningful research in light of this unknowable reality.

Introduction

Atomism, the idea that the universe is composed of fundamental components known as atoms, is perhaps the most influential philosophy leading scientific research. Richard Feynman considered atomism to be the most important thinking we should pass on to the next generation (Feynman et al., 2011), as various physical changes and chemical reactions can be explained by supposing the movements and interactions of atoms (Feynman et al., 2011).

Atomism has also had a strong influence on cognitive neuroscience. Psychologists have divided cognition into several elements, including perception, learning, memory, and decision-making (Baldwin, 1893). Each of these elements can be further divided into several sub-elements from different perspectives. For example, perception can be divided into the perception of space and time or into visual and auditory perception. Memory can be divided into short-term and long-term memory, or episodic and semantic memory, among other things. After investigating the brain activity when the subject is performing each element of cognition, neuroscientists aim to understand the biological backend of cognition by collecting all these pieces together (Fig. 1a). From this atomistic perspective, studying a single cognitive element is the foundation for understanding cognition, which is why I name this research protocol to be *basic*.

To perform basic cognitive studies, researchers elaborately designed simple and well-controlled experimental conditions to study a single cognitive element while teasing apart the influence from other elements. For example, to study working memory, researchers trained monkeys to recall a visual cue after a delay period (Constantinidis et al., 2001) (Fig. 1b). To study decision-making, researchers trained monkeys to watch two types of dots moving toward opposite directions and then decide which type had more dots (Roitman & Shadlen, 2002) (Fig. 1c).

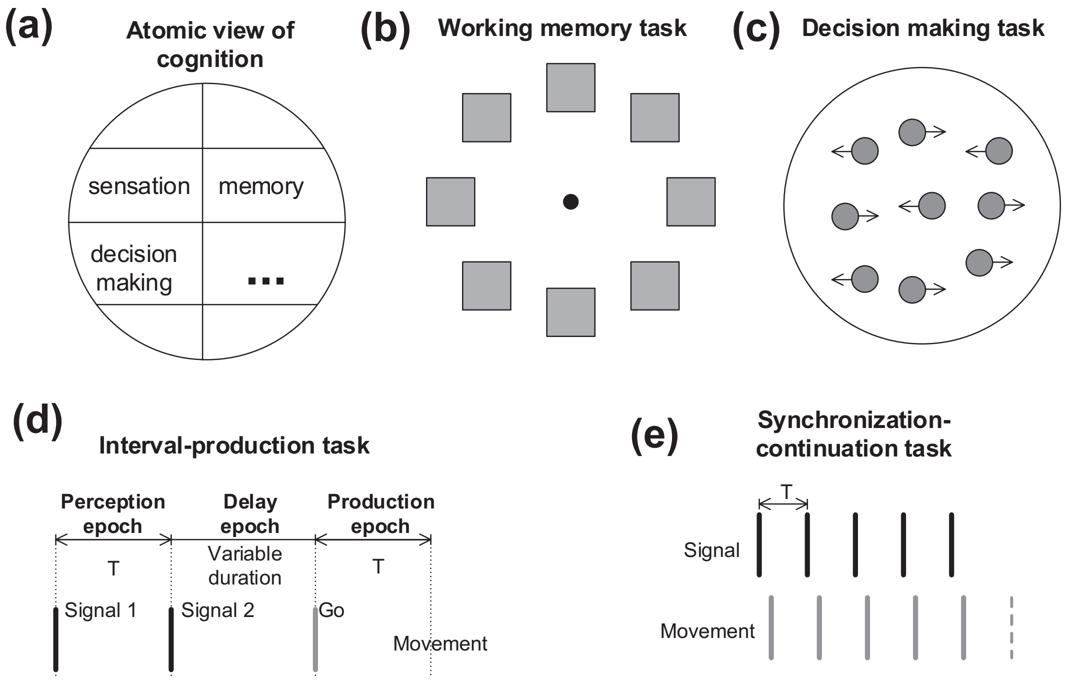


Fig. 1 Illustration of basic cognitive studies. (**a**) Basic cognitive studies are guided by the philosophy of atomism, which divides cognition into many elements, each of which is studied separately. Atomists believe that by understanding each element, we can eventually understand the whole of cognition. (**b**) A classical experiment to study working memory. The subject fixates on a central point, and a visuospatial cue (one of the eight gray boxes) is presented briefly, followed by a mnemonic delay. After the delay, the subject must make a saccadic eye movement to the remembered location. (**c**) A classical experiment to study decision-making. There are two types of random dots, one moving leftward and the other moving rightward. The subject must decide which type has more dots. (**d**) Schematic of the time production task. The subject receives two signals (black bars) separated by a time interval *T*; after a delay epoch with variable duration, a go cue (gray bar) appears, and the subject must move at time *T* after the go cue. (**e**) Schematic of the synchronization-continuation task. The subject must move (gray bars) immediately following a sequence of signals (black bars) with period *T*. The subject must still move with period *T* (dashed bar) after the signal was removed

Another example is the study of time cognition, which also stems from atomism. To focus on the processing of time while disentangling other cognitive elements (such as the perception of spatial information), psychologists or neuroscientists train subjects to perform simple but deliberately designed timing tasks. In a classical experiment (Rakitin et al., 1998), participants were presented with specific time intervals delimited by stimuli and then were asked to reproduce the interval (Fig. 1d). When subjects were performing these simple and deliberately designed tasks, researchers recorded subjects' brain activity to propose neural network mechanisms underpinning basic elements of cognition.

While some scientists think that the pure aim of science is to satisfy our curiosity about the world, I believe that scientific results must be implemented in technology and benefit the mass of people before scientific results complete their mission. However, the status and prospects of the technological applications of basic cognitive studies have seldom been discussed. In this paper, I will discuss the technological applications of basic cognitive studies, starting with a review of cognitive studies of time processing in the brain (i.e., basic timing studies) as examples of basic cognitive studies. Time processing is an indispensable dimension of cognition (Merchant et al., 2013), and time is also an indispensable dimension of any real-world signal to be processed by technology. Therefore, one may expect that the results of basic timing studies lay down the foundations for processing temporal signals in brain-related technology. Unfortunately, after reviewing two fields of brain-related technology, neuroengineering for brain health and brain-inspired artificial intelligence, which are two promising application fields of neuroscience suggested by the China Brain Project (Poo et al., 2016), I found that the results of basic timing studies are hardly helpful in solving practical problems.

I will attempt to clarify this awkward situation and offer suggestions for future research (Fig. 2). In my view, the challenge of applying basic timing studies (and, more broadly, basic cognitive studies) to technology stems from their lack of generalizability. In other words, the results of these studies are contingent on the specific conditions and tasks of the laboratory experiments that produced them and may not be applicable in other contexts. This lack of generalizability may be rooted in the fundamental unknowability of the world, including cognition. In other words, the capability of knowledge to describe the world is fundamentally limited, so the generalizability of our knowledge to various situations in the world is fundamentally limited, and therefore, the capability of knowledge to guide technological creation to change the world is also fundamentally limited.

I suggest three ways to improve future research (Fig. 2). Firstly, to improve the generalizability of results, researchers should analyze brain activity in real-life settings, rather than simple tasks in well-controlled experimental conditions, and examine their results under various situations. Secondly, to deal with the unknowability of the world, researchers should engineer surrogates of the object under investigation, so that they can predict the behavior of the investigated object using the surrogate, even without understanding how the object under investigation works. Finally,

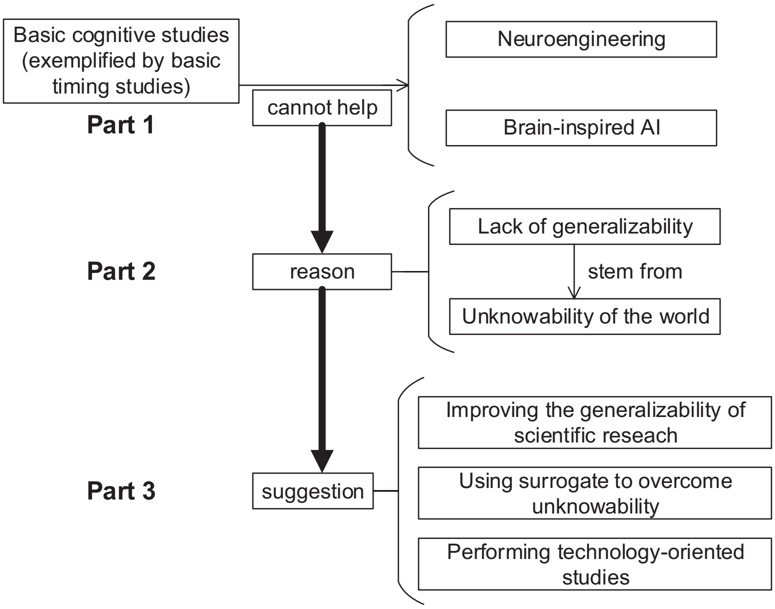


Fig. 2 Overview of this paper. In Part 1, the main results of basic timing studies and two application fields (neuroengineering and brain-inspired AI) are reviewed, showing that basic timing studies (and perhaps more generally, basic cognitive studies) cannot help the application fields of neuroscience. In Part 2, it is proposed that this situation is due to the lack of generalizability of the basic timing studies and, more fundamentally, the unknowability of the world. Finally, in Part 3, researchers are suggested to improve the generalizability of their results, engineer surrogates to overcome the unknowability of the world, and perform technology-oriented studies

due to the fundamental unknowability of the world, I suggest that researchers focus on technology-oriented research with the aim of creating new technologies, rather than science-oriented research with the aim of discovering knowledge.

Basic Timing Studies

This section provides an overview of basic timing studies, which use two main paradigms to study time cognition. The first paradigm is interval timing, which involves training the subject to perceive or produce a single time interval (Fig. 1d). The second paradigm is beat timing, which involves training the subject to perceive or produce a sequence of time intervals rhythmically (Fig. 1e). In the interval-production task (Rakitin et al., 1998), an example of the first paradigm, the subject is presented with a specific time interval delimited by stimuli and is then asked to reproduce the interval (Fig. 1d). In the synchronization-and-continuation task (Gámez et al., 2019), an example of the second paradigm, the subject is required to act following a sequence of rhythmic stimuli and continue to act rhythmically after the removal of the stimuli (Fig. 1e). By recording brain activity during these tasks, researchers can discover features of brain dynamics related to time cognition and gain insight into the neural network mechanisms underpinning time cognition.

In the interval-production task (Fig. 1d), a neural network perceives time by evolving its state along a stereotypical trajectory in the perception epoch, maintains time intervals in working memory using a manifold of line attractor in the delay epoch, and predicts a coming event by evolving its state along isomorphic trajectories with the speed of state evolution inversely scaling with the to-be-produced time interval in the production epoch (Bi & Zhou, 2020a) (Fig. 3a). These dynamic features align with experimental findings from other interval-timing tasks (Jin et al., 2009; Mita et al., 2009; Wang et al., 2018). In the synchronization-continuation task (Fig. 1e), the network encodes different beating periods T using circular trajectories (Gámez et al., 2019) (Fig. 3b). The radii of these circular trajectories increase with the period T, but the speed of state evolution with time remains constant across different values of T.

In both the perception and production epochs of the interval-production task, as well as in the beating intervals in the synchronization-continuation task, the neural network relies on state evolution along trajectories to sense the passage of time. This state evolution can be achieved through several mechanisms, including the pacemaker-accumulator model (Buhusi & Meck, 2005) (recently supported in (Cook et al., 2022)), in which an accumulator counts the number of pulses received from a pacemaker (Fig. 3c, left); the synfire chain model (Zeki & Balci, 2019), in which a chain of neurons is sequentially excited (Fig. 3c, middle); and the striatal beat-frequency model (Matell & Meck, 2004), in which a group of oscillators with heterogeneous frequencies have their phases reset by the stimulus (Fig. 3c, right).

Anatomically, several brain areas have been identified as participating in timing, including the basal ganglia (Jin et al., 2009), supplementary motor area (SMA) (Mita et al., 2009), sensory cortex (Shuler & Bear, 2006), and prefrontal cortex (Wang et al., 2018). There is ongoing debate about whether timing relies on dedicated circuits in the brain or on intrinsic computation that emerges from the inherent dynamics of neural circuits (Paton & Buonomano, 2018; Ivry & Schlerf, 2008). A prevailing viewpoint is that timing depends on the interaction of core timing areas, such as the basal ganglia and SMA, which are consistently involved in temporal processing across various contexts, and other areas, such as the prefrontal cortex, sensory cortex, and cerebellum, which are activated in a context-dependent manner (Merchant et al., 2013).

At the behavioral level, the most well-known timing principle is the scaling property, which posits that the variance of time interval estimation is proportional to the mean of the estimation (Allman et al., 2014).

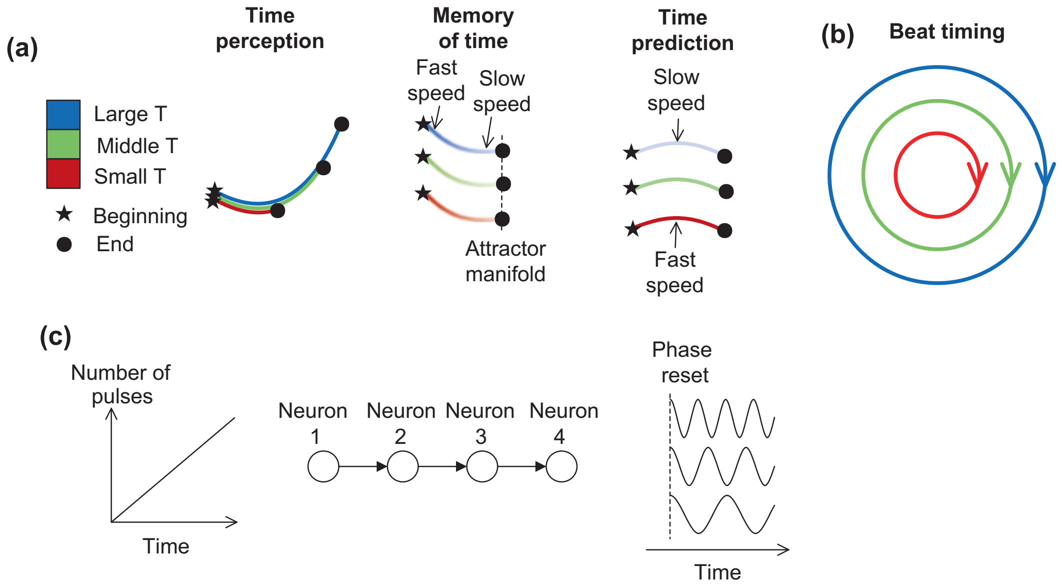


Fig. 3 Some results of basic timing studies. (**a**) Dynamic features of neural networks in the interval-production task. Left: For time perception in the perception epoch (see Fig. 1d), the network exhibits a stereotypical trajectory whose final position determines the perceived time interval *T* (see Fig. 1d). Lines with blue, green, and red colors, respectively, represent the trajectories when *T* is large, middle, and small. Asterisk and circle, respectively, represent the beginning and end of the trajectory. Middle: In the delay epoch, time intervals are maintained in the working memory as positions (black dots) in an attractor manifold. The speed of state evolution with time decreases near the attractor (indicated by the lighter color near the attractor). Right: In the production epoch, time prediction is performed when the network state evolves along isomorphic trajectories, with the speed of state evolution inversely scaling with the to-be-produced interval *T*. (**b**) Dynamic features of neural networks in the synchronization-continuation task. With different periods *T* (see Fig. 1e), the network state evolves along different circular trajectories at the same speed, but the radius of the circular trajectory increases with *T*. (**c**) Some computational models for the neural mechanisms of time sensing. Left: In the pacemaker-accumulator model, time is measured by the accumulated number of pulses emitted from the pacemaker. Middle: In the synfire chain model, time can also be measured by the sequential firing of a chain of neurons. Right: In the beat-frequency model, time is measured by the activity pattern of a group of oscillators with heterogeneous frequencies after phase resetting

Brain-Related Technology

Time processing is a fundamental aspect of cognition (Merchant et al., 2013), and time is also an indispensable dimension of any real-world signal to be processed in technology. Therefore, one might expect that studies on time processing in cognition would profoundly influence brain-related technology. This section will review two fields of brain-related technology, neuroengineering for brain health and brain-inspired artificial intelligence, which are two promising application fields of neuroscience suggested by the China Brain Project (Poo et al., 2016). Unfortunately, we will see that the results from basic timing studies are hardly helpful in solving practical problems.

Neuroengineering for Brain Health

Neuroengineering involves designing interfaces between living neural tissue and non-living constructs in order to understand, repair, replace, or enhance neural systems (Hetling, 2008). In this paper, I will review neuroengineering techniques used for the therapy of Parkinson's disease through deep brain stimulation, the diagnosis of epilepsy through neuroimaging, and the development of speech prostheses through machine translation of brain activity into language.

Parkinson's disease is closely related to pathological changes in the basal ganglia (Poewe et al., 2017), a core timing area of the brain (Merchant et al., 2013). Epilepsy also recruits timing-related regions such as the thalamus, basal ganglia, and frontal lobe (Bertram, 2009; Wu et al., 2019). As a result, patients with either Parkinson's disease or epilepsy may experience distortion of timing perception (Gu et al., 2016; Greyson et al., 2014; Cainelli et al., 2019). Besides, language has rich hierarchical temporal structures, and the processing of language may share a similar neural substrate with the processing of music (Patel, 2003; Janata & Grafton, 2003; Hickok, 2012). Therefore, it is reasonable to assume that basic timing studies could be of great help in the therapy and diagnosis of Parkinson's disease and epilepsy, as well as the machine translation of brain activity into language. However, I will show below that this is not the case.

Neuroengineering Is Driven by Clinical Data and Experience

Deep brain stimulation (DBS) therapy for Parkinson's disease was pioneered by Lawrence Pool, who implanted an electrode into the caudate nucleus of a female patient in 1948 (Pool, 1954). Traditional DBS is an open loop, where the clinician sets parameters of the controller that deliver short-duration (60–180 ms) and high-frequency (typically 130–185 Hz) pulses of electrical stimulation to alleviate symptoms (Benabid et al., 1994; Limousin et al., 1995; Siegfried & Lippitz, 1994) (Fig. 4a, left). However, this type of DBS cannot adapt its stimulation according to the feedback from the patients and has several drawbacks, such as adverse effects such as dyskinesia and high battery consumption (Bouthour et al., 2019; Krauss et al., 2021). Recently developed closed-loop DBS overcomes these problems by delivering stimulation only when pathological biomarkers are detected (Bouthour et al., 2019; Krauss et al., 2021) (Fig. 4a, right).

Interestingly, despite the broad success and application of DBS, the mechanism by which DBS ameliorates Parkinson's disease is still not fully understood, although some mechanisms related to neuronal circuits, astrocytes, and neurogenesis have been proposed (Okun, 2012). Due to the lack of understanding of the mechanism, the technical details of DBS have been established mainly through empirical means. For example, the optimal stimulation waveform shape in open-loop DBS was determined by systematically varying stimulation parameters and examining the therapeutic effects (Rizzone et al., 2001; Kuncel et al., 2006). The most prominent biomarker used in closed-loop DBS, excessively synchronized beta oscillation, was also discovered through empirical comparisons between normal and diseased brains (Oswal et al., 2013; Cheyne, 2013). Therefore, mechanical insight, which is the aim of basic timing studies (Fig. 3), is not the primary driving force behind the development of DBS.

Though the mechanical insights provided by basic timing studies may not currently be helpful in the research of DBS, one might still expect that they could be useful in the future. However, recent research trends suggest a dominance of data-driven automatic design in the development of DBS technology, rather than a rational implementation of mechanical knowledge. As mentioned earlier, closed-loop DBS delivers stimulation into the brain only when pathological activities (i.e., biomarkers) are detected. Traditionally, excessive beta oscillation was predetermined as the key biomarker of patients' tremors in Parkinson's disease (Bouthour et al., 2019; Krauss et al., 2021). However, in two recent studies (Shah et al., 2018; Tan et al., 2019), the authors recorded patients' body movements using accelerometers and recorded local field potentials (LFPs) using electrodes. They then trained binary classifiers to detect the LFPs during tremor or non-tremor periods. Here, the detector (i.e., the binary classifier) is trained by clinical data, instead of being rationally designed using our knowledge of the mechanisms of Parkinson's disease. A similar data-driven approach has also been used to detect biomarkers of depression (Scangos et al., 2021a, b), where a classifier was trained to map stereoelectroencephalography (SEEG) recordings to depression scores measured by a psychological questionnaire.

This data-driven approach is also the mainstream of other neuroengineering techniques. For example, in recent studies on epilepsy diagnosis, neural network models were built to simulate the large-scale dynamics of the brain. The models

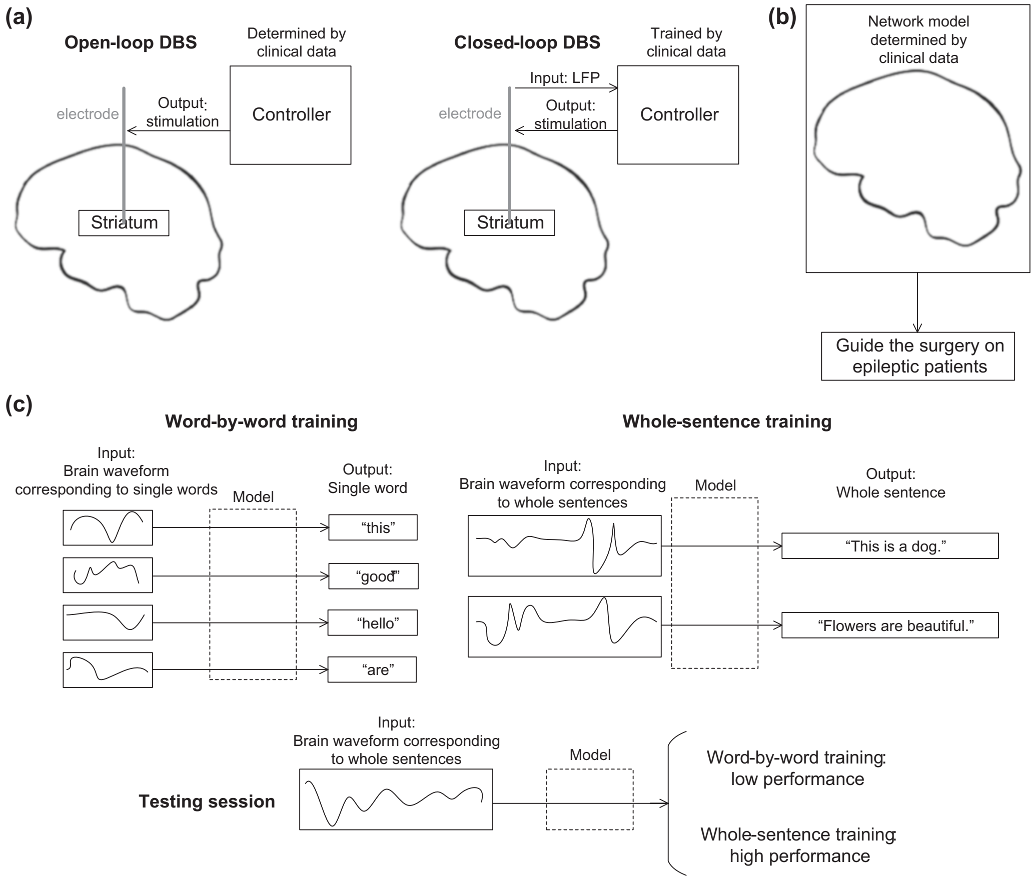


Fig. 4 Some neuroengineering techniques. (**a**) In open-loop deep brain stimulation (DBS) (left), the stimulation waveform is preset in the controller. In closed-loop DBS (right), the stimulation waveform can be adjusted according to the local field potential (LFP) of the brain. In both types of DBS, the controller is programmed based on clinical data, rather than on mechanical understandings from basic cognitive studies. (**b**) Neural network models that simulate large-scale dynamics of the brain have been used to guide the surgery on epileptic patients. The parameters of the model are also determined by clinical data, rather than mechanical understandings from basic cognitive studies. (**c**) There are two strategies to build the model to translate brain activities to language for speech prosthesis. In the word-by-word training strategy (upper left), the model is trained to map the brain waveforms when the patient is speaking single words to single spoken words. In the whole-sentence strategy (upper right), the model is trained to map the brain waveforms when the patient is speaking whole sentences to the whole spoken sentences. In the test session (lower), the trained model is used to translate the brain waveforms corresponding to whole sentences. The whole-sentence training strategy results in better performance than the word-by-word strategy

were used to identify the ictogenic zone of seizures and guide the resection of brain areas in clinical surgery (Cao et al., 2022; Sinha et al., 2017) (Fig. 4b). In their neural network models, the connection strengths were determined through the fitting of empirical data, rather than being rationally designed based on mechanical insights into epilepsy or the information processing of the brain. Another example is the machine translation of brain activities to language, which can be used as a speech prosthesis for degenerative motor diseases such as amyotrophic lateral sclerosis and locked-in syndrome (Fig. 4c). Traditional approaches trained translation machines by mapping neuroimaging signals to individual words or even sub-word syllabic features (such as vowel harmonics and fricative consonants) (Pasley et al., 2012; Angrick et al., 2019) (Fig. 4c, upper left). However, the best brain-to-language translation performance is now realized by training recurrent neural networks in an end-to-end manner, mapping brain signals to entire sentences rather than single words or syllabic features (Makin et al., 2020; Cogan, 2020; Moses et al., 2019) (Fig. 4c, upper right). Although understanding the brain activities related to words and even syllables may seem more "basic" and "mechanical," implementing such understandings in the technique results in worse performance than directly training a neural network to map brain signals to entire sentences (Fig. 4c, lower).

Summary

Overall, basic timing studies (Fig. 3), though expected to be "basic," unfortunately do not lead to the progress of application-oriented research. This awkward situation is due to a gap in that basic timing studies aim for mechanistic explanations for simple timing tasks, but neuroengineering, which aims for good performance in practical use, is not mainly driven by mechanistic understandings of the brain, but by clinical data and experience. This gap not only exists between basic timing studies and neuroengineering, as we have discussed here, but, more generally, between basic cognitive studies (Fig. 1) and neuroengineering. Therefore, we may conclude that basic cognitive studies do not lead to the progress of neuroengineering.

Brain-Inspired Artificial Intelligence

Brain-inspired artificial intelligence (AI) is another potential application field of neuroscience. Brain-inspired AI aims to build strong AI (i.e., AI that has mental capabilities and functions that mimic the human brain, or in other words, can pass the Turing test) by mimicking the structure and function of the brain, through the implementation of neuroscience knowledge in AI engineering (Hassabis et al., 2017). I have a criticism of this brain-inspired approach to AI, though due to ethical concerns, detailed investigations on the human brain cannot be performed. As a result, brain-inspired AI can only closely mimic the brain of animals, which has low-level intelligence, rather than that of humans, whose high-level intelligence is the ultimate aim. Therefore, the brain-inspired approach should not be the leading approach to strong AI in the long run. I will talk about the possible approach to strong AI at the end of this subsection; at present, however, let us forget this criticism and think about how basic timing studies may contribute to brain-inspired AI. Unfortunately, I will show that basic timing studies are also of little help to this field.

The Inspiration for AI from Neuroscience

The inspiration for AI from neuroscience is found at the levels of neurons, synapses, and neural networks. This is exemplified below:

1. Single Neuron Level.
    (a) Biological neurons fire spikes, unlike artificial analog neurons, whose activities take continuous values. Implementing spiking neurons in hardware significantly reduces energy consumption compared to analog neurons (Frenkel, 2021). The reason is that the membrane voltage of spiking neurons stays near the resting state most of the time due to the sparsity of spiking periods, resulting in small leaky currents.
    (b) Biological neurons also have rich internal dynamics due to the interaction between the membrane voltage and ion channels (Dayan & Abbott, 2001), unlike artificial neurons, which are usually nonlinear filters of total synaptic currents. Such rich internal dynamics significantly improve the computational power of biological neurons (Beniaguev et al., 2021). Recently, it has been found that only 19 neurons with internal dynamics can make up a full-stack autonomous vehicle control system (Lechner et al., 2020).
2. Single Synapse Level.
    (a) Biological synapses have binary efficacies (O'Connor et al., 2005), unlike in artificial networks where synaptic weights typically take continuous values. Binary-weight artificial neural networks have

been investigated and broadly used due to their low computation and memory cost, as well as performance that is comparable with continuous-weight networks (Courbariaux et al., 2016).

   (b) Biological synapses also have hidden states other than synaptic efficacy, which arise from the complex interactions of proteins in synapses (Graupner & Brunel, 2010). Adding hidden synaptic states in artificial neural networks improves memory capacity and learning performance (Baldassi et al., 2007; Kirkpatrick et al., 2017). The reason is that a high hidden state of a synapse can indicate that this synapse is important for the good performance of a task; therefore, protecting the efficacy of synapses with high hidden states from being changed in the further training process can maintain the performance of the neural network during further training.

3. Neural Network Level.
   (a) Memory replay, found in the hippocampus and cortex (Ji & Wilson, 2007), is a phenomenon in which the neuronal firing sequence in sleep or at rest closely matches the firing sequence in the real experience just before. Memory replay inspires DQN (Mnih et al., 2015), a well-known deep reinforcement learning algorithm that guides actions according to perceptual inputs in order to maximize future rewards. Besides, memory replay is also used in the Dyna algorithm (Sutton & Barto, 2018) to train a mental model of the environment. After training, the agent can predict the outcome of an action in situations never seen before using this mental model, facilitating the agent to adapt to more complicated environments.
   (b) Biological neurons are subject to gain modulation, which means that one input, the modulatory one, affects the sensitivity of a neuron to another input (Salinas & Thier, 2000; Salinas & Sejnowski, 2001). Gain modulation is the neural mechanism of attention. With attention mechanisms, a neural network looks at an image or input sequence and decides which parts of the image or sequence are important for the task at hand and then sends only the important parts to subsequent information processing. Attention mechanisms have become an indispensable component of advanced image and language processing models (Vaswani et al., 2017; Devlin et al., 2019).
   (c) Context-dependent gating (Cichon & Gan, 2015) means that different sparse sets of dendritic branches are disinhibited when the brain is involved in different tasks. This mechanism allows the brain to recruit different dendritic branches for different tasks, so that the synaptic weights learned for one task will not interfere with the configuration learned for another task. Such context-dependent gating has been implemented in artificial neural networks to avoid catastrophic forgetting during continual learning (Manning et al., 2020; Zeng et al., 2019).

Basic Timing Studies Hardly Inspire AI

From the examples provided (also see (Hassabis et al., 2017) for a detailed review), it is clear that basic timing studies do not have a significant impact on the development of brain-inspired AI, despite time processing being a fundamental aspect of brain cognition. Similarly, other basic cognitive topics, such as working memory and decision-making (Fig. 1b, c), though attract great interest in the neuroscience community, and they also contribute little to brain-inspired AI. There are two possible reasons for the limited impact of basic cognitive studies in AI applications:

1. Lack of Generalizability (Fig. 5a).

   All the neural mechanisms implemented in AI have a common property: They are not task-specific. In other words, if a neural mechanism exists only when the brain is performing a simple task like Fig. 1b–e, but does not exist if the brain is performing another more complicated task, this neural mechanism will not be used in AI implementation. The reason

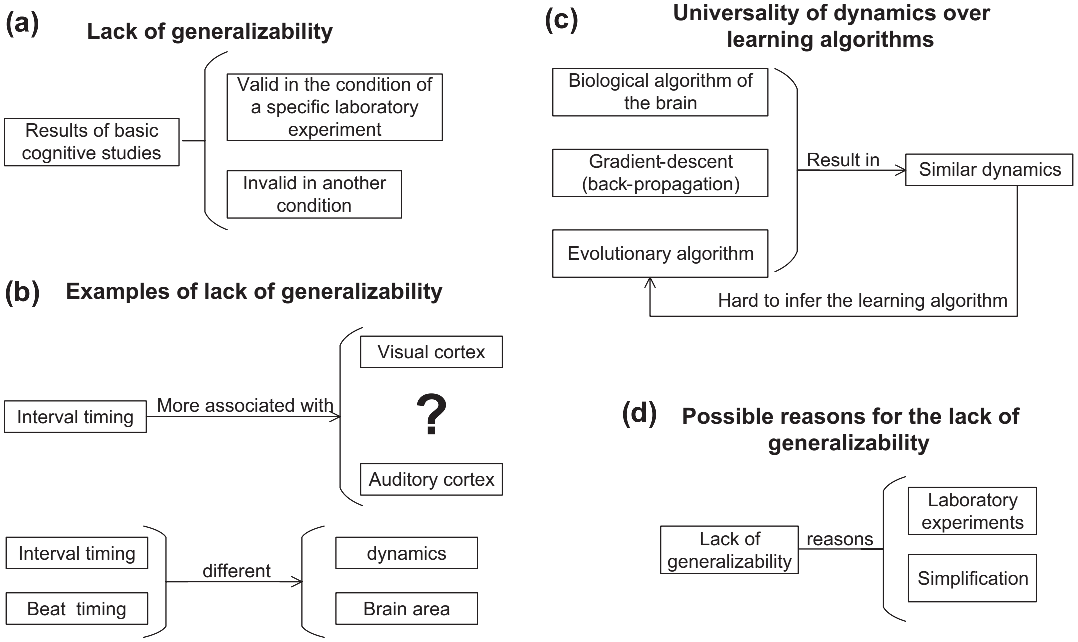


Fig. 5 Pitfalls of basic cognitive studies. (**a**) Basic cognitive studies may lack generalizability, so that the results are valid in a specific laboratory experiment condition but invalid in another condition. (**b**) Examples of lack of generalizability. Upper: Some experiments suggest that interval timing is more associated with the visual cortex, but other experiments suggest that interval timing is more associated with the auditory cortex. Lower: Interval timing and beating timing, though both are pure timing tasks, have different dynamic features and involve different brain areas. (**c**) Different learning algorithms result in similar neuronal population dynamics when the neural network is trained on the same basic cognitive task. Therefore, we cannot infer the learning algorithm that the brain uses through the dynamics when the brain is performing basic cognitive tasks. (**d**) The lack of generalizability may be due to the methodology of laboratory experiments in basic cognitive studies and also the misunderstanding of the ideology of simplification

is simple: AI aims to solve complicated real-world problems, instead of toy problems like Fig. 1b–e designed by neuroscientists.

The lack of generalizability to real-world situations is the main shortcoming of basic cognitive studies. The dynamics of the brain when performing complex tasks cannot be deduced from the dynamics observed when the brain is performing simple tasks. In other words, even if we have a good understanding of the dynamics involved in numerous simple tasks like the one illustrated in Fig. 1b–e, we still do not know the brain dynamics in complex tasks. For example, suppose we let a patient perform simple tasks of speaking single words. Even if we record the brain activity related to numerous single-word speaking, we still do not know the patient's brain activity when speaking a whole sentence: because by dividing a sentence into single words, we are neglecting the syntactic structure of the sentence. This is why the translation of brain activity to language for speech prosthesis achieves better performance when the neural network is trained to translate one sentence at a time instead of one word at a time (Makin et al., 2020; Cogan, 2020; Moses et al., 2019) (Fig. 4c). More generally, cognition requires the coordination of all "basic" elements: perception, memory, decision-making, and so on. Even if we study each "basic" element in isolation, we will still not be able to understand how the brain performs complex real-world tasks that require the coordination of these elements. We will discuss more about the lack of generalizability further in the next section.

2. Lack of Insight into Brain Learning Mechanism (Fig. 5c).

   One may wonder whether the dynamic features observed when the brain performs simple tasks (Fig. 3) depend on the specific learning algorithm of the brain. If a particular learning algorithm results in the experimentally observed features (Fig. 3), while other algorithms do not, we may be able to infer the brain's learning algorithm through these dynamic features. This learning mechanism could then be implemented into AI design. Unfortunately, accumulating evidence suggests that similar dynamic features universally emerge when neural networks are trained on the same basic cognitive task using different learning algorithms. This implies that we may not be able to infer critical information about the learning algorithm that the brain uses through the dynamics observed when the brain is performing simple tasks.

   For example, although error backpropagation (BP) algorithms lack convincing experimental support (Lillicrap et al., 2020), artificial neural networks trained by BP exhibit biologically plausible dynamics in image classification tasks (Hong et al., 2016), language tasks like next-word prediction (Goldstein et al., 2022), and other simple tasks in basic cognitive studies (Bi & Zhou, 2020a; Mante et al., 2013). Recently, I trained recurrent neural networks using an evolutionary algorithm to perform the context-dependent decision-making task (Bi et al., 2022) and found that the resulting network exhibited dynamics closely analogous to those observed in monkey experiments and the dynamics observed in artificial neural networks trained by BP (Mante et al., 2013). The reason for this universality of dynamics across different learning algorithms is unknown, but it is possibly because different algorithms universally tune the synaptic weights into a high-entropy region in the synaptic configuration space (Baldassi et al., 2015; Bi & Zhou, 2020b). Here, "high entropy" means that if we slightly perturb the synaptic weights found by an algorithm, the perturbed weights still probably result in good task performance. Therefore, the weights found by different algorithms are likely to be close to each other in a high-entropy region, which may be the reason for the universal dynamic property of the networks trained by different algorithms. Due to this universality, we cannot gain insight into the learning mechanism of the brain from the dynamic features found in basic cognitive studies, let alone implement the brain learning mechanism in AI.

So what is the approach to strong AI, the machine with intelligence equal to the human brain or even more powerful? In my opinion, the most important thing we should learn from biology is the colossal scale of the human brain. Comparative studies have shown that the human brain contains more neurons than any other animal, which is probably the reason for our superior cognitive abilities (Herculano-Houzel, 2012). Consistently, AI is undergoing a paradigm shift with the rise of colossal models (e.g., BERT (Devlin et al., 2019) and GPT-3 (Brown et al., 2020)) with over 100 billion parameters trained on oceans of data (Bommasani et al., 2021). Such models, trained unsupervisedly, develop geometric representation of knowledge (Manning et al., 2020; Rives et al., 2021), which versatilely serve as the common basis of many task-specific models via adaptation (Bommasani et al., 2021). Most impressively, as the size of the neural network increases, advanced functionalities such as in-context learning naturally emerge (Brown et al., 2020): In-context learning means that the neural network after training can be competent for a task never seen during training, after the trained neural network is instructed by a natural language description of the task. Such colossal models are becoming the trend of AI led by big tech companies such as Google, Microsoft, and Huawei, with broad applications in the text (Devlin et al., 2019), images (Ramesh et al., 2021), protein design (Rives et al., 2021), and chemical reactions (Schwaller et al., 2021). Recently, at a conference, a manager of the colossal model project of Huawei told me that the progress of colossal models is also gradually getting stagnant, because we cannot afford the huge energy consumption to train the model if the model is too large (Strubell et al., 2019). He believed that the next generation of colossal models should be led by the revolution of the computation paradigm, such as quantum computing, which can speed up some kinds of computations by exponential order (Nielsen & Chuang, 2011). At another conference, a professor of quantum physics told me that quantum algorithms are getting mature; the bottleneck of quantum computing lies in its hardware implementation. If their opinions are correct, we may expect that strong AI will naturally emerge after the manufacturing technology of quantum computers is mature and if we train mega-colossal models using quantum computers.

Summary

Overall, basic cognitive studies cannot significantly contribute to AI due to their limitations on generalizability and insight into brain learning mechanism. The future of AI is likely to be led by colossal models.

Contemplation

Time processing is a fundamental aspect of cognition, and time is also an indispensable dimension of any real-world signal to be processed in technology. But why do basic timing studies, which aim to study time processing in the brain, which are also interesting and elegant, instead have little help to the progress of brain-related technology? Below, I will discuss possible reasons for this discrepancy.

Generalizability: The Shortcoming

Generalizability is a measure of how useful the results of a study are for broader situations. Generalizability is the critical hypothesis (and also the aim) of science. To understand this point, let us consider a simple example. If we want to test the effectiveness of a new drug, we will recruit several patients and test the drug on them. However, our aim is not only to investigate these several recruited patients, but to draw a general conclusion on the effect of the drug on the mass of people using these recruited patients, with the hypothesis that similar phenomena can also be observed if we recruit another group of patients. As another example, when physicists perform an experiment and conclude a physical law, their aim is not only to explain the very experiment they perform, but to conclude a law generalizably applicable to other experiments taken at another place and another time. However, we should not take such generalizability for granted. Many hard problems are because we do not have a generalizable understanding of the problem or a generalizable technique to deal with the problem. For example, cancer is a challenging disease to cure because we do not have a generalizable technique to efficiently kill all the cancer cells due to the high diversity of cancer cells (Morita et al., 2020; Black & McGranahan, 2021).

Lack of generalizability is a significant shortcoming of basic timing studies. Results obtained under one experimental condition often cannot predict the result under another condition. For example, if an auditory stimulus is associated with time duration T_a and a visual stimulus with duration T_v in a rat subject, presenting the auditory and visual stimuli simultaneously will make the rat subject time an expected duration of T_+, which is between T_a and T_v, but closer to T_v (Swanton & Matell, 2011; Matell & Kurti, 2014). Additionally, compared to an auditory stimulus, the association between a visual stimulus with a time duration can be better transferred to a subsequent operant response when tested in a Pavlovian-instrumental transfer procedure (Matell & Valle, 2017). These results imply that visual signals are more involved in interval timing than auditory signals. However, in a recent study on an action timing task, in which a mouse had to learn the timing of its action based on the sensory feedback caused by its own action, it was the deprivation of auditory input (not visual) that disrupted the learned action timing (Cook et al., 2022), contradicting previous understanding (Fig. 5b, upper). Furthermore, there are two frequently studied experimental paradigms of timing tasks: interval timing, in which the subject is to perceive or produce a single time interval (Wang et al., 2018; Karmarkar & Buonomano, 2007) (Fig. 1d), and beating timing, in which the subject is to perceive or produce regular beats (Gámez et al., 2019) (Fig. 1e). It has been found that the brain uses different neural substrates and mechanisms to process temporal information in these two paradigms (Gámez et al., 2019; Wang et al., 2018; Karmarkar & Buonomano, 2007; Teki et al., 2011), even though they are both pure timing tasks with no other information (such as spatial information) involved (Fig. 5b, lower). Overall, the lack of generalizability in basic timing studies makes it challenging to conclude how the brain processes temporal information.Below, I will explore two possible reasons for the lack of generalizability in basic timing studies (Fig. 5d):

1. Laboratory Experiments.

   Basic timing studies are performed in laboratory experiments, which have artificially designed and well-controlled experimental conditions (just like Fig. 1d, e) that may not reflect real-life situations. The lack of generalizability has long been recognized as the shortcoming of laboratory experiments in social science, including psychology (Brüggemann & Bizer, 2016; Hulstijn, 1997). Therefore, the limitations of basic timing studies discussed here are just examples of the general shortcoming of the laboratory experiment paradigm. Perhaps the only way to improve the generalizability of the results of laboratory experiments is to capture the common results of different experimental conditions through a literature review. For example, Bueti and Buonomano (2014) concluded that temporal learning transfers across different modalities, including visual and auditory modalities, different auditory pitches, and slightly different lengths of temporal intervals, by reviewing papers. However, literature review cannot always lead to a straightforward conclusion, especially if the brain has a complicated performance under different experimental conditions. For example, the transfer of learning may not exist under some conditions and may be strong or weak in other conditions.

2. Misunderstanding of Simplification.

   Simplification is a pervasive idea in the data analysis and computational models of basic timing studies (Fig. 3). The pacemaker-accumulator model (Buhusi & Meck, 2005) (Fig. 3c, left), the best-known timing model, contains only four components (pacemaker, accumulator, memory device, and comparator) to model the timing process. The dynamic features found by basic timing studies (Fig. 3a, b) are often discovered after reducing the dimension of the population dynamics of neural networks using principal component analysis (PCA). This PCA method also manifests the idea of simplification: simplifying the population dynamics by reducing its dimension.

   The idea of simplification, also named the principle of Occam's razor, tries to explain the world using as few entities as possible. However, the advantage of this principle of simplification must be understood before using it. One widely accepted advantage of simplification is that simple theories tend to be more testable and, therefore, easier to falsify (Baker, 2016; Sober & Knowles, 1991). In other words, the primary advantage of a simple theory is not that it can better predict the experiment, but instead lies in its ease of falsification, which is believed to be a necessary property of a scientific theory (Popper, 1959). Another advantage of simplification (with controversy) is that it improves induction: choosing a simple theory after numerous observations reduces the chance of changing the theory after more future observations (Baker, 2016). This induction advantage is closely related to the concept of generalizability we discuss here because reducing the change in theory after future observations means improving the generalizability of the theory. However, "induction" means that the theory must be concluded after numerous observations, which is apparently not the case for the results (Fig. 3) in basic timing studies, which are usually proposed based on single laboratory experiments under simple and well-controlled situations. In other words, if we indeed want a simple timing theory that is generalizable to real-world situations, numerous observations in real-world situations are necessary.

Unknowability: The Reality

The methodology of basic cognitive studies involves recording brain activity while subjects perform deliberately designed tasks in order to understand the neural mechanisms of cognition. This methodology is based on the following philosophical understanding of science (Fig. 6a): Science investigates the world, generates knowledge, and then technology uses the knowledge generated by science to change the world. However, this philosophy fails to consider the possibility that the capability of knowledge (and therefore science) to describe the world may be fundamentally limited, meaning that some parts of the world are unknowable. If this is the case, the knowledge generated by science will not be able to well guide the design of technology to change the world effectively (Fig. 6a). The lack of generalizability discussed before may also stem from the unknowability of the world (including cognition): If the capability of knowledge to describe the world is fundamentally limited, we should not dream of the luxury that our knowledge has the possibility to generalize to every situation.

The Inspiration from AI

To discuss the limited capability of knowledge, let us start with an interesting empirical finding in AI. In AI, knowledge is usually represented by *(subject; relation; object)* triplets, representing the relationship between a subject and an object (Hogan et al., 2021). For example, the sentence "dog is animal" can be represented by a triplet *(dog; be; animal)*. A collection of a large number of triplets is called a knowledge base. It has been found that adding knowledge bases to deep learning models can improve the performance of natural language processing (Guo et al., 2022; Annervaz et al., 2018). However, interestingly, well-known colossal models (such as GPT-3 (Brown et al., 2020) of Microsoft or Pangu (Zeng et al., 2021) of Huawei) are pure deep neural networks without a knowledge base. A possible explanation for why well-known colossal models do not contain a knowledge base is that the performance improvement after adding a knowledge base to colossal models is marginal (below 4%, see Table 5 of Colon-Hernandez et al., 2021) (Fig. 6b). I discussed this interesting phenomenon with an AI expert in NetEase, who believed that this is because colossal models are trained by oceans of texts collected from the Internet, which contain far richer information than knowledge bases can provide, so adding knowledge bases to colossal models can hardly increase the information used to train the colossal models. Notice that people have invested great efforts to develop knowledge bases: Well-known knowledge bases such as YAGO and Freebase contain over 1 billion triplets. Despite such efforts, these knowledge bases are still hardly useful in the core AI technology of colossal models.

What can we learn from this empirical finding in AI? Notice that science is a process of generating knowledge from experiments (Fig. 6a): For example, basic timing studies aim to establish the relationship between the dynamics of the brain and the behavioral task. Also, notice that AI represents the future of technology. Therefore, if knowledge bases cannot help AI, we may conclude that science will not help technology in the future!

The Inspiration from Philosophy and Physics

The recognition of the limited capability of knowledge has a long history in philosophy. David Hume believed that causality cannot be justified because we can only observe that one thing, *A*, happened after another thing, *B*, but we cannot observe the underlying causal mechanism that made *A* happen after *B* (David Hume, https://en.wikipedia.org/wiki/David_Hume). Immanuel Kant believed that there exist things (the so-called things-in-themselves) that are unperceivable and unknowable. What we can perceive are mere "appearances" of these unknowable things, and a theory of the world develops when the per-

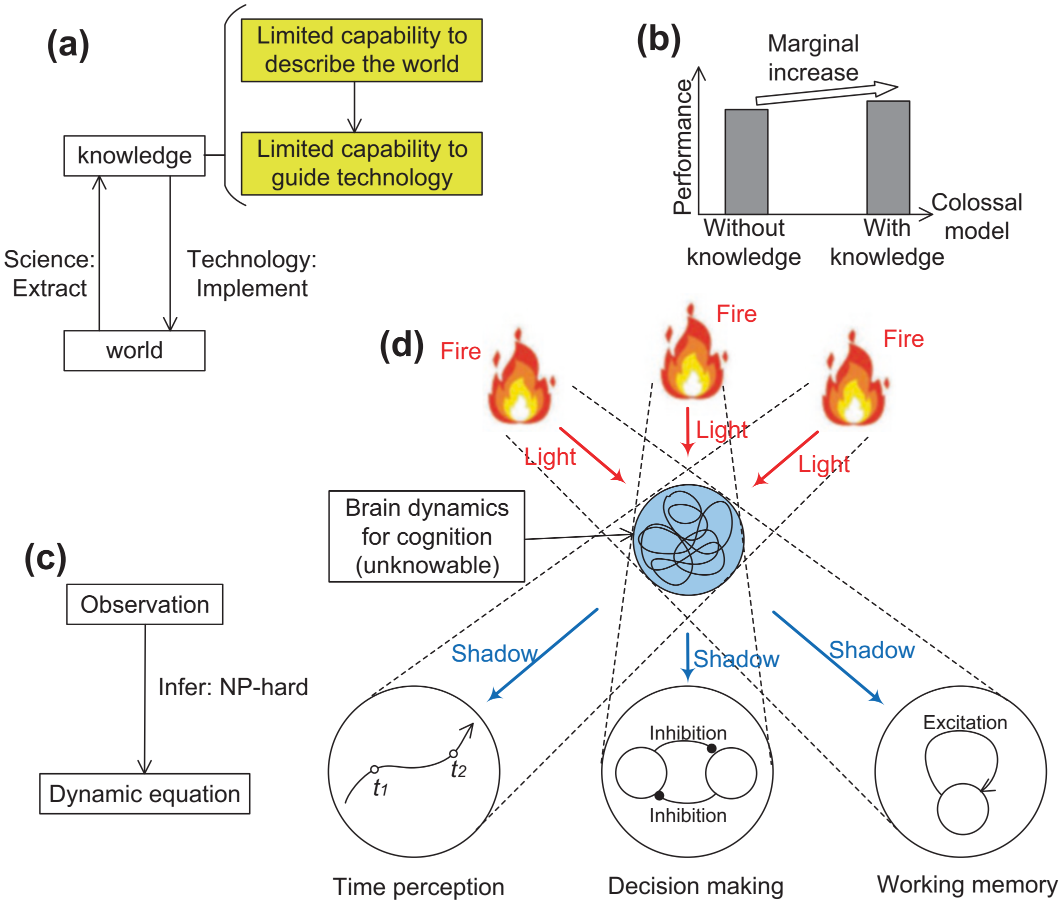


Fig. 6 The unknowability of the world. (**a**) Science extracts knowledge from the world, and technology implements knowledge to change the world. However, if knowledge has limited capability to describe the world, it will also have limited capability to guide the creation of technology (yellow boxes). (**b**) Adding knowledge to colossal models only marginally improves their performance. (**c**) Inferring the dynamic equation from experimental observation is an NP-hard problem. (**d**) A Platonic understanding of cognition. The dynamics of the brain in various tasks (such as the stereotypical trajectory for time perception (Bi & Zhou, 2020a; Karmarkar & Buonomano, 2007), competing dynamics by mutual inhibition for decision-making (Wong & Wang, 2006), and self-excitation for working memory (Lim & Goldman, 2013)) are various shadows of an unknowable object (the brain dynamics for cognition) under fires at different positions. This Platonic viewpoint implies that we will still not understand cognition after studying the dynamics in various tasks, unlike the atomic viewpoint (Fig. 1a), which believes that we will understand cognition after studying each element of cognition

ceived things conform to our spatial and temporal forms of intuition (Immanuel Kant, https://en.wikipedia.org/wiki/Immanuel_Kant). In 1963, Frederic Fitch proposed a logic paradox that asserts that if all truths were knowable, it would follow that all truths are already known (Fitch's paradox of knowability, https://en.wikipedia.org/wiki/Fitch%27s_paradox_of_knowability). Therefore, if we acknowledge that not all truths are already known, we have to acknowledge that not all truths are knowable. Fitch's paradox sets up a fundamental limitation on the capability of experiments: There exists truth that cannot be known using experiments, no matter how advanced the techniques we use.

A recent study in the field of physics provides further evidence for the notion of unknowability. The study demonstrates that identifying the underlying dynamical equation, or physical reality, from any amount of experimental observations is provably NP-hard for both classical and quantum mechanical systems (Cubitt et al., 2012a, b) (Fig. 6c). In simpler terms, if NP/ = P, which is the prevailing belief among computer scientists, identifying the underlying dynamical equation will require an exponentially long amount of time relative to the dimension of the system. Therefore, the dynamical equation would effectively be unknowable if the system has a large dimension.

There have been extensive studies on NP problems using models, such as spin glass, derived from statistical physics (Mézard & Montanari, 2009), which provide insight into the nature of the computational difficulty in solving these problems. The main conclusion is that the computational difficulty is closely related to the (some kind of) correlation between degrees of freedom in the system. To understand this concept, consider a system described by a state vector $x = (x_1, x_2, \ldots, x_n)$. If the different x_is ($i = 1, 2, \ldots, n$) do not interact with each other, we can find the optimal state x_{opt} of the system with respect to a problem by sequentially optimizing each $x_{i,}$ respectively. However, if the different x_is strongly interact with each other, we may have to adjust a large number of x_is simultaneously during the optimization process, making it more challenging to find x_{opt}.

Basic cognitive studies (Fig. 1) aim to understand the dynamics of the brain underlying cognition by observing the brain's activities when the brain is performing simple tasks. Therefore, basic cognitive studies address the same type of NP-hard problem studied in (Cubitt et al., 2012a, b) that infers dynamics from observation. We have mentioned in the last paragraph that the difficulty of this problem lies in the correlation between different degrees of freedom. Therefore, the atomic philosophy (Fig. 1a), which aims to understand cognition by studying each individual cognitive element (such as perception, memory, and decision-making), is actually unsuitable for guiding cognition research. This is because the coordination between different cognitive elements is vital for performing real-life tasks, so it is important to consider the whole task simultaneously. We have mentioned a good example before (Fig. 4c): The translation of brain activity to language for speech prosthesis achieves better performance when training the neural network to translate one sentence at a time instead of one individual word at a time (Makin et al., 2020; Cogan, 2020; Moses et al., 2019).

Unfortunately, atomism is just the very philosophy that guides basic cognitive studies (including basic timing studies), which is possibly the reason for the difficulty we encounter in understanding cognition. Despite decades of research, we still do not have a complete understanding of how the brain processes time. Results from basic timing studies can sometimes contradict each other (Fig. 5b) and cannot provide guidance for the design of technology. The study of the hippocampus is another example of this issue. While it has been found that the hippocampus transfers memory into the cortex (Goto et al., 2021) and performs inferential reasoning (Barron et al., 2020), the hippocampus encodes place (Sosa & Giocomo, 2021), head directions (Sosa & Giocomo, 2021), time (Eichenbaum, 2014), visual and auditory stimuli (Goto et al., 2021; Turk-Browne, 2019), and abstract knowledge (Nieh et al., 2021), we still do not have a clear understanding of its functional role. In other words, we cannot predict the hippocampus' functional role in a new experimental condition. What is the mechanism of the brain to process time? What is the functional role of the hippocampus? Perhaps, they are essentially unknowable.

How can we make sense of the kaleidoscopic observations in timing and hippocampal studies? In his famous allegory of the cave, Plato likens our understanding of the world to the shadows on the wall of a cave, cast by objects in front of a fire (Allegory of the cave, https://en.wikipedia.org/wiki/Allegory_of_the_cave). Inspired by this allegory, I think the best way to understand the observations in timing or hippocampal studies is to regard the brain dynamics in different experimental conditions as the shadows cast by an object from fires at different positions (Fig. 6d).

The object represents the reality of the neural mechanism of cognition, which is unknowable, but what we can observe are only the dynamics of the brain when performing a specific task. When the fire is at different positions, the projection on the wall is different, just like the kaleidoscopic dynamics of the brain when performing different tasks. This Platonic viewpoint suggests that we may still not understand cognition after studying the dynamics in various tasks, in contrast to the atomic viewpoint (Fig. 1a), which posits that we will understand cognition after studying each element of cognition. Plato encouraged us to walk out of the cave and know the reality of the world through reason. However, inferring the reality from observation is an NP-hard problem (Cubitt et al., 2012a, b), so the reality may be essentially unknowable.

Summary

The results of basic cognitive studies have limited applicability to the development of brain-related technology due to their lack of generalizability. This lack of generalizability may be attributed to the fundamental unknowability of cognition.

Outlook

What can we learn from the understandings above to guide our future research? I give three suggestions, explained in three subsections below (Fig. 7).

Improving Generalizability

As previously mentioned, generalizability is a central aim of science. We want our results to be valid in broader conditions, not just in the specific experimental conditions we investigated (Fig. 7a). Basic cognitive studies (Fig. 1b–e) are typically performed in laboratory experiments, where the experimental conditions are artificially designed and well-controlled, rather than in real-life settings. As previously noted, the lack of generalizability has long been recognized as a shortcoming of laboratory experiments (Brüggemann & Bizer, 2016; Hulstijn, 1997). Therefore, one possible way to improve the generalizability of our results is to extract the features of brain dynamics when subjects are performing real-life tasks, rather than tasks deliberately designed for experiments. Additionally, to

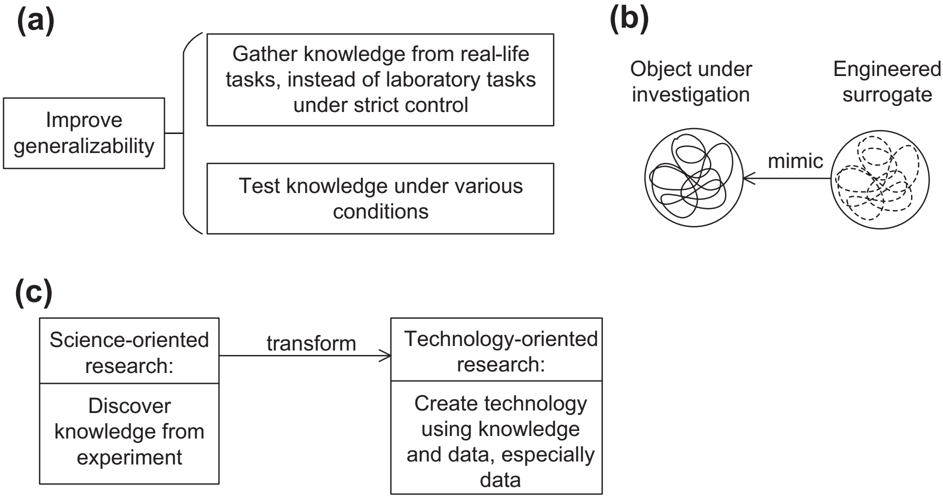


Fig. 7 My suggestions for future research. (**a**) Improve generalizability. (**b**) Engineer an easily accessible surrogate to mimic the object under investigation, so that we can predict the behavior of the object by investigating the surrogate without experimenting on the object. (**c**) Transform our research style from science-oriented, which aims to discover knowledge, to technology-oriented, which aims to create technology

improve the generalizability of a result, it is necessary to verify the result under various conditions.

Making Use of Surrogate

The concept of unknowability suggests that it may be impossible to develop a universal theory that is applicable to every situation. In light of this, how should we proceed with research? One emerging methodology to address the issue of unknowability is the use of surrogates. Instead of attempting to understand the underlying mechanism of a complex system, we create a surrogate of the system that behaves similarly to the original system in situations of interest (Fig. 7b). By observing the behavior of the surrogate, we can predict the behavior of the original system in a new situation. Surrogates are often more accessible than the original system, and this approach can be implemented even when we do not fully understand the mechanism underlying the original system.

Surrogating is the exact idea of neural network models. After observing the input signal I_i ($i = 1, 2, 3, \ldots$), and the output O_i of a system in response to I_i, a deep neural network model can be constructed by training the network to produce the output O_i given the input I_i. The resulting deep network serves as a surrogate for the original system and can predict the output of the system when given a new input signal, I^j, as long as I^j is not significantly different from the set of observed inputs $\{I_i\}_i$. In this approach, the response mechanism of the original system is not explicitly studied; instead, this understanding is encoded in the parameters of the trained neural network. While this knowledge may be difficult to interpret, it can still be used effectively. This type of knowledge is known as “dark knowledge” (Jia, 2019; Hinton et al., 2015), which stands in contrast to “light knowledge” that can be expressed through language and formulas.

Surrogates have been extensively used in brain research. It has been discovered that artificial neural networks, after being trained on a task, exhibit similar dynamics to the brain when performing the same task (Hong et al., 2016; Goldstein et al., 2022; Mante et al., 2013). Therefore, artificial neural networks can be used as surrogates to study the brain, as has been done in the study of timing tasks (Bi & Zhou, 2020a). Additionally, as previously mentioned, neural network models have been utilized as surrogates for epileptic brains to guide clinical surgery (Cao et al., 2022; Sinha et al., 2017).

Furthermore, the concept of surrogating has been applied in fields beyond brain research. For example, self-organized 3-dimensional tissue cultures derived from stem cells, known as organoids, have been used to model various organs, personalize disease treatment, and develop new drugs (Chiaradia & Lancaster, 2020; Kim et al., 2020). Another example is digital twins, which are virtual models designed to accurately represent physical objects and updated in real time with collected data. Digital twins are used to design, manufacture, monitor, and diagnose large equipment such as bridges, aircraft, and power generators (Liu et al., 2021). These examples demonstrate the use of surrogates to investigate and manipulate an easier-to-understand system in order to study the original system, even though the surrogate may also be too complex to be fully understood (e.g., the “dark knowledge” found in artificial neural networks).

Being Technology-Oriented

Science is the process of exploring new knowledge through observation and experiments. Technology is the process of applying scientific knowledge for various purposes. However, the fundamental unknowability of the world presents a limitation on the capability of science to understand the world and guide technology (Fig. 6a). Therefore, in my opinion, future studies should be technology-oriented (Fig. 7c), which has the following two meanings:

1. Instead of being driven by the interest in how nature works, scientists should perform their research with practical applications in mind. A blueprint or at least a rough sketch of how

their findings could be applied to solve practical problems would be beneficial. Without the guidance of technology, scientific results may not be useful in guiding practical applications, as demonstrated by basic timing studies' limited impact on neuroengineering and brain-inspired AI.

2. Technology tends to be created without the guidance of scientific knowledge. There is a growing trend to create technology through human-guided self-organization rather than through the implementation of knowledge via rational design. This shift away from rational design may be due to the unknowability of the world, which renders knowledge increasingly useless in dealing with complex problems. Self-organization is a process in which collective order arises from local interactions between parts of an initially disordered system (Self-organization, https://en.wikipedia.org/wiki/Self-organization; Spontaneous order, https://en.wikipedia.org/wiki/Spontaneous_order). The training of deep artificial neural networks is a self-organization process under human guidance: We adjust the interactions between artificial neurons by adjusting the synaptic weights, rather than directly designing the activity of each neuron, but the collective dynamics of the neural network when performing tasks emerge from these interactions. A good example of the paradigm shift of technology from rational design to human-guided self-organization is natural language processing. The traditional method of translating one language to another was to recognize the grammatical structure of an input sentence and then translate the sentence based on this structure using human-designed rules (Cambria & White, 2014). Today, however, the language translation is based on end-to-end training of neural networks, with the grammatical structure and translation rules automatically and implicitly emerging during training (Goldberg, 2017). As previously mentioned, this automatic and implicit feature extraction by neural networks has also been used to recognize pathological biomarkers in closed-loop deep brain stimulation (Scangos et al., 2021a, b) and translate brain activity into natural language (Makin et al., 2020; Cogan, 2020) (Fig. 4).

How can we guide the self-organization of a complex system to create technology? The current dominating methodology, deep learning, involves adjusting the synaptic weights of a deep network by gradient-based algorithms while fixing the network architecture at the form preassigned by humans (Goodfellow et al., 2016). However, evolutionary algorithms have the potential advantage of allowing for the adjustment of both synaptic weights and network architecture, without requiring human design input (Stanley et al., 2019). In a neural network created by evolutionary algorithms, everything emerges from self-organization, minimizing the interference of human rational design, whose capability is limited due to the unknowability of the world, potentially leading to superior technology (Stanley & Lehman, 2015). Furthermore, human-guided evolution is not only an algorithm that runs on computers but also a practice in laboratories. We create high-yield plants and animals by selective breeding (Selective breeding, https://en.wikipedia.org/wiki/Selective_breeding), and we also discover drugs and functional proteins by directing the evolution of engineered microbes (Davis et al., 2017; Romero & Arnold, 2009). Human-guided evolution, without the need for rational design, may be the ultimate method to create something to our desired end in this unknowable world.

Conclusion

In this paper, I review the main results of basic timing studies and highlight their limited applicability in solving practical problems in the fields of neuroengineering and brain-inspired AI. Basic timing studies extract knowledge from deliberately designed simple tasks, whereas neuroengineering is mainly driven by clinical data and AI is driven by training colossal models using oceans of data collected from the Internet. The limitation of basic timing studies may be due to the lack of generalizability of their results, which stems from the fundamental unknowability of the world, including cognition. The reason for this limitation is also true for, more generally, basic cognitive studies. As a result, I question and criticize the usefulness and prospect of the research protocol of basic cognitive studies (Fig. 1), which involves recording brain activity when the subject is performing deliberately designed experiments to understand the neural mechanism of cognition. I then suggest three ways to guide future research: improving the generalizability of results, considering using surrogates to overcome the unknowability, and performing technology-oriented studies.

The neuroscience problem identified in this paper is part of a larger trend in biology where mass-scale technology, such as multi-omic databases and supercomputing power, is increasingly being used to solve practical problems with AI (Subramanian et al., 2020). The knowledge necessary for AI to solve these problems is not implemented by humans through rational design, but instead emerges self-organizedly during the problem-solving process in a hidden manner. This knowledge is encoded in the AI system, such as in synaptic weights, but is unknowable by humans. We can imagine that in the far future, when AI becomes far more powerful than human intelligence, we may feel hard to understand the logic behind AI's problem-solving even if AI tries to explain it to us. Therefore, the use of hidden knowledge, something we can use but not understand, should be a gradually dominating paradigm in scientific and technological research.

Acknowledgments Z. B. was supported by the National Natural Science Foundation of China (32000694).

References

Allegory of the cave. https://en.wikipedia.org/wiki/Allegory_of_the_cave

Allman, M. J., Teki, S., Griffiths, T. D., & Meck, W. H. (2014). Properties of the internal clock: First- and second-order principles of subjective time. *Annual Review of Psychology, 65*, 743–771.

Angrick, M., et al. (2019). Speech synthesis from ECoG using densely connected 3D convolutional neural networks. *Journal of Neural Engineering, 16*, 036019.

Annervaz, K. M., Chowdhury, S. B. R., & Dukkipati, A. (2018). Learning beyond datasets: Knowledge graph augmented neural networks for natural language processing. *arXiv*, 1802.05930.

Baker, A. (2016). Simplicity. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Stanford University.

Baldassi, C., Braunstein, A., Brunel, N., & Zecchina, R. (2007). Efficient supervised learning in networks with binary synapses. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 11079–11084.

Baldassi, C., Ingrosso, A., Lucibello, C., Saglietti, L., & Zecchina, R. (2015). Subdominant dense clusters allow for simple learning and high computational performance in neural networks with discrete synapses. *Physical Review Letters, 115*, 128101.

Baldwin, J. M. (1893). *Elements of psychology*. Macmillan and Co.

Barron, H. C., et al. (2020). Neuronal computation underlying inferential reasoning in humans and mice. *Cell, 183*, 228–243.

Benabid, A. L., et al. (1994). Acute and long-term effects of subthalamic nucleus stimulation in Parkinson's disease. *Stereotactic and Functional Neurosurgery, 62*, 76–84.

Beniaguev, D., Segev, I., & London, M. (2021). Single cortical neurons as deep artificial neural networks. *Neuron, 109*, 2727–2739.

Bertram, E. H. (2009). Temporal lobe epilepsy: Where do the seizures really begin? *Epilepsy & Behavior, 14*(Suppl 1), 32–37.

Bi, Z., & Zhou, C. (2020a). Understanding the computation of time using neural network models. *Proceedings of the National Academy of Sciences of the United States of America, 117*, 10530–10540.

Bi, Z., & Zhou, C. (2020b). Understanding the computational difficulty of a binary-weight perceptron and the advantage of input sparseness. *Journal of Physics A: Mathematical and Theoretical, 53*, 035002.

Bi, Z., Chen, G., Yang, D., Zhou, Y., & Tian, L. (2022). Evolutionary learning in the brain by heterosynaptic plasticity. *bioRxiv*, 2021.12.14.472260.

Black, J. R. M., & McGranahan, N. (2021). Genetic and non-genetic clonal diversity in cancer evolution. *Nature Reviews. Cancer, 21*, 379–392.

Bommasani, R., et al. (2021). On the opportunities and risks of foundation models. *arXiv*, 2108.07258.

Bouthour, W., et al. (2019). Biomarkers for closed-loop deep brain stimulation in Parkinson disease and beyond. *Nature Reviews. Neurology, 15*, 343–352.

Brown, T. B., et al. (2020). Language models are few-shot learners. *arXiv*, 2005.14165.

Brüggemann, J., & Bizer, K. (2016). Laboratory experiments in innovation research: A methodological overview and a review of the current literature. *Journal of Innovation and Entrepreneurship, 5*, 24.

Bueti, D., & Buonomano, D. V. (2014). Temporal perceptual learning. *Timing and Time Perception, 2*, 261–289.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews. Neuroscience, 6*, 755–765.

Cainelli, E., Mioni, G., Boniver, C., Bisiacchi, P. S., & Vecchi, M. (2019). Time perception in childhood absence epilepsy: Findings from a pilot study. *Epilepsy & Behavior, 99*, 106460.

Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational Intelligence Magazine, 9*, 48–57.

Cao, M., et al. (2022). Virtual intracranial EEG signals reconstructed from MEG with potential for epilepsy surgery. *Nature Communications, 13*, 994.

Cheyne, D. O. (2013). MEG studies of sensorimotor rhythms: A review. *Experimental Neurology, 245*, 27–39.

Chiaradia, I., & Lancaster, M. A. (2020). Brain organoids for the study of human neurobiology at the interface of in vitro and in vivo. *Nature Neuroscience, 23*, 1496–1508.

Cichon, J., & Gan, W.-B. (2015). Branch-specific dendritic Ca2+ spikes cause persistent synaptic plasticity. *Nature, 520*, 180–185.

Cogan, G. B. (2020). Translating the brain. *Nature Neuroscience, 23*, 469–472.

Colon-Hernandez, P., Havasi, C., Alonso, J., Huggins, M., & Breazeal, C. (2021). Combining pre-trained language models and structured knowledge. *arXiv*, 2101.12294.

Constantinidis, C., Franowicz, M. N., & Goldman-Rakic, P. S. (2001). Coding specificity in cortical microcircuits: A multiple-electrode analysis of primate prefrontal cortex. *The Journal of Neuroscience, 21*, 3646–3655.

Cook, J. R., et al. (2022). Secondary auditory cortex mediates a sensorimotor mechanism for action timing. *Nature Neuroscience, 25*, 330–344.

Courbariaux, M., Hubara, I., Soudry, D., El-Yaniv, R., & Bengio, Y. (2016). Binarized neural networks: Training deep neural networks with weights and activations constrained to +1 or −1. *arXiv*, 1602.02830.

Cubitt, T. S., Eisert, J., & Wolf, M. M. (2012a). Extracting dynamical equations from experimental data is NP hard. *Physical Review Letters, 108*, 120503.

Cubitt, T. S., Eisert, J., & Wolf, M. M. (2012b). The complexity of relating quantum channels to master equations. *Communications in Mathematical Physics, 310*, 383–418.

David Hume. https://en.wikipedia.org/wiki/David_Hume

Davis, A. M., Plowright, A. T., & Valeur, E. (2017). Directing evolution: The next revolution in drug discovery? *Nature Reviews. Drug Discovery, 16*, 681–698.

Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. The MIT Press.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv*, 1810.04805.

Eichenbaum, H. (2014). Time cells in the hippocampus: A new dimension for mapping memories. *Nature Reviews. Neuroscience, 15*, 732–744.

Feynman, R. P., Leighton, R. B., & Sands, M. (2011). *The Feynman lectures on physics* (Vol. 1, New Millennium ed.). Basic Books.

Fitch's paradox of knowability. https://en.wikipedia.org/wiki/Fitch%27s_paradox_of_knowability

Frenkel, C. (2021). Sparsity provides a competitive advantage. *Nature Machine Intelligence, 3*, 742–743.

Gámez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biology, 17*, e3000054.

Goldberg, Y. (2017). *Neural network methods in natural language processing*. Morgan & Claypool Publishers.

Goldstein, A., et al. (2022). Shared computational principles for language processing in humans and deep language models. *Nature Neuroscience, 25*, 369–380.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. The MIT Press.

Goto, A., et al. (2021). Stepwise synaptic plasticity events drive the early phase of memory consolidation. *Science, 374*, 857–863.

Graupner, M., & Brunel, N. (2010). Mechanisms of induction and maintenance of spike-timing dependent plasticity in biophysical synapse models. *Frontiers in Computational Neuroscience, 4*, 136.

Greyson, B., Fountain, N. B., Derr, L. L., & Broshek, D. K. (2014). Out-of-body experiences associated with seizures. *Frontiers in Human Neuroscience, 8*, 65.

Gu, B.-M., Jurkowski, A. J., Shi, Z., & Meck, W. H. (2016). Bayesian optimization of interval timing and biases in temporal memory as a function of temporal context, feedback, and dopamine levels in young, aged and Parkinson's disease patients. *Timing and Time Perception, 4*, 315–342.

Guo, Q., et al. (2022). A survey on knowledge graph-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering, 34*, 3549–3568.

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron, 95*, 245–258.

Herculano-Houzel, S. (2012). The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 10661–10668.

Hetling, J. R. (2008). Comment on 'what is neural engineering?'. *Journal of Neural Engineering, 5*, 360.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews. Neuroscience, 13*, 135–145.

Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the knowledge in a neural network. *arXiv*, 1503.02531.

Hogan, A., et al. (2021). Knowledge graphs. *arXiv*, 2003.02320.

Hong, H., Yamins, D. L. K., Majaj, N. J., & DiCarlo, J. J. (2016). Explicit information for category orthogonal object properties increases along the ventral stream. *Nature Neuroscience, 19*, 613–622.

Hulstijn, J. H. (1997). Second language acquisition research in the laboratory: Possibilities and limitations. *Studies in Second Language Acquisition, 19*, 131–143.

Immanuel Kant. https://en.wikipedia.org/wiki/Immanuel_Kant

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences, 12*, 1606–1609.

Janata, P., & Grafton, S. T. (2003). Swinging in the brain: Shared neural substrates for behaviors related to sequencing and music. *Nature Neuroscience, 6*, 682–687.

Ji, D., & Wilson, M. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nature Neuroscience, 10*, 100–107.

Jia, W. W. (2019). *Dark knowledge: How machine cognition subverts business and society*. CITIC Press Group.

Jin, D. Z., Fujii, N., & Graybiel, A. M. (2009). Neural representation of time in cortico-basal ganglia circuits. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 19156–19161.

Karmarkar, U. R., & Buonomano, D. V. (2007). Timing in the absence of clocks: Encoding time in neural network states. *Neuron, 53*, 427–438.

Kim, J., Koo, B.-K., & Knoblich, J. A. (2020). Human organoids: Model systems for human biology and medicine. *Nature Reviews. Molecular Cell Biology, 21*, 571–584.

Kirkpatrick, J., et al. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences of the United States of America, 114*, 3521–3526.

Krauss, J. K., et al. (2021). Technology of deep brain stimulation: Current status and future directions. *Nature Reviews. Neurology, 17*, 75–87.

Kuncel, A. M., et al. (2006). Clinical response to varying the stimulus parameters in deep brain stimulation for essential tremor. *Movement Disorders, 21*, 1920–1928.

Lechner, M., et al. (2020). Neural circuit policies enabling auditable autonomy. *Nature Machine Intelligence, 2*, 642–652.

Lillicrap, T. P., Santoro, A., Marris, L., Akerman, C. J., & Hinton, G. (2020). Backpropagation and the brain. *Nature Reviews. Neuroscience, 21*, 335–346.

Lim, S., & Goldman, M. S. (2013). Balanced cortical microcircuitry for maintaining information in working memory. *Nature Neuroscience, 16*, 1306–1314.

Limousin, P., et al. (1995). Effect on parkinsonian signs and symptoms of bilateral subthalamic nucleus stimulation. *Lancet, 62*, 91–95.

Liu, M., Fang, S., Dong, H., & Xu, C. (2021). Review of digital twin about concepts, technologies, and industrial applications. *Journal of Manufacturing Systems, 58*, 346–361.

Makin, J. G., Moses, D. A., & Chang, E. F. (2020). Machine translation of cortical activity to text with an encoder- decoder framework. *Nature Neuroscience, 23*, 575–582.

Manning, C. D., Clark, K., Hewitt, J., Khandelwal, U., & Levy, O. (2020). Emergent linguistic structure in artificial neural networks trained by self-supervision. *Proceedings of the National Academy of Sciences of the United States of America, 117*, 30046–30054.

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature, 503*, 78–84.

Matell, M. S., & Kurti, A. N. (2014). Reinforcement probability modulates temporal memory selection and integration processes. *Acta Psychologica, 147*, 80–91.

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cognitive Brain Research, 21*, 139–170.

Matell, M. S., & Valle, R. B. D. (2017). Temporal specificity in Pavlovian-to-instrumental transfer. *Learning & Memory, 25*, 8–20.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience, 36*, 313–336.

Mézard, M., & Montanari, A. (2009). *Information, physics, and computation*. Oxford University Press.

Mita, A., Mushiake, H., Shima, K., Matsuzaka, Y., & Tanji, J. (2009). Interval time coding by neurons in the presupplementary and supplementary motor areas. *Nature Neuroscience, 12*, 502–507.

Mnih, V., et al. (2015). Human-level control through deep reinforcement learning. *Nature, 518*, 529–533.

Morita, K., et al. (2020). Clonal evolution of acute myeloid leukemia revealed by high-throughput single-cell genomics. *Nature Communications, 11*, 5327.

Moses, D. A., Leonard, M. K., Makin, J. G., & Chang, E. F. (2019). Real-time decoding of question-and-answer speech dialogue using human cortical activity. *Nature Communications, 10*, 3096.

Nieh, E. H., et al. (2021). Geometry of abstract learned knowledge in the hippocampus. *Nature, 595*, 80–84.

Nielsen, M. A., & Chuang, I. L. (2011). *Quantum computation and quantum information*. Cambridge University Press.

O'Connor, D. H., Wittenberg, G. M., & Wang, S. S.-H. (2005). Graded bidirectional synaptic plasticity is composed of switch-like unitary events. *Proceedings of the National Academy of Sciences of the United States of America, 102*, 9679–9684.

Okun, M. S. (2012). Deep-brain stimulation for Parkinson's disease. *The New England Journal of Medicine, 367*, 1529–1538.

Oswal, A., Brown, P., & Litvak, V. (2013). Synchronized neural oscillations and the pathophysiology of Parkinson's disease. *Current Opinion in Neurology, 26*, 662–670.

Pasley, B. N., et al. (2012). Reconstructing speech from human auditory cortex. *PLoS Biology, 10*, e1001251.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature Neuroscience, 6*, 674–681.

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

Poewe, W., et al. (2017). Parkinson disease. *Nature Reviews. Disease Primers, 3*, 17013.

Poo, M., et al. (2016). China brain project: Basic neuroscience, brain diseases, and brain-inspired computing. *Neuron, 92*, 591–596.

Pool, J. L. (1954). Psychosurgery in older people. *Journal of the American Geriatrics Society, 2*, 456–466.

Popper, K. (1959). *The logic of scientific discovery*. Hutchinson.

Rakitin, B. C., Gibbon, J., Penney, T. B., & Malapani, C. (1998). Scalar expectancy theory and peak-interval timing in humans. *Journal of Experimental Psychology. Animal Behavior Processes, 24*, 15–33.

Ramesh, A., et al. (2021). Zero-shot text-to-image generation. *arXiv*, 2102.12092.

Rives, A., et al. (2021). Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences. *Proceedings of the National Academy of Sciences of the United States of America, 118*, e2016239118.

Rizzone, M., et al. (2001). Deep brain stimulation of the subthalamic nucleus in Parkinson's disease: Effects of variation in stimulation parameters. *Journal of Neurology, Neurosurgery, and Psychiatry, 71*, 215–219.

Roitman, J. D., & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *The Journal of Neuroscience, 22*, 9475–9489.

Romero, P., & Arnold, F. (2009). Exploring protein fitness landscapes by directed evolution. *Nature Reviews. Molecular Cell Biology, 16*, 866–876.

Salinas, E., & Sejnowski, T. J. (2001). Gain modulation in the central nervous system: Where behavior, neurophysiology, and computation meet. *The Neuroscientist, 7*, 430–440.

Salinas, E., & Thier, P. (2000). Gain modulation: A major computational principle of the central nervous system. *Neuron, 27*, 15–21.

Scangos, K. W., et al. (2021a). Closed-loop neuromodulation in an individual with treatment-resistant depression. *Nature Medicine, 27*, 1696–1700.

Scangos, K. W., Makhoul, G. S., Sugrue, L. P., Chang, E. F., & Krystal, A. D. (2021b). State-dependent responses to intracranial brain stimulation in a patient with depression. *Nature Medicine, 27*, 229–231.

Schwaller, P., et al. (2021). Mapping the space of chemical reactions using attention-based neural networks. *Nature Machine Intelligence, 3*, 144–152.

Selective breeding. https://en.wikipedia.org/wiki/Selective_breeding

Self-organization. https://en.wikipedia.org/wiki/Self-organization

Shah, S. A., Tinkhauser, G., Chen, C. C., Little, S., & Brown, P. (2018). Parkinsonian tremor detection from subthalamic nucleus local field potentials for closed-loop deep brain stimulation. *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2320–2324.

Shuler, M. G., & Bear, M. F. (2006). Reward timing in the primary visual cortex. *Science, 311*, 1606–1609.

Siegfried, J., & Lippitz, B. (1994). Bilateral chronic electrostimulation of ventroposterolateral pallidum: A new therapeutic approach for alleviating all parkinsonian symptoms. *Neurosurgery, 35*, 1126–1129.

Sinha, N., et al. (2017). Predicting neurosurgical outcomes in focal epilepsy patients using computational modelling. *Brain, 140*, 319–332.

Sober, E., & Knowles, D. (1991). *Let's Razor Ockham's Razor. Royal Institute of Philosophy Supplements*. Cambridge University Press.

Sosa, M., & Giocomo, L. M. (2021). Navigating for reward. *Nature Reviews. Neuroscience, 22*, 472–487.

Spontaneous order. https://en.wikipedia.org/wiki/Spontaneous_order

Stanley, K. O., & Lehman, J. (2015). *Why greatness cannot be planned: The myth of the objective*. Springer.

Stanley, K. O., Clune, J., Lehman, J., & Miikkulainen, R. (2019). Designing neural networks through neuroevolution. *Nature Machine Intelligence, 1*, 24–35.

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in nlp. *arXiv*, 1906.02243.

Subramanian, I., Verma, S., Kumar, S., Jere, A., & Anamika, K. (2020). Multi-omics data integration, interpretation, and its application. *Bioinformatics and Biology Insights, 14*, 1177932219899051.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. The MIT Press.

Swanton, D. N., & Matell, M. S. (2011). Stimulus compounding in interval timing: The modality-duration relationship of the anchor durations results in qualitatively different response patterns to the compound cue. *Journal of Experimental Psychology. Animal Behavior Processes, 37*, 94–107.

Tan, H., et al. (2019). Decoding voluntary movements and postural tremor based on thalamic LFPs as a basis for closed-loop stimulation for essential tremor. *Brain Stimulation, 12*, 858–867.

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration based and beat-based auditory timing. *The Journal of Neuroscience, 31*, 3805–3812.

Turk-Browne, N. B. (2019). The hippocampus as a visual area organized by space and time: A spatiotemporal similarity hypothesis. *Vision Research, 165*, 123–130.

Vaswani, A., et al. (2017). Attention is all you need. *arXiv*, 1706.03762.

Wang, J., Narain, D., Hosseini, E. A., & Jazayeri, M. (2018). Flexible timing by temporal scaling of cortical responses. *Nature Neuroscience, 21*, 102–110.

Wong, K.-F., & Wang, X.-J. (2006). A recurrent network mechanism of time integration in perceptual decisions. *The Journal of Neuroscience, 26*, 1314–1328.

Wu, X., et al. (2019). Altered intrinsic brain activity associated with outcome in frontal lobe epilepsy. *Scientific Reports, 9*, 8989.

Zeki, M., & Balci, F. (2019). A simplified model of communication between time cells: Accounting for the linearly increasing timing imprecision. *Frontiers in Computational Neuroscience, 12*, 111.

Zeng, G., Chen, Y., Cui, B., & Yu, S. (2019). Continual learning of context-dependent processing in neural networks. *Nature Machine Intelligence, 1*, 369–372.

Zeng, W., et al. (2021). Pangu-α: Large-scale autoregressive pretrained Chinese language models with auto-parallel computation. *arXiv*, 2104.12369.

Part III

Timing in Humans

The Motor of Time: Coupling Action to Temporally Predictable Events Heightens Perception

Jennifer T. Coull, Inga Korolczuk,
and Benjamin Morillon

Abstract

Timing and motor function share neural circuits and dynamics, which underpin their close and synergistic relationship. For instance, the temporal predictability of a sensory event optimizes motor responses to that event. Knowing when an event is likely to occur lowers response thresholds, leading to faster and more efficient motor behavior though in situations of response conflict can induce impulsive and inappropriate responding. In turn, through a process of active sensing, coupling action to temporally predictable sensory input enhances perceptual processing. Action not only hones perception of the event's onset or duration, but also boosts sensory processing of its non-temporal features such as pitch or shape. The effects of temporal predictability on motor behavior and sensory processing involve motor and left parietal cortices and are mediated by changes in delta and beta oscillations in motor areas of the brain.

Keywords

Temporal predictions · Active sensing · Temporal orienting · Impulsivity · Motor rhythm · Delta

Introduction

Estimating the duration of an event engages regions of the brain traditionally associated with motor function, such as supplementary motor area, basal ganglia, and cerebellum, even when the temporal estimation process is purely perceptual (Wiener et al., 2010; Naghibi et al., 2023). This neuroanatomical overlap suggests that timing might share functional mechanisms with motor processing and that we may even acquire our sense of time through action (Coull & Droit-Volet, 2018). The inherent link between temporal processing and motor systems is exemplified by the universal and innate act of dancing to the beat of music (Mehr et al., 2019). While music is most often considered an auditory phenomenon, from an ecological and phylogenetic perspective it is tightly coupled to dance (Fitch, 2016). This uniquely human behavior involves synchronizing body movements with the musical rhythm through audio-motor interaction (Merchant et al.,

---

J. T. Coull (✉)
Centre for Research in Psychology and Neuroscience (UMR 7077), Aix-Marseille Université & CNRS, Marseille, France
e-mail: jennifer.coull@univ-amu.fr

I. Korolczuk
Department of Pathophysiology, Medical University of Lublin, Lublin, Poland

B. Morillon
Aix Marseille Université, INSERM, INS, Institut de Neurosciences des Systèmes, Marseille, France

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_11

2015; Patel & Iversen, 2014; Vuust et al., 2022; Zatorre et al., 2007). Yet not all forms of music induce dance equally. The musical quality and psychological construct associated with dance and the pleasurable wanting-to-move experience is called groove (Janata et al., 2012). The mere fact that humans spontaneously dance to specific types of musical stimuli speaks in favor of the dynamic and integrated nature of cognitive processes. In particular, it emphasizes the closed-loop nature of the action-perception cycle and exemplifies active sensing frameworks (see below; (Crapse & Sommer, 2008; Schroeder et al., 2010)). Investigation of the groove phenomenon shows that motor contributions to auditory perception involve temporal processing and arise when precise temporal expectations (or "priors") are violated by sensory evidence (Vuust et al., 2022). In the context of music, these temporal violations (or "prediction errors") are captured by the degree of syncopation—defined as the appearance of a beat on a metrically weak (unexpected) accent preceding a rest on a metrically strong (expected) accent. The experience of groove occurs during perception of recurring syncopated rhythmic patterns, with the relationship between the degree of syncopation and perceived groove being non-linear (Vuust & Witek, 2014; Zalta et al., 2024). This relation can be modeled with dynamical system approaches or as the interaction between the degree of syncopation and the precision (or metrical certainty) of the internal model of temporal structure (Vuust et al., 2022; Zalta et al., 2024).

Temporal Predictability Optimizes Motor Processing

While the temporal predictability of musical rhythms induces spontaneous movement for pleasure, the temporal predictability of sensory input can be used in a more practical way to optimize behavior. For instance, going through the same traffic light every day builds an association between duration (the length of time for which the light stays red) and action (accelerating away once the light changes color). This association creates a temporal expectation (or "prior") that allows you to predict the moment at which the light will turn green so that you can accelerate away more quickly. Experimentally, the behavioral benefits of temporal predictability have been extensively explored and documented. Studies repeatedly show that knowing when an event will happen improves both perceptual and motor processing of that event: temporally predictable events are perceived more easily and more quickly; they are better encoded into working and long-term memory and are responded to more quickly (Nobre & van Ede, 2018, 2023). Yet the motor benefits of temporal predictability are not restricted to response speed. Several studies have examined the effects of temporal preparation on motor mechanisms by manipulating the length of the interval (or "foreperiod," FP) between a warning cue and a target in a simple RT task. In the variable FP paradigm, the FP varies from one trial to another. As the conditional probability of target appearance gradually increases as the FP elapses, participants both respond more quickly (Niemi & Näätänen, 1981) and exert less force on the response button (Mattes & Ulrich, 1997; Jaśkowski & Verleger, 1993). In other words, temporal predictability speeds responses while simultaneously reducing the muscular effort needed to make these responses, suggesting that it might improve motor efficiency. Support for this hypothesis has come from neurophysiological studies of the fixed FP paradigm in which the FP is consistently short in one block but consistently long in another block. Since, according to Weber's law, temporal variance increases with the length of the interval, temporal estimates of short FPs are more precise than estimates of long ones, meaning that participants are better prepared in short FP blocks. Accordingly, RTs are faster in short FP blocks than long FP ones (Niemi & Näätänen, 1981). In addition to these performance benefits, the peripheral motor units that contract the muscle of the responding hand are better synchronized in short FP blocks than long ones (Hasbroucq et al., 1995) and activation of the primary motor cortex contralateral to the response hand is lower (Tandonnet et al., 2006). A strikingly similar pattern of findings was revealed by single unit recordings of the primary motor cortex in monkeys. Over several months of training on a fixed FP task, RTs to the target got steadily faster indicating that monkeys were learning to predict the time of target onset. In parallel, neural firing in primary motor cortex became tuned to the temporal structure of the task. At the predicted time of target onset, neural firing was better synchronized and the mean firing rate in the population of recorded neurons was lower overall (Kilavik et al., 2009). The performance benefits of temporal predictability might therefore be mediated by enhanced network efficiency: better synchronization and lower firing rates.

Even when we can learn to predict the onset time of a target within a single experimental session, RTs get progressively faster, and the muscular force used to make the response gets steadily smaller, again indicating improved motor efficiency (Thomas et al., 2019). In this study, participants performed a whole-body pointing task toward a response button that was just slightly out of reach. They responded to the presentation of a visual target that appeared after either a fixed FP in one session, or after a variable FP in another. During the pointing movement, electromyographic (EMG) recordings of the tibialis anterior in the ankle allowed changes in activation of the muscle to be tracked over the course of the session (50 trials). When the target was presented after a fixed FP, RTs to the target got steadily faster and the EMG amplitude of tibialis activation got progressively lower. By contrast, when the time of target onset could not be predicted in the variable FP session, there was no change in RT or EMG amplitude. Temporal predictability therefore optimized motor efficiency, both improving performance and reducing muscular effort. Notably, these muscular changes occurred even in muscles that were far from the primary response effector, indicating a distributed, whole-body motor effect (Thomas et al., 2019).

Temporal predictability not only optimizes motor activity related to limb movements, but also affects various types of oculomotor behavior. For instance, saccades (Dankner et al., 2017), microsaccades (Amit et al., 2019; Denison et al., 2019), and blinks (Amit et al., 2019; Abeles et al., 2020) are inhibited just prior to the presentation of temporally predictable, but not unpredictable, targets, even when no response to the target is required (Tal-Perry & Yuval-Greenberg, 2021). Moreover, microsaccade inhibition occurs whether the temporal dynamics of stimulus processing are shaped implicitly by temporal probabilities (Amit et al., 2019) or explicitly by temporal cues (Denison et al., 2019; Tal-Perry & Yuval-Greenberg, 2020). These results demonstrate how temporal predictions influence basic motor mechanisms to optimize behavior. Blinks and microsaccades during stimulus presentation impair perception, therefore inhibiting such oculomotor activity at temporally probable or relevant moments in time would improve visual discrimination of any stimuli occurring at those times. Nonetheless, oculomotor inhibition has even been demonstrated during discrimination of tactile (Badde et al., 2020) or auditory targets (Abeles et al., 2020), with greater inhibition the more temporally predictable the targets are. Since oculomotor inhibition occurs even in the absence of a visual target, it might represent a generalized supramodal marker of temporal predictions. Such covert markers of temporal prediction, occurring before the target has even appeared, could complement more overt measures of temporal prediction, such as RTs, which are measurable only *after* the target has appeared. For instance, individuals with ADHD do not show the usual RT benefits of temporal predictability and also fail to show saccade inhibition prior to the predicted time, indicating that they have difficulty processing temporal regularities in stimulus presentation over and above any concomitant changes in performance (Dankner et al., 2017).

Temporal Predictability Can Trigger Impulsive Behavior

Unfortunately, temporal predictability is not always beneficial for behavior. Although RTs to tactile targets in choice discrimination tasks are faster when targets appear after a fixed or highly probable FP (Lange & Röder, 2006; Badde et al., 2020), tactile perceptual sensitivity in a psychophysical task is impaired by the temporal regularity of target appearance (Kusnir et al., 2020). Kusnir et al. (2020) found that when brief changes in auditory or tactile intensity were presented at fixed, rather than variable, intervals, detection thresholds were better for auditory stimuli but worse for tactile stimuli. The authors attributed this dissociation to the fact that, by contrast with visual and auditory perception, our perception of touch is often determined by the force of our own motor acts. Self-generated action induces sensory suppression mechanisms that cause the predicted sensations of our own motor acts (the "efference copy") to be perceived as weaker than those imposed by external stimuli (e.g., Blakemore et al., 1999). By analogy, passive tactile stimuli occurring at predictable moments in time may be similarly subject to tactile suppression mechanisms, leading to reduced perceptual sensitivity.

Even within the visual modality, in which temporal predictability has typically been shown to benefit performance, it can be detrimental when stimulus characteristics trigger potentially competing responses. For instance, Correa et al. (2010) showed that temporal predictability exacerbates the interfering effects of incompatible stimulus-response associations in classic response conflict paradigms, such as the flanker or Simon tasks. In the flanker task, participants make left or right response buttons according to the direction indicated by a central arrow. This arrow is flanked on either side by arrows whose direction either match (compatible) or not (incompatible) the direction of the central arrow. Typically, RTs are slower for targets with incompatible flankers due to the conflicting responses induced by the directions of the task-relevant central arrow and the flanking distractors. In the Simon task, left or right button presses are associated with a specific stimulus feature, such as color (press left for blue stimuli and right for red) or shape (press left for x and right for +). Crucially, the side of the screen on which stimuli are presented either matches (compatible) or not (incompatible) the response side indicated by the target. Again, RTs are slower for incompatible targets due to the response conflict induced by the task-relevant stimulus feature and the more salient, though task-irrelevant, target location. Correa et al. (2010) adapted these classic paradigms by presenting stimuli at times that were either temporally probable (75%) or improbable (25%) within a given block, with the hypothesis that temporal predictability would mitigate the RT costs of response conflict. However, in both flanker and Simon tasks, RTs to incompatible stimuli were even slower when they appeared at expected, rather than unexpected, times (Correa et al., 2010). In other words, temporal predictability exacerbated the motor cost of response conflict (see also van Driel et al., 2015 for a complementary approach). Correa et al. (2010) suggested that temporal predictability increased motor readiness to response to both task-relevant and task-irrelevant features. As such, responses to compatible stimuli would accelerate due to the combined influence of the task-relevant and irrelevant features. By contrast, responses to incompatible stimuli would be both slower and more error-prone because the interfering influence of the task-irrelevant feature would counteract that of the task-relevant feature.

We have followed up this research in a series of studies in which temporal predictability in the Simon task was manipulated with temporal cues rather than stimulus probabilities. Specifically, a temporal cue predicted whether the subsequent target would occur after a short or a long interval whereas a neutral cue provided no predictive information, with the target equally likely to occur after the short or long delay (Fig. 1a). Since short and long interval trials were intermixed within blocks, participants could use temporal cues to voluntarily orient their focus of temporal attention to different moments in time from one trial to the next. We first confirmed Correa et al.'s (2010) findings that the RT cost of incompatible targets was greater when they were temporally predictable. Second, we demonstrated that in the fastest section of the RT distribution, when participants are responding particularly quickly and making more errors to incompatible stimuli, these errors

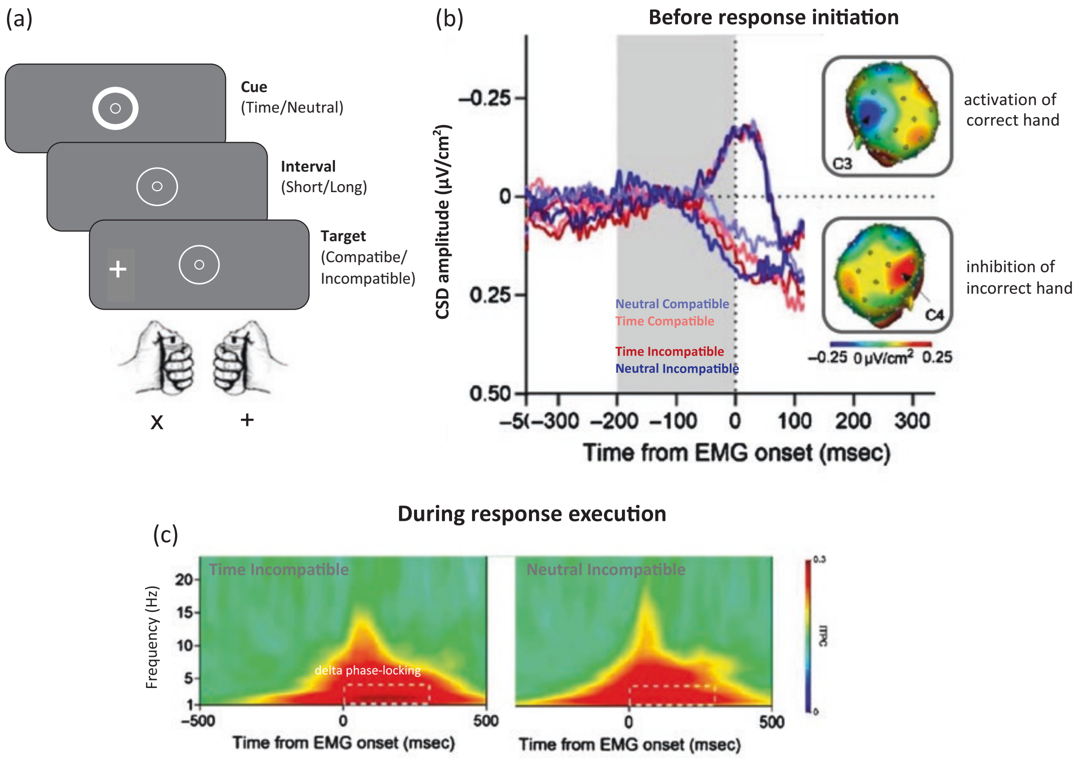


Fig. 1 Temporal predictions modulate motor cortex activity during response conflict. (**a**) Temporal Simon task: Centrally presented "Time" cues indicated whether the target would occur after a short (small inner circle) or long (large outer circle) interval. "Neutral" cues (both circles) gave no information about target onset. Target shape was associated with left/right hand responses (counterbalanced across participants). Target location could be either compatible or incompatible with the correct hand response. In this example, the + appears on the left but should be responded to with the right hand and so is an "incompatible" target type. If it had appeared on the right of the screen, it would have been a "compatible" target. (**b**) After target presentation, but before initiation of the response (gray square), there was stronger inhibition (i.e., steeper slopes) of activity measured over the primary motor cortex contralateral to the *incorrect* response hand (lower panel) for temporally predictable (compared to neutral) *compatible* targets, but less inhibition for temporally predictable (compared to neutral) *incompatible* targets. By contrast, temporal predictability had no effect on motor cortex activity contralateral to the *correct* response hand (upper panel) prior to response initiation. (**c**) During response execution, delta phase locking measured over the primary motor cortex contralateral to the correct response hand was stronger for incompatible than compatible targets (not shown) and was stronger still when incompatible targets were temporally predictable (left panel) rather than unpredictable (right panel). (*From Korolczuk et al.* (2022))

were even more frequent in the temporal cue condition than the neutral cue condition (Korolczuk et al., 2018). In other words, when the time of target appearance could be predicted in advance, it was harder to stop the automatic impulse to make a response with the hand associated to the more salient, but task-irrelevant, feature. This result indicates that temporal predictability induces a greater number of fast, impulsive errors and supports the hypothesis of increased response readiness.

Further support for this hypothesis came from EMG recordings of the response hands. EMG allowed us to detect small, sub-threshold "twitches" in the muscles of the *incorrect* response hand that occasionally occur before the response with the *correct* hand is given. These twitches, termed "partial errors," are usually caught and corrected before becoming fully fledged supra-threshold errors (Burle et al., 2002; Servant et al., 2015). Nevertheless, the number of these covert partial errors provides a measure of susceptibility to automatic response capture by the task-irrelevant stimulus feature, independent of any subsequent correction (inhibition) processes. Korolczuk et al. (2020) found that temporal cueing not only increased the number of overt impulsive errors, confirming prior results, but also increased the number of covert partial errors. Moreover, these partial errors were made more quickly when the target was temporally predictable. By contrast, temporal predictability did not affect the number of covert partial errors that were subsequently corrected. Taken together, this pattern of results indicates that participants quickly began to activate the wrong response hand when targets were temporally predictable but were still able to correct their mistake before it became an overt motor response. Very similar findings were reported by Menceloglu et al. (2021) using a fixed FP flanker task, in which participants used a computer mouse to reach toward the left or right response side rather than simply pressing a left or right button. Analysis of reaching trajectories allowed covert response tendencies to be tracked. If flankers were incompatible with the direction indicated by the central arrow, trajectories initially curved toward the side indicated by the task-irrelevant flankers before then being adjusted back toward the correct response side. Temporal predictability exacerbated the effects of response conflict on reaching dynamics: trajectories to incompatible targets were even more curved toward the competing side when participants knew when the target was going to appear. Together, these studies demonstrate how careful analysis of movement parameters (EMG or reach trajectory) helps uncover the effects of temporal predictability on motor control more completely. Online inhibition of these automatic, yet covert, response impulses would never have been revealed by simply measuring mean RT. Instead, analyses of sub-threshold EMG activity or trajectory dynamics show that temporal predictability increases the urge to initiate premature, yet potentially erroneous, prepotent responses, which can nevertheless still be corrected before the final response is delivered.

Temporal predictability doesn't only modulate motor activity during the *execution* of the motor response. We recently combined EMG with electroencephalography (EEG) to show that temporal predictability can modulate motor activity related to the resolution of response conflict before response execution has even begun. We recorded activity over left and right primary motor cortices, which were contralateral to either the correct or incorrect response hand in a temporally cued Simon task (Korolczuk et al., 2022). Simultaneous EMG recording allowed the onset of motor responses to be pinpointed with high temporal precision. We hypothesized that temporal predictability would improve RTs to compatible targets by increasing motor activation in the hemisphere contralateral to the correct hand and/or decreasing motor inhibition in the hemisphere contralateral to the incorrect hand. Results showed that, in fact, temporal predictability had no effect on activity in the motor cortex contralateral to the correct hand in this response conflict paradigm (Fig. 1b, top). Instead, it differentially modulated inhibitory activity in the motor cortex contralateral to the *incorrect* response hand as a function of target compatibility (Fig. 1b, bottom). When target features were compatible, temporal predictability increased motor inhibition in the hemisphere controlling the incorrect response hand and performance improved. Conversely, when target features were incompatible, triggering two conflicting responses, temporal predictability decreased motor inhibition in the hemisphere controlling the incorrect hand, and impulsive responding increased. This neurophysiological dissociation explains both the performance benefits of temporal predictability for non-conflict stimuli and performance costs for conflict stimuli. These inhibitory effects occurred after the target had been presented but approximately 100 ms before the EMG-defined onset of the motor response, indicating a clear effect of temporal predictability on motor planning rather than response execution.

The effects of temporal predictability on motor inhibition are further corroborated by the results of a temporally cued version of the Stop-Signal Task (Korolczuk et al., 2018). The Stop-

Signal task is a visual choice RT task in which 25% of trials contain an auditory stop signal, presented very soon after the target appears, signaling that participants should withhold their response to the target on that trial. Korolczuk et al. (2018) adapted this paradigm by including temporal or neutral cues (similar to those shown in Fig. 1a) at the beginning of each trial, which either predicted (temporal cue) or not (neutral cue) when the target would occur, i.e., after a short or long interval. As expected, temporal (versus neutral) cues speeded RTs on the 75% of "go" trials that did not contain a stop signal. By contrast, when participants had to withhold their response in the "stop" trials, temporal cues *impaired* performance: it took participants longer to stop their response to the target when the time at which it would appear was entirely predictable. The complementary pattern of results demonstrates that temporal predictability facilitates motor responding to a target in go trials but makes it harder to inhibit the response in stop trials. In other words, temporal predictability lowers the response threshold, sometimes to detrimental effect.

Rhythmic Movements Enhance Sensory Processing

Temporal predictions not only guide motor control but can also facilitate sensory processing (Nobre & van Ede, 2018, 2023). Behavioral experiments demonstrate that anticipating the temporal occurrence of an upcoming event optimizes its processing by improving the quality of visual (Cravo et al., 2013; Doherty et al., 2005; Rohenkohl et al., 2012) or auditory (Jaramillo & Zador, 2011; Morillon et al., 2016; Wollman & Morillon, 2018) information. Indeed, naturalistic acoustic signals, such as speech and music, also exhibit reliable temporal regularities that can be used to generate temporal predictions (Ding et al., 2017). While the temporal structure of environmental signals has been mostly overlooked until now, recent results crucially reveal the unique role of the motor system in the analysis of perceptual temporal dynamics. One of the contributions of the motor cortex to sensory processing would be in the analysis of the slow temporal dynamics of the perceptual stream, likely providing a contextual temporal framework to parse sensory information, thereby improving the quality of perceptual processing. This contribution of motor areas is likely modality- and domain-general and would occur during perception of any sensory signal temporally structured with low-frequency dynamics. Accordingly, coupling action to temporally predictable sensory streams, such as auditory rhythms, further enhances sensory processing. From early childhood (< 1 year), infants' movements influence their perception of auditory rhythms (Phillips-Silver & Trainor, 2005), suggesting that the contribution of the motor system to time perception is innate. Rhythmic action training helps children aged 5 years old to accurately reproduce a rhythmic sequence and, more interestingly, helps both 5- and 8-year-olds to discriminate the trained rhythm in a purely perceptual timing task (Monier et al., 2019). Moving in time to a rhythmic stimulus helps adults (Phillips-Silver & Trainor, 2007), especially non-musicians (Su & Pöppel, 2012; Zalta et al., 2020), find the beat in ambiguous rhythms and improves perceptual estimates of duration (Manning & Schutz, 2013, 2016). Even simply imagining rhythmic movement strengthens the performance benefits of the temporal predictability of a stimulus stream (Fautrelle et al., 2015). The temporal information implicitly embedded within rhythmic movements therefore helps construct an explicit, independent, and flexible representation of duration that can be used to make accurate perceptual temporal discriminations.

In a set of studies, we investigated the influence of overt movements on the quality of auditory temporal attention (Morillon & Baillet, 2017; Morillon et al., 2014). We focused on the auditory modality because of its relative bottom-up disconnection from the motor system. Indeed, while bottom-up and top-down motor influences are contingent in most sensory modalities, audition is the exception since we cannot selectively move our ears. Hence, bottom-up auditory processing is remarkably divorced from movement, whereas movement can modulate auditory processing through top-down corollary discharge (efference copy) signals (see below; Morillon et al., 2015; Schroeder et al., 2010). We measured the ability of human participants to extract relevant melodic information embedded in rhythmic streams that were interleaved with distractors, as in the cocktail-party effect (Cherry, 1953), where listeners must "tune in" to one conversation in a noisy scene. We used a perceptual decision-making task to measure the effects of temporal attention over an extended time period. Participants were asked to categorize sequences of eight pure tones as being higher or lower pitched, on average, than a reference frequency. To drive rhythmic fluctuations in attention, the tones (targets) were delivered in phase with a reference beat, presented at the beginning of each trial, and in antiphase with irrelevant yet physically indistinguishable tones (distractors). Hence, only the temporal dimension, namely an accurate internal representation of the beat, allowed participants to discriminate targets from distractors (see Fig. 2a). Participants were instructed either to use their motor system overtly, by tapping in time with the reference beat, or to keep the rhythm covertly, in both cases as a way to optimize the allocation of temporal attention (predictions) in a purely internally driven fashion. Our findings showed that the active implication of the motor system improved the precision of temporal attention and the quality of sensory processing. In other words, moving in time to a periodic stimulus improves the temporal segmentation of auditory information and improves perceptual discrimination of non-temporal features (here, pitch) of stimuli occurring on the beat. Moreover, the motor-related improvement in auditory segmentation depended parametrically on the temporal predictions made by the motor system: the more rhythmic the auditory sequence and the better able participants were to tap rhythmically, the more they benefitted from overt motor rhythms to extract auditory information (Morillon et al., 2014). Our results therefore reinforce the idea that the motor system is intrinsically engaged in the analysis of temporal sequences (see below).

In a more recent study, we developed a paradigm to behaviorally quantify the sampling capacity of periodic temporal attention during auditory and visual perception (Zalta et al., 2020). Sequences of stimuli were presented on each trial, from 2 to ~20 s. Three reference stimuli defining the tempo (or beat frequency) of the

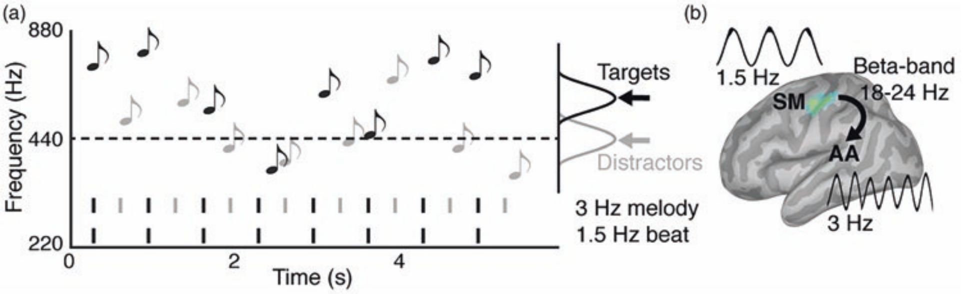


Fig. 2 Temporal predictions mediated by delta-beta coupled oscillations. (**a**) Dynamic selective attention paradigm. Quasi-rhythmic sequences of tones were presented binaurally on each trial. Four reference tones (not shown) preceded an alternation of eight target and eight distractor tones of variable frequencies. Targets occurred in phase with the preceding references, whereas distractors occurred in antiphase. Participants had to decide whether the mean frequency of targets was higher or lower than the reference frequency. While auditory input occurs at 3 Hz (melody), participants temporally modulate their attention at 1.5 Hz (beat). (**b**) In this design, neural activity dedicated to stimulus processing or temporal prediction is hence dissociable based on its temporal dynamics (3 Hz vs. 1.5 Hz, respectively). While auditory association cortex (AA) tracked stimulus dynamics (3 Hz; not shown), temporal predictions were encoded in the left sensorimotor cortex (SM) in delta-beta phase-amplitude coupled (1.5 – 18-24 Hz) oscillations, functionally directed toward auditory regions to modulate the 3 Hz auditory input. (*From Morillon and Baillet* (2017))

isochronous event sequence preceded a mixture of on-beat and off-beat stimuli. Participants performed a beat discrimination task at the end of each trial, by deciding whether the last stimulus of the sequence, a deviant, was on or off beat. While on-beat stimuli reinforced the temporal structure, crucially, off-beat stimuli had a distracting influence. This interleaved delivery of sensory events forced participants to track the beat throughout the entire duration of the sequence while minimizing the interference of aperiodic events. This protocol thus ensured that their attentional focus was temporally modulated over an extended time period. The density of distractors (i.e., number of distractors per beat) was adjusted for each participant prior to the experiment to reach threshold performance for a 2 Hz beat frequency. The quality of temporal attention was estimated for different beat frequencies, ranging from ~0.5 Hz to ~4 Hz across conditions, to span most of the range of discernible tempi (Fraisse, 1948; McAuley, 2010; Moelants, 2002; Repp & Su, 2013; Woodrow, 1951). In each modality, we first investigated temporal attention during passive perception—i.e., without overt motor involvement—and then quantified in another set of experiments the motor contribution to temporal attention. Through six interrelated behavioral experiments, we revealed the existence of a limited sampling capacity of temporal attention which, importantly, was sensory-specific (~1.4 Hz in audition, ~0.7 Hz in vision). In addition, we demonstrated that the motor contribution to temporal attention was also sensory-specific and derives from the compatibility of temporal dynamics underlying motor and sensory-specific attentional processes. Indeed, we showed that the motor contribution to temporal attention scales with motor rhythmic precision, which is maximal at ~1.5–2 Hz (Fraisse, 1948; Woodrow, 1951). Moreover, this effect crucially depends on the temporal simultaneity of motor acts relative to the beat (Morillon et al., 2014; Zalta et al., 2020). Finally, in line with previous models of beat perception and temporal attention processes (Jones, 1976; Large & Jones, 1999; Large et al., 2015), we show that our results are reproduced by a simple dynamical model involving three coupled oscillators. While the optimal sampling rate of temporal attention is directly reflected in the natural frequency of the (auditory or visual) attentional oscillator, the quality of the motor modulation of temporal attention crucially depends on the time delay in the coupling between the stimulus and the motor oscillator. These results suggest that specific rhythmic sampling rates emerge from the specific configuration of large-scale neural networks encompassing sensory regions, in addition to attentional and motor areas (Donner & Siegel, 2011; Siegel et al., 2012). Future models and experiments need to further investigate how these multiple neural structures—devoted to specific algorithmic processes and having specific temporal constraints—are dynamically coordinated to subserve or facilitate perceptual processing.

Active Sensing as a Fundamental Mechanism of Perception

Active sensing describes the fact that perception is intrinsically shaped by motor dynamics (Morillon et al., 2015; Schroeder et al., 2010). Indeed, our sensory organs are not passive receptacles for stimulation but are part of an action-perception closed-loop system (Ahissar & Assa, 2016). Sensory inputs are acquired through overt motor sampling behaviors, such as hand or eye movements in primates and whisking or sniffing in rodents. Consequently, motor acts and associated neural dynamics temporally structure the activity of sensory cortices and, consequently, the processing of incoming sensory inputs.

Attention is an essential component of the process, helping to impose a motor sampling pattern on the relevant sensory stream (visual, tactile…). In addition to driving activity in sensory areas by volleys of ascending input (through movement of the sensory organ), there is top-down (corollary discharge and/or attentional) modulation, all yoked to movement (Schroeder et al., 2010). The corollary discharge signals provided by motor areas to sensory systems convey system-specific information content (spatial, spectral, etc.) as well as contextual temporal information (i.e., the underlying temporal framework (Crapse & Sommer, 2008)). Thus, perceptual information is sampled according to the rhythms of overt action, but also the rhythms of covert motor and/or attentional dynamics. As illustrated by whisking and sniffing in rodents and saccadic sampling during free viewing in primates, motor-driven information sampling routines follow temporally periodic patterns (McAuley et al., 1999; Wachowiak, 2011).

Covert-attentional sampling is a form of active sensing that takes over when overt-motoric sampling is not possible. In the auditory domain, a form of overt active sensing is lacking (Morillon et al., 2015; Schroeder et al., 2010) but a *covert* form of active sensing is present, with oscillatory influences from motor cortex modulating activity in auditory regions during perception (Arnal, 2012; Merchant et al., 2015; Morillon & Schroeder, 2015; Morillon et al., 2015; Patel & Iversen, 2014). When perceiving several rhythmic sensory streams simultaneously (e.g., one conversation at a cocktail party), dynamic attentional filtering facilitates the processing of one task-relevant stream at the expense of all others. This occurs whether attention is selecting between modalities (Lakatos et al., 2009), within a modality (Lakatos et al., 2013), or integrating across modalities (Zion Golumbic et al., 2013). Motor system neural dynamics are implicated in covert-attentional, as well as overt-motoric, sampling of sensory input. For instance, motor and premotor cortices consistently emerge as active regions in studies of rhythmic (Morillon et al., 2015) or temporally cued (Coull & Nobre, 1998; Davranche et al., 2011; Coull et al., 2013) attentional selection. According to the premotor theory of attention, covert shifts in spatial attention are governed by the same circuitry that controls overt shifts in eye position (Rizzolatti et al., 1987). Although one can argue specific tenets of the theory (Smith & Schenk, 2012), it clearly underscores a fundamental relationship between motor systems and active/attentive sensory processing. This relationship comes to the fore when the temporal predictability of sensory input is being used to optimize behavior (traffic light, musical beat, etc.). Accordingly, it has been proposed that time perception relies on the neural recycling of action circuits and is implemented by internal, non-conscious “simulation” of movements in most ecological situations (Schubotz, 2007; Coull & Droit-Volet, 2018; Arnal, 2012; Patel & Iversen, 2014; De Kock et al., 2021; Robbe, 2023). This idea mirrors the central proposition of the premotor theory of attention in which the transmission of temporal information from motor to sensory regions could be implemented through corollary discharge signals (Crapse & Sommer, 2008; Nobre & van Ede, 2018; Khalilian-Gourtani et al., 2022).

Motor Cortex Rhythms as Intrinsic Temporal Constraints of Perception

Cortical rhythms correspond to the periodic shifting of neuronal populations between states of high and low excitability, which coordinates neural communication (Wang, 2010). In the motor cortex, most of the existing data point to beta oscillations (~12–30 Hz) as a predominant and specific rhythm during rest and to coordinate information transfer in action planning and execution. However, using finer-grained spectral analysis methods to analyze recordings from the motor cortex reveals a much more complex picture (Morillon et al., 2019). In humans for example, ongoing oscillatory activity of the primary motor cortex is characterized by consistent spectral peaks, principally in the delta (0.5–4 Hz) and beta frequency ranges (Keitel & Gross, 2016). Interestingly, these oscillatory rhythms are hierarchically structured in time (Lakatos et al., 2005). For example, a functionally relevant delta-beta phase-amplitude coupling has been shown in the motor cortex during auditory (Arnal et al., 2015; Keitel et al., 2018; Morillon & Baillet, 2017) and visual (Saleh et al., 2010) perception. This specific spectro-spatial pattern of activity represents temporal information and is directly related to behavioral performance in tasks involving perceptual temporal processing. Increases in the sensory quality of information presented at predictable moments in time (Nobre & van Ede, 2018) are reflected in the reorganization of low-frequency neural oscillations, which temporally modulates the excitability of task-relevant neural populations and thus acts as an instrument of sensory selection (Schroeder & Lakatos, 2009). This neural reorganization is visible in sensory cortices, but also in higher-order associative attentional and motor regions (Besle et al., 2011).

Multiple findings converge to support the relevance of intrinsic delta oscillations, at the functional and behavioral levels, whenever rhythmicity, i.e., temporal predictability, is observed. For instance, walking, a most basic motor act, is fundamentally rhythmic and operates within the delta range (2 Hz; (MacDougall & Moore, 2005)). Strikingly, this mirrors spontaneous oscillatory activity in the motor cortex (~2 Hz). Spontaneous rhythmic motor behaviors such as finger tapping also function at a preferred tempo of ~1.5–2 Hz, and motor tapping has an optimal temporal precision at around ~0.8–2.5 Hz (Fraisse, 1948). Moreover, during production of speech, an extremely complex motor behavior, the coordination of articulatory movements is encoded in kinematic trajectories characterized by delta oscillatory dynamics (Chartier et al., 2018). Even during reaching, a non-periodic motor behavior, motor trajectories are encoded in patterns of neural dynamics that oscillate at around 1–2 Hz (Churchland et al., 2012). Delta dynamics in motor areas also anticipate the timing of informative cues in motor planning (Saleh et al., 2010; Cadena-Valencia et al., 2018) and are tuned to temporally predictable targets requiring resolution of response conflict in the Simon task (Korolczuk et al., 2022; Fig. 1c). Delta oscillatory dynamics are also argued to be responsible for cognitive phenomena such as attentional blink, inhibition of return, and the psychological refractory period (Schroeder & Lakatos, 2009; Wyart et al., 2012). Overall, delta oscillations shape the dynamics of neural activity in the motor cortex and of motor behaviors and appear to impose a temporal framework that both enables and constrains the sampling of perceptual information. In other words, a proactive simulation process is effected by delta oscillations which constrain the speed at which temporal processing occurs.

The modulatory effects of temporal predictability on both perceptual and motor processing have consistently been shown to originate at least in part in motor cortex (Morillon & Baillet, 2017) and regions associated with motor planning, such as left inferior parietal cortex (Bolger et al, 2014; Coull, 2014; Naghibi et al., 2023). In a magnetoencephalography (MEG) experiment, we asked human participants to perform auditory categorizations of sequences of pure tunes, composed of an interleaved delivery of targets and distractors (Morillon et al., 2014). This paradigm mirrors the cocktail-party effect (Cherry, 1953), with the “noisy” signal occurring at a rate of ~3 Hz while temporal attention is guided to temporally predictable targets at ~1.5 Hz (see Fig. 2 and above). We observed that while the acoustic signal drives cortical dynamics in and around auditory regions, left-lateralized sensorimotor cortex encodes the temporal predictions that allow for precise temporal anticipation of forthcoming sensory inputs at 1.5 Hz (see also Bolger et al., 2014). Moreover, this encoding is associated with bursts of beta (18–24 Hz) neural oscillations that are directed toward auditory regions. In addition, when participants were instructed to use their motor system overtly, by pressing a button in time with the reference beat (similar to (Morillon et al., 2014)), the quality of temporal predictions increased and enhanced auditory task performance even more. These behavioral changes were associated with increased signaling of temporal predictions in right-lateralized frontoparietal associative regions, involved in melodic processing and auditory memory for pitch (Zatorre et al., 1994). This study indicates a covert form of auditory active sensing and emphasizes the fundamental role of left-lateralized motor brain areas and overt motor behavior in sensory processing. Attentional modulation of auditory information thus depends on the downward propagation of such temporal information represented in delta-beta phase-amplitude coupled oscillations in motor cortex (Arnal et al., 2015; Morillon & Baillet, 2017; Saleh et al., 2010).

Conclusion

Temporal regularities in sensory input can be automatically or voluntarily extracted and exploited to optimize both perceptual processing and motor behavior. Moreover, the influence between the motor system and timing is mutual. The temporal predictability of sensory input affects a wide variety of motor behaviors, including speed, force, and impulsivity. Reciprocally, motor behavior hones temporal precision and leads to a more robust representation of time. By boosting the precision of the temporal representation through movement, temporal predictions become more accurate and the motor and sensory processing of events occurring at predicted moments in time is heightened. Importantly, faster responses to temporally predictable events will further tighten the coupling between action and time, which could then feed back within a closed-loop system to enhance the temporal precision of predictions even more, and so continually fine-tune sensorimotor processing. While enhanced sensory processing of predictable events will generally serve to optimize behavior, it may sometimes induce maladaptive impulsive behavior by triggering responses to salient but irrelevant events. The intertwined relationship between motor activity and temporal information processing may be mediated by delta-range oscillatory activity in the motor cortex, such that oscillatory dynamics become tuned to the temporal predictability of sensory input, thereby shaping the sampling of perceptual information. More generally, this outline is directly compatible with the active sensing framework, which attributes the motor system with a fundamental role in perception.

References

Abeles, D., Amit, R., Tal-Perry, N., Carrasco, M., & Yuval-Greenberg, S. (2020). Oculomotor inhibition precedes temporally expected auditory targets. *Nature Communications, 11*(1), 3524.

Ahissar, E., & Assa, E. (2016). Perception as a closed-loop convergence process. *eLife, 5*, e12830.

Amit, R., Abeles, D., Carrasco, M., & Yuval-Greenberg, S. (2019). Oculomotor inhibition reflects temporal expectations. *NeuroImage, 184*, 279–292.

Arnal, L. H. (2012). Predicting “when” using the motor system’s beta-band oscillations. *Frontiers in Human Neuroscience, 6*, 225.

Arnal, L. H., Doelling, K. B., & Poeppel, D. (2015). Delta-beta coupled oscillations underlie temporal prediction accuracy. *Cerebral Cortex, 25*, 3077–3085.

Badde, S., Myers, C. F., Yuval-Greenberg, S., & Carrasco, M. (2020). Oculomotor freezing reflects tactile temporal expectation and aids tactile perception. *Nature Communications, 11*(1), 3341.

Besle, J., Schevon, C. A., Mehta, A. D., Lakatos, P., Goodman, R. R., McKhann, G. M., Emerson, R. G., & Schroeder, C. E. (2011). Tuning of the human neocortex to the temporal dynamics of attended events. *The Journal of Neuroscience, 31*, 3176–3185.

Blakemore, S. J., Frith, C. D., & Wolpert, D. M. (1999). Spatiotemporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience, 11*, 551–559.

Bolger, D., Coull, J. T., & Schön, D. (2014). Metrical rhythm implicitly orients attention in time as indexed by improved target detection and left inferior parietal activation. *Journal of Cognitive Neuroscience, 26*(3), 593–605. https://doi.org/10.1162/jocn_a_00511

Burle, B., Possamaı, C.-A., Vidal, F., Bonnet, M., Hasbroucq, T. (2002). Executive control in the Simon effect: an electromyographic and distributional analysis. *Psychological Research, 66*, 324–336.

Cadena-Valencia, J., Garcia-Garibay, O., Merchant, H., Jazayeri, M., & de Lafuente, V. (2018). Entrainment and maintenance of an internal metronome in supplementary motor area. *eLife, 7*, e38983.

Chartier, J., Anumanchipalli, G. K., Johnson, K., & Chang, E. F. (2018). Encoding of articulatory kinematic trajectories in human speech sensorimotor cortex. *Neuron, 98*, 1042–1054.e4.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America, 25*, 975.

Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., & Shenoy, K. V. (2012). Neural population dynamics during reaching. *Nature, 487*, 51–56.

Correa, A., Cappucci, P., Nobre, A. C., & Lupiáñez, J. (2010). The two sides of temporal orienting: facilitating perceptual selection, disrupting response selection. *Experimental Psychology, 57*(2), 142–148.

Coull, J. T. (2014). Getting the timing right: experimental protocols for investigating time with functional neuroimaging and psychopharmacology. *Advances in Experimental Medicine and Biology, 829*, 237–264.

Coull, J. T., & Droit-Volet, S. (2018). Explicit understanding of duration develops implicitly through action. *Trends in Cognitive Sciences, 22*(10), 923–937.

Coull, J. T., & 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. *The Journal of Neuroscience, 18*(18), 7426–7435.

Coull, J. T., Davranche, K., Nazarian, B., & Vidal, F. (2013). Functional anatomy of timing differs for production versus prediction of time intervals. *Neuropsychologia, 51*(2), 309–319.

Crapse, T. B., & Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nature Reviews. Neuroscience, 9*, 587–600.

Cravo, A. M., Rohenkohl, G., Wyart, V., & Nobre, A. C. (2013). Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *The Journal of Neuroscience, 33*, 4002–4010.

Dankner, Y., Shalev, L., Carrasco, M., & Yuval-Greenberg, S. (2017). Prestimulus inhibition of saccades in adults with and without attention-deficit/hyperactivity disorder as an index of temporal expectations. *Psychological Science, 28*(7), 835–850.

Davranche, K., Nazarian, B., Vidal, F., & Coull, J. (2011). Orienting attention in time activates left intraparietal sulcus for both perceptual and motor task goals. *Journal of Cognitive Neuroscience, 23*(11), 3318–3330.

De Kock, R., Gladhill, K. A., Ali, M. N., Joiner, W. M., & Wiener, M. (2021). How movements shape the perception of time. *Trends in Cognitive Sciences, 25*, 950–963.

Denison, R. N., Yuval-Greenberg, S., & Carrasco, M. (2019). Directing voluntary temporal attention increases fixational stability. *The Journal of Neuroscience, 39*(2), 353–363.

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience and Biobehavioral Reviews, 81*, 181–187.

Doherty, J. R., Rao, A., Mesulam, M. M., & Nobre, A. C. (2005). Synergistic effect of combined temporal and spatial expectations on visual attention. *The Journal of Neuroscience, 25*, 8259–8266.

Donner, T. H., & Siegel, M. (2011). A framework for local cortical oscillation patterns. *Trends in Cognitive Sciences, 15*, 191–199.

Fautrelle, L., Mareschal, D., French, R., Addyman, C., & Thomas, E. (2015). Motor activity improves temporal expectancy. *PLoS One, 10*, e0119187.

Fitch, W. T. (2016). Dance, music, meter and groove: A forgotten partnership. *Frontiers in Human Neuroscience, 10*, 64.

Fraisse, P. (1948). II.-Rythmes auditifs et rythmes visuels. *L'année Psychologique, 49*, 21–42.

Hasbroucq, T., Mouret, I., Seal, J., & Akamatsu, M. (1995). Finger pairings in two-choice reaction time taskscolon: Does the between-hands advantage reflect response preparation? *Journal of Motor Behavior, 27*(3), 251–262.

Janata, P., Tomic, S. T., & Haberman, J. M. (2012). Sensorimotor coupling in music and the psychology of the groove. *Journal of Experimental Psychology. General, 141*, 54–75.

Jaramillo, S., & Zador, A. M. (2011). The auditory cortex mediates the perceptual effects of acoustic temporal expectation. *Nature Neuroscience, 14*, 246–251.

Jaśkowski, P., & Verleger, R. (1993). A clock paradigm to study the relationship between expectancy and response force. *Perceptual and Motor Skills, 77*, 163–174.

Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention, and memory. *Psychological Review, 83*, 323–355.

Keitel, A., & Gross, J. (2016). Individual human brain areas can be identified from their characteristic spectral activation fingerprints. *PLoS Biology, 14*, e1002498.

Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLoS Biology, 16*, e2004473.

Kilavik, B. E., Roux, S., Ponce-Alvarez, A., Confais, J., Grün, S., & Riehle, A. (2009). Longterm modifications in motor cortical dynamics induced by intensive practice. *The Journal of Neuroscience, 29*(40), 12653–12663.

Khalilian-Gourtani, A., Wang, R., Chen, X., Yu, L., Dugan, P., Friedman, D., … & Flinker, A. (2022). A corollary discharge circuit in human speech. BioRxiv, Sept 2022. https://doi.org/10.1101/2022.09.12.507590.

Korolczuk, I., Burle, B., & Coull, J. T. (2018). The costs and benefits of temporal predictability: impaired inhibition of prepotent responses accompanies increased activation of task-relevant responses. *Cognition, 179*, 102–110.

Korolczuk, I., Burle, B., Coull, J. T., & Smigasiewicz, K. (2020). Mechanisms of impulsive responding to temporally predictable events as revealed by electromyography. *Neuroscience, 428*, 13–22.

Korolczuk, I., Burle, B., Coull, J. T., & Śmigasiewicz, K. (2022). Time for action: Neural basis of the costs and benefits of temporal predictability for competing response choices. *Journal of Cognitive Neuroscience, 34*(2), 273–289.

Kusnir, F., Pesin, S., Moscona, G., & Landau, A. N. (2020). When temporal certainty doesn't help. *Journal of Cognitive Neuroscience, 32*(2), 315–325.

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of Neurophysiology, 94*, 1904–1911.

Lakatos, P., O'Connell, M. N., Barczak, A., Mills, A., Javitt, D. C., & Schroeder, C. E. (2009). The leading sense: Supramodal control of neurophysiological context by attention. *Neuron, 64*, 419–430.

Lakatos, P., Musacchia, G., O'Connel, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron, 77*, 750–761.

Lange, K., & Röder, B. (2006). Orienting attention to points in time improves stimulus processing both within and across modalities. *Journal of Cognitive Neuroscience, 18*(5), 715–729.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*, 119–159.

Large, E. W., Herrera, J. A., & Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Frontiers in Systems Neuroscience, 9*, 159.

MacDougall, H. G., & Moore, S. T. (2005). Marching to the beat of the same drummer: The spontaneous tempo of human locomotion. *Journal of Applied Physiology, 99*, 1164–1173.

Manning, F., & Schutz, M. (2013). "Moving to the beat" improves timing perception. *Psychonomic Bulletin & Review, 20*(6), 1133–1139.

Manning, F. C., & Schutz, M. (2016). Trained to keep a beat: Movement-related enhancements to timing perception in percussionists and non-percussionists. *Psychological Research, 80*, 532–542.

Mattes, S., & Ulrich, R. (1997). Response force is sensitive to the temporal uncertainty of response stimuli. *Perception & Psychophysics, 59*, 1089–1097.

McAuley, J. D. (2010). Tempo and rhythm. In M. Riess Jones, R. R. Fay, & A. N. Popper (Eds.), *Music perception* (pp. 165–199). Springer.

McAuley, J. H., Rothwell, J. C., & Marsden, C. D. (1999). Human anticipatory eye movements may reflect rhythmic central nervous activity. *Neuroscience, 94*, 339–350.

Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., Lucas, C., Jacoby, N., Egner, A. A., Hopkins, E. J., et al. (2019). Universality and diversity in human song. *Science, 366*, 20191122.

Menceloglu, M., Suzuki, S., & Song, J. H. (2021). Revealing the effects of temporal orienting of attention on response conflict using continuous movements. *Attention, Perception & Psychophysics, 83*(4), 1463–1478.

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: A neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 370*, 20140093.

Moelants, D. (2002). Preferred tempo reconsidered. In *Proceedings of the seventh international conference on music perception and cognition*, Sydney.

Monier, F., Droit-Volet, S., & Coull, J. T. (2019). The beneficial effect of synchronized action on motor and perceptual timing in children. *Developmental Science, 22*, e12821.

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings National Academy of Sciences. United States of America, 114*, E8913–E8921.

Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences, 1337*, 26–31.

Morillon, B., Schroeder, C. E., & Wyart, V. (2014). Motor contributions to the temporal precision of auditory attention. *Nature Communications, 5*, 5255.

Morillon, B., Hackett, T. A., Kajikawa, Y., & Schroeder, C. E. (2015). Predictive motor control of sensory dynamics in auditory active sensing. *Current Opinion in Neurobiology, 31*, 230–238.

Morillon, B., Schroeder, C. E., Wyart, V., & Arnal, L. H. (2016). Temporal prediction in lieu of periodic stimulation. *The Journal of Neuroscience, 36*, 2342–2347.

Morillon, B., Arnal, L. H., Schroeder, C. E., & Keitel, A. (2019). Prominence of delta oscillatory rhythms in the motor cortex and their relevance for auditory and speech perception. *Neuroscience and Biobehavioral Reviews, 107*, 136–142.

Naghibi, N., Jahangiri, N., Khosrowabadi, R., Eickhoff, C. R., Eickhoff, S. B., Coull, J. T., & Tahmasian, M. (2023). Embodying time in the brain: A multi-dimensional neuroimaging meta-analysis of 95 duration processing studies. *Neuropsychology Review*. https://doi.org/10.1007/s11065-023-09588-1

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin, 89*, 133–162.

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

Nobre, A. C., & van Ede, F. (2023). Attention in flux. *Neuron, 111*(7), 971–986.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: The Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience, 8*, 57.

Phillips-Silver, J., & Trainor, L. J. (2005). Feeling the beat: Movement influences infant rhythm perception. *Science, 308*, 1430.

Phillips-Silver, J., & Trainor, L. J. (2007). Hearing what the body feels: Auditory encoding of rhythmic movement. *Cognition, 105*, 533–546.

Repp, B. H., & Su, Y.-H. (2013). Sensorimotor synchronization: A review of recent research (2006–2012). *Psychonomic Bulletin & Review, 20*, 403–452.

Rizzolatti, G., Riggio, L., Dascola, I., & Umiltá, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia, 25*, 31–40.

Robbe, D. (2023). Lost in time: Relocating the perception of duration outside the brain. *Neuroscience and biobehavioral reviews, 153*, 105312. https://doi.org/10.1016/j.neubiorev.2023.105312

Rohenkohl, G., Cravo, A. M., Wyart, V., & Nobre, A. C. (2012). Temporal expectation improves the quality of sensory information. *The Journal of Neuroscience, 32*, 8424–8428.

Saleh, M., Reimer, J., Penn, R., Ojakangas, C. L., & Hatsopoulos, N. G. (2010). Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. *Neuron, 65*, 461–471.

Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences, 32*, 9–18.

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Current Opinion in Neurobiology, 20*, 172–176.

Schubotz, R. I. (2007). Prediction of external events with our motor system: Towards a new framework. *Trends in Cognitive Sciences, 11*, 211–218.

Servant, M., White, C., Montagnini, A., & Burle, B. (2015). Using covert response activation to test latent assumptions of formal decision-making models in humans. *The Journal of Neuroscience, 35*(28), 10371–10385.

Siegel, M., Donner, T. H., & Engel, A. K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews. Neuroscience, 13*, 121–134.

Smith, D. T., & Schenk, T. (2012). The premotor theory of attention: Time to move on? *Neuropsychologia, 50*, 1104–1114.

Su, Y.-H., & Pöppel, E. (2012). Body movement enhances the extraction of temporal structures in auditory sequences. *Psychological Research, 76*, 373–382.

Tal-Perry, N., & Yuval-Greenberg, S. (2020). Pre-target oculomotor inhibition reflects temporal orienting rather than certainty. *Scientific Reports, 10*(1), 21478.

Tal-Perry, N., & Yuval-Greenberg, S. (2021). Prestimulus inhibition of eye movements reflects temporal expectation rather than time estimation. *Attention, Perception & Psychophysics, 83*(6), 2473–2485.

Tandonnet, C., Burle, B., Vidal, F., & Hasbroucq, T. (2006). Knowing when to respond and the efficiency of the cortical motor command: A Laplacian ERP study. *Brain Research, 1109*(1), 158–163.

Thomas, E., French, R., Alizee, G., & Coull, J. T. (2019). Having your cake and eating it: Faster responses with reduced muscular activation while learning a temporal interval. *Neuroscience, 410*, 68–75.

van Driel, J., Swart, J. C., Egner, T., Ridderinkhof, K. R., & Cohen, M. X. (2015). (No) time for control: Frontal theta dynamics reveal the cost of temporally guided conflict anticipation. *Cognitive, Affective, & Behavioral Neuroscience, 15*(4), 787–807.

Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology, 5*, 1111.

Vuust, P., Heggli, O. A., Friston, K. J., & Kringelbach, M. L. (2022). Music in the brain. *Nature Reviews. Neuroscience, 23*, 287–305.

Wachowiak, M. (2011). All in a sniff: Olfaction as a model for active sensing. *Neuron, 71*, 962–973.

Wang, X.-J. (2010). Neurophysiological and computational principles of cortical rhythms in cognition. *Physiological Reviews, 90*, 1195–1268.

Wiener, M., Turkeltaub, P., & Coslett, H. B. (2010). The image of time: A voxel-wise meta-analysis. *NeuroImage, 49*(2), 1728–1740.

Wollman, I., & Morillon, B. (2018). Organizational principles of multidimensional predictions in human auditory attention. *Scientific Reports, 8*, 13466.

Woodrow, H. (1951). Time perception. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1224–1236). Wiley.

Wyart, V., de Gardelle, V., Scholl, J., & Summerfield, C. (2012). Rhythmic fluctuations in evidence accumulation during decision making in the human brain. *Neuron, 76*, 847–858.

Zalta, A., Petkoski, S., & Morillon, B. (2020). Natural rhythms of periodic temporal attention. *Nature Communications, 11*, 1051.

Zalta, A., Large, E. W., Schön, D., & Morillon, B. (2024). Neural dynamics of predictive timing and motor engagement in music listening. *Science Advances, 10*(10), eadi2525. https://doi.org/10.1126/sciadv.adi2525.

Zatorre, R. J., Evans, A. C., & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *The Journal of Neuroscience, 14*, 1908–1919.

Zatorre, R. J., Chen, J. L., & Penhune, V. B. (2007). When the brain plays music: Auditory motor interactions in music perception and production. *Nature Reviews. Neuroscience, 8*, 547–558.

Zion Golumbic, E. M., Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., Goodman, R. R., Emerson, R., Mehta, A. D., Simon, J. Z., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a "cocktail party". *Neuron, 77*, 980–991.

Coordinate-Based Meta-Analyses of the Time Perception Network

Martin Wiener

Abstract

The study of time perception has advanced over the past three decades to include numerous neuroimaging studies, most notably including the use of functional Magnetic Resonance Imaging (fMRI). Yet, with this increase in studies, there comes the desire to draw broader conclusions across datasets about the nature and instantiation of time in the human brain. In the absence of collating individual studies together, the field has employed the use of Coordinate-Based Meta-Analyses (CBMA), in which foci from individual studies are modeled as probability distributions within the brain, from which common areas of activation-likelihood are determined. This chapter provides an overview of these CBMA studies, the methods they employ, the conclusions drawn by them, and where future areas of inquiry lie. The result of this survey suggests the existence of a domain-general "timing network" that can be used both as a guide for individual neuroimaging studies and as a template for future meta-analyses.

Keywords

Neuroimaging · Meta-analysis · fMRI · Timing network

Introduction

"Time" is ubiquitous, yet timing studies are not. Indeed, for studying the subject, "time" is not one thing. This is because timing studies occupy a diverse landscape of possible experimental task designs (Allman et al., 2014; Vatakis et al., 2018). Temporal discrimination, production, and reproduction can all be used to measure explicit, prospective timing, yet so can (self)paced finger tapping, target anticipation, and oddball detection (Coull & Nobre, 2008). This diversity complicates the pursuit of neuroimaging studies of timing: should we focus on what is common to all timing tasks? Or should we focus on what makes timing varied, labile, and adaptive?

No easy answer exists to this question (Matthews & Meck, 2014; Salet et al., 2022). Yet, there is a richness in the diversity of neuroimaging studies conducted on the study of time perception. For this chapter, I will focus on explicit timing studies. That is, those studies where "time" is the to-be-attended dimension. While studies investigating implicit timing do exist (Wiener et al., 2010a, b), there is a diversity

---

M. Wiener (✉)
George Mason University, Fairfax, VA, USA
e-mail: mwiener@gmu.edu

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_12

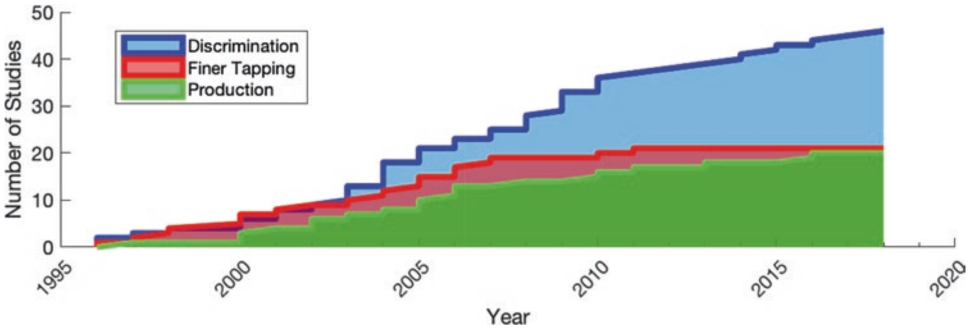


Fig. 1 Two decades of timing studies in neuroimaging. Displayed are cumulative distributions of the number of timing studies for three of the main task designs in time perception: Discrimination, finger tapping, and production. All three studies have increased in prevalence, yet discrimination studies have outpaced the other two types since 2003, whereas the other two main types have plateaued. (Data drawn from (Nani et al., 2019))

among them that may indeed be greater than that found in explicit timing tasks. Part of this is due to the difficulties in building a taxonomy of timing tasks, from which common domains can be determined for inquiry (Paton & Buonomano, 2018; Merchant et al., 2013).

Early attempts in the neuroimaging of time perception highlight the problem in approaching even an area as conserved as explicit timing. Those studies conducted in the 1990s typically focused on paced finger tapping as a measure of timing, which may have grown out of early robust findings in the fMRI literature in examining the neural effects of motor movements (Biswal et al., 1995; Rao et al., 1993). With the advent of event-related fMRI, studies in the 2000s turned to a wider array of experimental designs (Rao et al., 2001), including temporal discrimination and reproduction (Fig. 1). This trend has continued through the 2010s to the present moment this chapter is being written in. Among these newer studies, there have been attempts to examine so-called context effects in time perception (Merchant et al., 2013; Bueti et al., 2008). That is, situations in which time can change as a result of different experimental parameters. This may include the effects of signal emotion, reward, velocity, or magnitude on perceived duration (Matthews & Meck, 2016; Allman et al., 2014). Indeed, beyond finding *where* time can be observed in the brain, these studies highlight attempts to better understand *how* time is encoded (Bueti, 2011). A wider review of these latter studies can be found in other chapters in this volume, but highlight the future of single-imaging studies for time perception, which can include connectionist, multivariate, and encoding-model types.

Coordinate-Based Meta-Analytic Methods

As the neuroimaging literature grew, there were early attempts to provide a better "overview" of findings. Indeed, the exponential rise of fMRI studies led to some concern among researchers for how findings would be concatenated (Fox et al., 1998). This was further compounded by concerns regarding sample sizes in fMRI and the difficult task of determining effect sizes (McGonigle et al., 2000), an issue that is still present today (Grady et al., 2021). Further, concern regarding the generalizability of neuroimaging studies was also present; how certain could researchers be that their findings regarding a particular function would apply to other studies investigating that same function?

To address the above issues, early steps were taken to survey the literature and generate databases of neuroimaging findings. The brainmap database (www.brainmap.org) represents one result of this, in which neuroimaging findings could be categorized and catalogued in a way that other researchers could easily access them as a record (Laird et al., 2005a, b, c). A critical aspect of this was to have appropriate metadata; that is, terms that could sufficiently describe the functions or tasks of interest for a particular study.

From this effort, the main goal was to provide a way to synthesize results from neuroimaging studies that could provide insight regarding the consistency of findings for a particular area. Up until this point, any attempts at meta-analyses for neuroimaging relied on so-called "label-based" methods, in which activated regions that had been labeled by an atlas were collected across studies and those labels that occurred most often were deemed most likely (Laird et al., 2005a, b, c). This method was useful for describing neuroimaging findings, but suffered from being qualitative in nature. With no statistical test to rigorously interrogate the findings, how certain could a researcher be that the meta-analysis was accurate?

To address this, two primary methods were developed independently yet simultaneously: Activation-Likelihood Estimation (ALE) and Multilevel Kernel Density Analysis (MKDA) (Wager et al., 2009). Both methods were concerned with addressing the *likelihood* of activation for any given brain region associated with a particular function or state [p(activation|function)]. Additionally, both methods relied on using as a starting point the three-dimensional coordinates reported for the peaks of activation clusters in neuroimaging studies. For ALE, the approach sought to answer *what is the probability that a given voxel was active in at least one of the included studies*. To answer this, activation foci from reported studies were all assumed to have an activation probability of 1, but each one was then smoothed with a 3D Gaussian function, such that the probability of activation dropped off in every direction (Turkeltaub et al., 2002; Chein et al., 2002). From there, the sum of these functions was taken at each individual voxel, thus representing the ALE statistic. For MKDA, the approach instead asked how many studies reported activation at a given voxel. To answer this, a 3D uniform distribution of 1 s was spread out in a 10 mm radius from each reported activation foci. These values were then summed across studies, such that the final value represented the number of studies reporting activation at a given location (Wager et al., 2004). One notable distinction between ALE and MKDA values is that the latter provided a more readily interpretable statistic; by looking at any region, one could get a sense instantly of how many studies were reporting activation (Bartra et al., 2013). By contrast, ALE values are in themselves difficult to interpret, as their value will depend on numerous factors, including the smoothing kernel for the Gaussian, the number of foci reported, and the distance between those foci. Yet, an advantage of the ALE method is that the graded probability distributions when summed can provide a *relative* difference in activation-likelihood between different voxels and regions that is more nuanced than MKDA, and so one can thus determine which regions are more likely to be activated over others. Regardless, once generated, both methods provided a similar means of assessing statistical significance, in which a random or null distribution was nonparametrically generated by randomizing the reported foci locations and conducting the generating the ALE/MKDA values again with a high number of repetitions (~10,000). Because both methods relied on producing brain maps from reported coordinates, they were referred to as Coordinate-Based Meta-Analyses (CBMA).

Since the advent of primary CBMA methods, a number of advances have been made as the technique has proliferated (Fox et al., 2014). For both ALE and MKDA, stronger inferences were allowed by providing algorithms for assessing false discovery rate and familywise error, as well as cluster-forming thresholds (Laird et al., 2005a, b, c; Eickhoff et al., 2012). Further work also provided a change from fixed-effects models to random-effects, by incorporating the number of subjects within each study as a covariate to modify individual ALE maps (Eickhoff et al., 2009). Other changes were also made to adjust for errors in the design; for example, both ALE and the original version of MKDA (known as "KDA") were sensitive to studies that reported large number of activation foci compared to those that reported fewer ones. However, updates to both algorithms were able to account for this by restricting their statistics to the likelihood of activation *across* studies (Turkeltaub et al., 2012). Other additions to the methods allowed for use of subtraction analyses, in which two ALE maps could be contrasted with one another to examine if one type of task was more likely to have activation at a particular voxel than another type of task (Laird et al., 2005a, b, c). All of these additions improved the robustness of CBMA methods, providing stronger inferences regarding brain activation. However, despite these improvements, there are substantial weaknesses to both CBMA approaches. First, and foremost, is that all CBMA methods rely on modeling the uncertainty associated with activation foci. This modeling, once thresholded, provides a map that may appear similar to fMRI activation maps, as both incorporate smoothness into their images. Yet, CBMA methods have no access to the original shape activation, and as such likely do not reflect the "true" activation probability across studies. Indeed, a study that addressed this possibility by comparing CBMA methods to a meta-analysis that incorporated actual statistical maps from a group of experiments found that these methods only matched the true activation pattern by 45% (Salimi-Khorshidi et al., 2009). However, it was noted that, of the methods tested, ALE provided the relatively closest similarity. As a second weakness, both ALE and MKDA do not take into account differences in effect size between studies and activation foci (Radua & Mataix-Cols, 2009). Rather, all activation foci are treated equally. Yet, in practice this is never the case, as marked differences in the size of an effect will differ across activation peaks. Finally, a third major weakness is that both methods are biased to include only those studies that were published, which naturally ignores those studies that were not. This so-called "file drawer" problem means that CBMA methods likely inflate the likelihood of true activation. Notably, a method to correct for this in ALE has recently been developed (Acar et al., 2018).

Finally, while outside the scope of this review, it should be noted that many other CBMA methods were developed and used (Wager et al., 2009; Samartsidis et al., 2017). Indeed, the basic principle is such that anyone could generate their own CBMA using similar means kable (Bartra et al., 2013). Of importance to mention is that both ALE and MKDA include *user-defined* sets of coordinates. That is, the person conducting the CBMA is the one responsible for finding the activation foci from the particular studies they are interested in. This stands in contrast to *automatic* meta-analytic methods, the most prevalent of which is by Neurosynth (Yarkoni et al., 2011). In the Neurosynth method, rather than running a CBMA on a given set of coordinates, the algorithm attempts to search across the *entire* corpus of neuroimaging studies in online journals, to scan the text of these papers to find a term of interest to the user (i.e., "timing"), extract automatically the reported activation foci from those papers, and then generate a CBMA of those coordinates like MKDA. However, from here, Neurosynth compares this activation map to the remaining corpus—that is, those studies lacking the term of interest—and compares them with a chi-square test. The result is two different activation maps: one which provides the probability of activation for a given function or state, and the other which provides the probability of a function or state for a given activation [p(function|activation)]. This latter term provides a so-called "reverse inference" map, in which one can attempt to ask if certain regions are more likely to be activated for particular functions (Poldrack, 2006).

Previous Meta-Analyses of Time Perception Networks

The first CBMA of time perception was conducted in 2010 (Wiener et al., 2010a, b). Before that point, three label-based meta-analyses had been conducted. Of these three, each incorporated a different set of studies and reached somewhat different conclusions. Lewis and Miall (2003) suggested that the cerebellum and supplementary motor area (SMA) were the most likely to be activated, whereas Penney & Vaitilingham (2008) suggested it was the cerebellum and right inferior frontal gyrus (rIFG), and (Macar et al., 2002) suggested a range of cortical and subcortical structures. With the first meta-analysis, these questions were quantitatively addressed by dividing the corpus of timing studies into those that measured explicit timing at subsecond and suprasecond ranges, as well as whether the task was motor (i.e., paced finger tapping, reproduction) or perceptual (i.e., discrimination, estimation) in nature. Here, the findings demonstrated marked differences in activation-likelihood across sub- and suprasecond ranges, with the former more likely to activate subcortical structures such as the basal ganglia and cerebellum, and the latter recruiting more cortical regions. Crucially, across all timing task variations, the SMA and rIFG were found to be the most commonly active.

Following these initial results, a number of other meta-analyses were run, yet each to address different questions. Indeed, a strength of the CBMA method is that it can ask questions of commonality or differences across studies that may be difficult to ask within an individual study. Ortuño et al. (2011) ran an ALE meta-analysis examining explicit timing studies in both healthy control and schizophrenia patients. For the analysis of healthy control subjects, significant ALE values were found in the SMA, left precentral gyrus, basal ganglia, and thalamus, with reduced activation-likelihood in these same approximate regions for Schizophrenia patients. An additional study also compared activation-likelihood between subjects with Attention-Deficit Hyperactivity Disorder (ADHD) and healthy controls performing timing tasks (Hart et al., 2012). Notably, this study employed another CBMA method, known as Effect-size Signed Differential Mapping (SDM; now known as Seed-based d Mapping). In this method, spatial maps are generated that also take into effect the size and sign of the effect (for example, by incorporating reported *t* statistic values for each peak), and so can account for both direction and magnitude. The results of this meta-analysis demonstrated reduced likelihood in left-hemispheric regions including the IFG, inferior parietal, cerebellum, and insula. Additionally, the right dorsolateral prefrontal cortex (DLPFC) was found to vary depending on medication status across studies in ADHD. A similar study employing SDM was conducted by Radua et al. (2014) that incorporated both time perception and cognitive effort (i.e., working memory and attention). Here, a large overlap between time perception and cognitive effort was observed for many cortical regions associated with time, including the SMA, parietal and prefrontal cortices, with exclusive timing likelihood remaining in the basal ganglia. A crucial insight gained from this study is that many of the regions associated with "timing" were likely engaged in multiple, overlapping functions, noting that the specificity of any one region was difficult to assess.

While the results of Radua and colleagues may suggest that numerous areas associated with "timing" are engaged in other processes during a timing task, it should be noted that CBMA can still afford some insights into the functional subdivisions of these areas. For example, a secondary study followed up on the original Wiener 2010 results by conducting an in-depth analysis of the studies likely to activate solely the SMA (Schwartze et al., 2012). Here, by dividing studies up between motor and perceptual components, the authors demonstrated that activation-likelihood shifted along a rostrocaudal gradient, with perceptual timing studies more likely to activate the anterior SMA, also known as the "pre"-SMA, and motor timing studies more likely activate posterior regions of the SMA "proper." Notably, this finding was also observed in the original 2010 findings (Wiener et al., 2011).

With further neuroimaging studies between 2009 and 2019, a second series of CBMAs for time perception have been run. The first, by (Teghil et al., 2019) divided neuroimaging studies of time perception between those that measured activation while subjects timed an exogenous cue (as in a temporal discrimination task, for example) and those that timed an endogenous cue (as in a self-paced finger tapping task, for example). Here, the general CBMA revealed a pattern of activation likelihood similar to the original 2010 results, including the SMA, bilateral prefrontal and parietal cortices, and the basal ganglia; notably absent was the cerebellum in this study. Between internally and externally driven stimuli, the authors observed that external stimuli were more likely to activate the SMA, rIFG, left precentral gyrus, and insula, suggesting that external stimuli are stronger drivers of time processing than internally based timekeeping.

The second of the "new" studies was conducted by Nani et al. (2019). Here, the authors conducted a more direct replication of the (Wiener et al., 2010a, b) results by dividing studies between sub- and suprasecond and motor and perceptual domains. Importantly, this study also incorporated numerous controls to measure the robustness of results, incorporating the null distribution correction suggested by Acar et al. (2018). Further, the authors measured correlations between ALE maps, as well as a hierarchical clustering to measure similarities. The results of this more conservative meta-analysis nonetheless revealed a similar pattern to the original 2010 findings, yet with a more conserved volume for each region. Notably, the cerebellum and inferior parietal cortices were less likely to be activated overall, with the former only being observed for subsecond motor timing. In terms of similarity, motor and perceptual studies were more similar to each other across duration ranges, yet at the suprasecond range, motor and perceptual timing studies were quite similar. These results supported the original 2010 findings, noting that the overall results had not changed much despite additional studies, and also provided a more nuanced view of the timing landscape.

The third recent CBMA for time perception was conducted by Cona et al. (2021). In this study, timing-related regions were compared to those that were likely to activate space perception and processing, which included spatial navigation, mental rotation, and spatial attention studies among others. Here, the main ALE analysis for time was again similar to the results of both Wiener et al. (2010a, b) and Nani et al. (2019). Between space and time, the former was more likely to activate posterior regions, including occipital and parietal cortices, whereas the latter was more likely to activate anterior regions, including prefrontal cortex, and subcortical regions including the basal ganglia and cerebellum. A conjunction analysis found significant activation-likelihood in the SMA, rIFG, left precentral gyrus, bilateral insula, and inferior parietal cortices. As an additional analysis, these authors examined "gradients" of activation-likelihood within conjunction regions, finding that the SMA, rIFG, right inferior parietal cortex all shifted in activation-likelihood between time and space studies in either a rostrocaudal or dorsoventral direction.

The final and most recent CBMA was performed by Naghibi et al. (2023). In this study, which collected the largest number of neuroimaging studies of timing to date, the authors segregated the studies according to a variety of classifications, including the duration of stimuli, the modality of the stimuli, whether intervals were presented in a sequence or in isolation, whether the task was perceptual or motor in nature, whether subjects were quantifying or predicting intervals, and the nature of the control task. The last comparison was of particular importance, as the choice of control task and its difficulty can have large differences in observed activation patterns (Livesey et al., 2007). As with other CBMA, a similar network of regions were observed; however, the pre-SMA and left anterior insula were the most robust among the different distinctions.

Does a Time Perception Network Exist?

Altogether, the results of the past 10+ years of meta-analyses for time perception have revealed striking consistency (Fig. 2). Indeed, apart from a few areas that have dropped in and out (i.e., cerebellum, parietal cortex), the remaining overall constellation has been fixed. This consistency raises the question of whether or not a true timing "network" exists in the brain. The existence of such a network would be helpful for future neuroimaging ventures of time perception. Indeed, if one already knows *where* activation is likely to be found across any given timing task, then studies can focus more on *how* those regions are involved (Bueti, 2011).

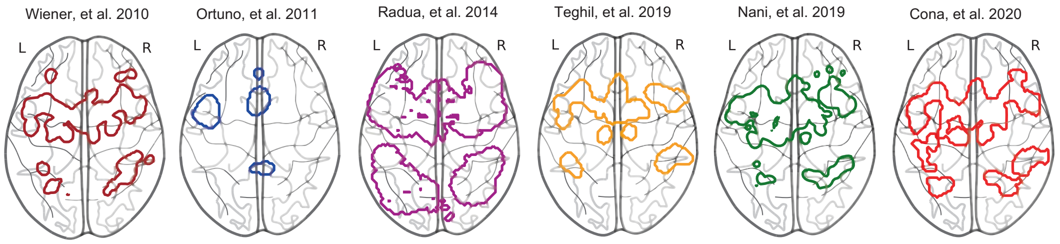


Fig. 2 Timing meta-analyses from 2010 to 2020. From left to right, meta-analyses are presented as contours on a glass brain. The majority of these studies have employed the use of ALE, which changed its methods after 2011. The exception is Radua et al. (2014), which employed Signed Differential Mapping. Across these meta-analyses, a common set of regions can be observed, spanning from the SMA to the basal ganglia, inferior frontal and parietal cortices. Notably, little change has been observed despite the larger number of studies included

In examining the region's most commonly activated across timing meta-analyses, a number of features are readily observed. First, the SMA is consistently the most likely structure to be active across explicit studies of timing. While there are certainly task contexts that influence its function, such as its motor or perceptual nature and spatial context, the region is specifically invoked. Here, then, is our first "node" in an explicit timing network, from which others may diverge. A second area commonly active is the rIFG, spanning pars triangularis and operculum. This region may overlap with the DLPFC, as commonly observed as well in individual studies, yet not commonly reported in meta-analyses. Beyond these two "primary" regions, a number of other nodes are commonly observed across the cortex. These include the bilateral inferior parietal cortices; however, an observed feature is that the right is favored more than the left, with a generally broader distribution that includes the supramarginal gyrus. At the subcortical level, the basal ganglia are also observed bilaterally, including caudate and putamen. Indeed, while these regions may not be commonly observed across *all* timing task variations, it is important to note that the basal ganglia are a set of heterogenous structures, rather than a single unit, and so a lack of activation-likelihood in this region may be due to different studies/contexts activating distinct parts (Wiener et al., 2011). The same context applies to the cerebellum, although here studies are most likely to find activation in subsecond motor paradigms, and most commonly surrounding the dentate gyrus. The thalamus is also commonly observed across neuroimaging studies, although not always as the highest subcortical region. Yet, the thalamus is a critical node for relaying patterns of activity between cortical and subcortical areas; indeed, the striatal beat frequency model of timing (SBF) directly invokes cortico-striato-*thalamic* loops (Matell & Meck, 2004). Finally, stratifying the border between subcortical regions and the cortex are the insular gyri, which are also observed bilaterally. Yet, due to their proximity to the IFG, it is difficult to ascertain at the level of a CBMA if these regions are truly active across timing studies, or merely a result of spreading activation-likelihood from a more lateral cortical source. Regardless, they are included as part of the timing network due to recent work suggesting that interoceptive processes are typically invoked for time processing (Wittmann, 2013).

Altogether, the consensus of timing meta-analyses provides a parcellation that can be called the "Timing Network" (Fig. 3). This network can be used to guide future neuroimaging studies, which can attempt to validate its existence or to further probe interactions between these regions across different task contexts. For example, the edge weights connecting these nodes may vary across different timing tasks or conditions. However, there is still an open question of whether or not a true timing network actually exists. Indeed, among "functional" network parcellations, a number of domain-general networks with distinct yet overlapping activation patterns

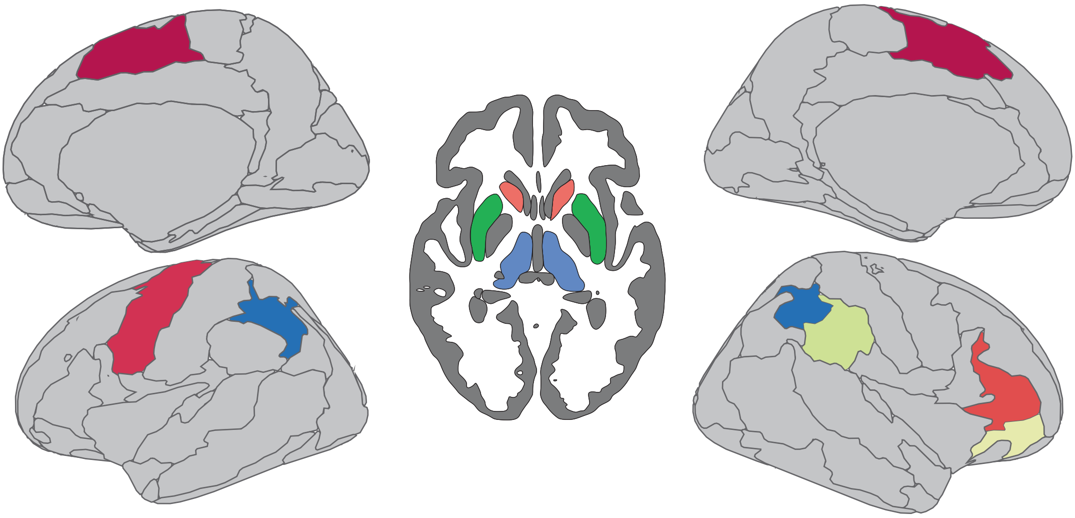


Fig. 3 The proposed timing network. A plot of parcels displayed on a rendered brain, with each colored parcel representing a node in the network. Note that colors here are arbitrary. Each parcel was drawn from the AAL atlas to include those regions most commonly reported across meta-analyses of explicit timing studies. The middle segment includes an axial slice with subcortical structures highlighted. Not displayed here is the cerebellum, which is included in the network. This network is available for download at https://neurovault.org/collections/13081/

are already well-known (Mattar et al., 2015). These include the default mode, dorsal attention, somatomotor, language, cingular-opercular, and multiple demand networks, among numerous others, which will certainly overlap with the proposed timing network. However, the timing network may have a unique collection of regions compared to these others, spanning both cortical and subcortical regions. A true dissociation of these networks from the timing network would require comparisons of structural, functional, and resting-state networks. But, the timing network as conceived here, based on consistent meta-analytic findings, is a reasonable place to start.

The Future of Timing Meta-Analyses

The discussion of prior CBMAs for time perception is meant to highlight two things: (1) the consistency among them for regions associated with time perception, leading to the conclusion of a generalized timing network, and (2) the possibility for future analyses that have not yet been done. Indeed, CBMAs have a strong utility in their ability to address questions of consistency. To that end, there are numerous other possibilities for CBMAs of the timing network that have not yet been employed.

First, among CBMA methods, only ALE and SDM have been used to measure timing networks. This leaves MKDA as a method that has not yet been employed. However, while this may present an opening for a novel meta-analysis, I suggest that the results of MKDA would likely not differ from the prior meta-analyses. This is because, at their core, all CBMA methods rely on the same general strategy for modeling activation based on reported foci (Fox et al., 1998). However, other methods for CBMA exist that may provide at least somewhat divergent findings. One recently developed method, the Analysis of brain coordinates (ABC) shows promise (Tench et al., 2022). Briefly, this method considers at the first-level what *clusters* of studies are most likely to occur across the brain volume, which differs from the ALE/MKDA approach in which clusters are later defined after statistical thresholding. The ABC method further thresholds these clusters based on the expected proportion that would occur by chance. As a result, significant clusters for ABC report clusters that are likely to *replicate* across the corpus of included studies.

Second, other extensions to the ALE algorithm exist that have not yet been tested. The first is termed Meta-Analytic Connectivity Modeling (MACM) (Laird et al., 2009). MACM is a type of connectivity analysis that shares the same strategy as that used by seed-based connectivity measures. Specifically, MACM works by isolating a particular voxel or region of the brain, and then searching for all studies that report activation for that particular region. Notably, this search may be restricted to only include studies that investigate a particular area (i.e., timing). The activation foci for these studies are extracted and a standard ALE analysis is run, with the resulting MACM map displaying those regions that are significantly associated with the seed region of interest. For example, one could specify the SMA as a node, and then examine all other regions that are commonly activated with it. Comparisons between regions in the timing network may yield details regarding how different regions interact across different task contexts.

While MACM provides a measure of *association,* it is important to stress that it does not provide a measure of *connectivity* in the sense applied to studies of resting-state or task-based fMRI, and even their claims of connections may be spurious (Leonardi & Van, 2015). A closer measure for CBMA is the recently developed Co-activation Probability Estimation (CoPE) method (Chu et al., 2015). In the CoPE method, activation foci are treated as probability distributions, similar to ALE, but with a smaller width. From here, values are normalized and the co-activation of each voxel is measured across studies. That is, which voxels are likely to be activated together across studies? The resulting measures are compared against a null distribution from Monte Carlo simulations for statistical significance, resulting in a map where clusters represent those that are co-active across studies. The important distinction of the CoPE method is that it can distinguish between local and long-range connectivity, and so be used to derive a connectivity matrix between regions. Applying CoPE for time perception would allow for a true measure of network properties observed across studies. Yet, as of this writing, no software package for CoPE is publicly available, limiting its use.

As an alternative to the CoPE method, a more recent CBMA connectivity measure has been proposed, for which a freely available software package exists (Tench et al., 2020). This method, known as Coordinate-based meta-analysis of networks (CBMAN), is a variation of the ABC method described above. Broadly, the CBMAN method works by measuring z-scores associated with reported activation peaks and examining their covariance structure across the included structures. As a result, multivariate normal distributions can be fit to the z-scores of the most likely clusters for activation, with the covariance used to estimate connectivity between clusters. As this method includes both activation foci and effect sizes, it provides a strong measure for inferring connectivity; ripe for the study of timing.

A final, untapped method is to examine reverse inferences for the timing network. That is, up until now, all of the CBMA methods here report the likelihood of activation *given* the set of included studies [p(activation|timing)]. However, they do not speak to the converse inference: what is the probability of a timing task having occurred, given activation is found in a particular region [p(timing|activation)]. For example, if a study finds a significant cluster in the SMA, was the subject timing? Knowing this probability can provide insight into the *specificity* of any one region for timing. However, given the ubiquity of timing studies, it is likely that no single region has a high *absolute* probability for timing, but rather there will be *relative* differences between regions (e.g., if SMA activation is observed, is it more likely that a subject was timing than if right parietal activation is found). As described above, the Neurosynth method provides a means to assess this. Yet, Neurosynth relies on automatic tagging of studies based on terms of interest, and the term “timing” likely includes those studies associated with time perception and those that aren’t. In fact, the Neurosynth website does include this term, but with no clear clusters available for reverse inference. However, given the

ALE values represent the forward probability of activation, it is possible via Bayes Theorem to construct the posterior probability. This method has recently been proposed for ALE (Costa et al., 2021), and provides a simple software plugin to accomplish it. The only requirement, however, is to have a set of "nontiming" studies to compare with it.

Recently, we noted this method was employed by our group (Mondok & Wiener, 2022). Here, we employed the timing studies used in our previous meta-analysis (Cona et al., 2021) and conducted the reverse-inference analysis as described above. Two main findings emerged from this analysis. First, the overall probability of a timing task having been conducted, given activation of a particular brain region, was low. Indeed, no single region offered high predictive value for determining if a study was employing a timing study over another task. However, it should be noted that *many* tasks offer low predictive value, especially when a large network of regional activations are possible (Yarkoni et al., 2011). Nonetheless, among those regions that were predictive, the SMA and the bilateral insula had the highest predictive power for timing tasks. In particular, we note an interesting convergence with the recent results of (Naghibi et al., 2023), who in their standard CBMA also found these regions as having the most consistent likelihood.

Conclusions

In discussing CBMA methods and their application for time perception, a final open question is whether or not there are new analyses available at the aggregate level that can yield insights to how timing is accomplished in the brain. Hopefully, the new methods described just above can be applied with important distinctions available, and when carefully applied can provide further details about the existence and flexibility of the timing network. Further, additional methods may come along that provide a new leap in our understanding of time at the collective level. Regardless, the foundation for any meta-analysis is the individual studies that support it. As timing studies continue to be done with neuroimaging, more detailed questions can be asked, and more can be learned.

References

Acar, F., Seurinck, R., Eickhoff, S. B., & Moerkerke, B. (2018). Assessing robustness against potential publication bias in Activation Likelihood Estimation (ALE) meta-analyses for fMRI. *PLoS One, 13*, e0208177.

Allman, M. J., Teki, S., Griffiths, T. D., & Meck, W. H. (2014). Properties of the internal clock: First- and second-order principles of subjective time. *Annual Review of Psychology, 65*, 743–771.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage, 76*, 412–427.

Biswal, B., Yetkin, F. Z., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magnetic Resonance in Medicine, 34*, 537–541.

Bueti, D. (2011). The sensory representation of time. *Frontiers in Integrative Neuroscience, 5*, 34.

Bueti, D., Walsh, V., Frith, C., & Rees, G. (2008). Different brain circuits underlie motor and perceptual representations of temporal intervals. *Journal of Cognitive Neuroscience, 20*, 204–214.

Chein, J. M., Fissell, K., Jacobs, S., & Fiez, J. A. (2002). Functional heterogeneity within Broca's area during verbal working memory. *Physiology & Behavior, 77*, 635–639.

Chu, C., Fan, L., Eickhoff, C. R., Liu, Y., Yang, Y., Eickhoff, S. B., & Jiang, T. (2015). Co-activation Probability Estimation (CoPE): An approach for modeling functional co-activation architecture based on neuroimaging coordinates. *NeuroImage, 117*, 397–407.

Cona, G., Wiener, M., & Scarpazza, C. (2021). From ATOM to GradiATOM: Cortical gradients support time and space processing as revealed by a meta-analysis of neuroimaging studies. *NeuroImage, 224*, 117407.

Costa, T., Manuello, J., Ferraro, M., Liloia, D., Nani, A., Fox, P. T., Lancaster, J., & Cauda, F. (2021). BACON: A tool for reverse inference in brain activation and alteration. *Human Brain Mapping, 42*, 3343–3351.

Coull, J., & Nobre, A. (2008). Dissociating explicit timing from temporal expectation with fMRI. *Current Opinion in Neurobiology, 18*, 137–144.

Eickhoff, S. B., Laird, A. R., Grefkes, C., Wang, L. E., Zilles, K., & Fox, P. T. (2009). Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human Brain Mapping, 30*, 2907–2926.

Eickhoff, S. B., Bzdok, D., Laird, A. R., Kurth, F., & Fox, P. T. (2012). Activation likelihood estimation meta-analysis revisited. *NeuroImage, 59*, 2349–2361.

Fox, P. T., Parsons, L. M., & Lancaster, J. L. (1998). Beyond the single study: Function/location metanalysis in cognitive neuroimaging. *Current Opinion in Neurobiology, 8*, 178–187.

Fox, P. T., Lancaster, J. L., Laird, A. R., & Eickhoff, S. B. (2014). Meta-analysis in human neuroimaging: Computational modeling of large-scale databases. *Annual Review of Neuroscience, 37*, 409–434.

Grady, C. L., Rieck, J. R., Nichol, D., Rodrigue, K. M., & Kennedy, K. M. (2021). Influence of sample size and analytic approach on stability and interpretation of brain-behavior correlations in task-related fMRI data. *Human Brain Mapping, 42*, 204–219.

Hart, H., Radua, J., Mataix-Cols, D., & Rubia, K. (2012). Meta-analysis of fMRI studies of timing in attention-deficit hyperactivity disorder (ADHD). *Neuroscience and Biobehavioral Reviews, 36*, 2248–2256.

Laird, A. R., Fox, P. M., Price, C. J., Glahn, D. C., Uecker, A. M., Lancaster, J. L., Turkeltaub, P. E., Kochunov, P., & Fox, P. T. (2005a). ALE meta-analysis: Controlling the false discovery rate and performing statistical contrasts. *Human Brain Mapping, 25*, 155–164.

Laird, A. R., Lancaster, J. L., & Fox, P. T. (2005b). BrainMap: The social evolution of a human brain mapping database. *Neuroinformatics, 3*, 065–078. https://doi.org/10.1385/ni:3:1:065

Laird, A. R., McMillan, K. M., Lancaster, J. L., Kochunov, P., Turkeltaub, P. E., Pardo, J. V., & Fox, P. T. (2005c). A comparison of label-based review and ALE meta-analysis in the Stroop task. *Human Brain Mapping, 25*, 6–21.

Laird, A. R., Eickhoff, S. B., Li, K., Robin, D. A., Glahn, D. C., & Fox, P. T. (2009). Investigating the functional heterogeneity of the default mode network using coordinate-based meta-analytic modeling. *The Journal of Neuroscience, 29*, 14496–14505.

Leonardi, N., & Van, D. V. D. (2015). On spurious and real fluctuations of dynamic functional connectivity during rest. *NeuroImage, 104*, 430–436.

Lewis, P. A., & Miall, R. C. (2003). Distinct systems for automatic and cognitively controlled time measurement: Evidence from neuroimaging. *Current Opinion in Neurobiology, 13*, 250–255.

Livesey, A. C., Wall, M. B., & Smith, A. T. (2007). Time perception: Manipulation of task difficulty dissociates clock functions from other cognitive demands. *Neuropsychologia, 45*, 321–331.

Macar, F., Lejeune, H., Bonnet, M., Ferrara, A., Pouthas, V., Vidal, F., & Maquet, P. (2002). Activation of the supplementary motor area and of attentional networks during temporal processing. *Experimental Brain Research, 142*, 475–485.

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Brain Research. Cognitive Brain Research, 21*, 139–170.

Mattar, M. G., Cole, M. W., Thompson-Schill, S. L., & Bassett, D. S. (2015). A functional cartography of cognitive systems. *PLoS Computational Biology, 11*, e1004533.

Matthews, W. J., & Meck, W. H. (2014). Time perception: The bad news and the good. *Wiley Interdisciplinary Reviews: Cognitive Science, 5*, 429–446.

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

McGonigle, D. J., Howseman, A. M., Athwal, B. S., Friston, K. J., Frackowiak, R. S., & Holmes, A. P. (2000). Variability in fMRI: An examination of intersession differences. *NeuroImage, 11*, 708–734.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time. *Annual Review of Neuroscience, 36*, 313–336.

Mondok, C., & Wiener, M. (2022). Selectivity of timing: A meta-analysis of temporal processing in neuroimaging studies using activation likelihood estimation and reverse inference. *Frontiers in Human Neuroscience, 16*, 1000995.

Naghibi, N., Jahangiri, N., Khosrowabadi, R., Eickhoff, C. R., Eickhoff, S. B., Coull, J. T., & Tahmasian, M. (2023). Embodying time in the brain: A multi-dimensional neuroimaging meta-analysis of 95 duration processing studies. *Neuropsychology Review, 34*, 277. https://doi.org/10.1007/s11065-023-09588-1

Nani, A., Manuello, J., Liloia, D., Duca, S., Costa, T., & Cauda, F. (2019). The neural correlates of time: A meta-analysis of neuroimaging studies. *Journal of Cognitive Neuroscience, 31*, 1796–1826.

Ortuño, F., Guillén-Grima, F., López-García, P., Gómez, J., & Pla, J. (2011). Functional neural networks of time perception: Challenge and opportunity for schizophrenia research. *Schizophrenia Research, 125*, 129–135.

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

Penney, T. B., & Vaitilingam, L. (2008). Imaging time. Psychology of time, 261–294.

Poldrack, R. A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences, 10*, 59–63.

Radua, J., & Mataix-Cols, D. (2009). Voxel-wise meta-analysis of grey matter changes in obsessive-compulsive disorder. *The British Journal of Psychiatry, 195*, 393–402.

Radua, J., del Pozo, N. O., Gómez, J., Guillen-Grima, F., & Ortuño, F. (2014). Meta-analysis of functional neuroimaging studies indicates that an increase of cognitive difficulty during executive tasks engages brain regions associated with time perception. *Neuropsychologia, 58*, 14–22.

Rao, S. M., Binder, J. R., Bandettini, P. A., Hammeke, T. A., Yetkin, F. Z., Jesmanowicz, A., Lisk, L. M., Morris, G. L., Mueller, W. M., & Estkowski, L. D. (1993). Functional magnetic resonance imaging of complex human movements. *Neurology, 43*, 2311–2318.

Rao, S. M., Mayer, A. R., & Harrington, D. L. (2001). The evolution of brain activation during temporal processing. *Nature Neuroscience, 4*, 317–323.

Salet, J. M., de Jong, J., & van Rijn, H. (2022). Still stuck with the stopwatch. *Behavioral Neuroscience, 136*, 453–466.

Salimi-Khorshidi, G., Smith, S. M., Keltner, J. R., Wager, T. D., & Nichols, T. E. (2009). Meta-analysis of neuroimaging data: A comparison of image-based and coordinate-based pooling of studies. *NeuroImage, 45*, 810–823.

Samartsidis, P., Montagna, S., Nichols, T. E., & Johnson, T. D. (2017). The coordinate-based meta-analysis of neuroimaging data. *Statistical Science, 32*, 580–599.

Schwartze, M., Rothermich, K., & Kotz, S. A. (2012). Functional dissociation of pre-SMA and SMA-proper in temporal processing. *NeuroImage, 60*, 290–298.

Teghil, A., Boccia, M., D'Antonio, F., Di Vita, A., de Lena, C., & Guariglia, C. (2019). Neural substrates of internally-based and externally-cued timing: An activation likelihood estimation (ALE) meta-analysis of fMRI studies. *Neuroscience and Biobehavioral Reviews, 96*, 197–209.

Tench, C. R., Tanasescu, R., Constantinescu, C. S., Cottam, W. J., & Auer, D. P. (2020). Coordinate based meta-analysis of networks in neuroimaging studies. *NeuroImage, 205*, 116259.

Tench, C. R., Tanasescu, R., Constantinescu, C. S., Auer, D. P., & Cottam, W. J. (2022). Easy to interpret coordinate based meta-analysis of neuroimaging studies: Analysis of brain coordinates (ABC). *Journal of Neuroscience Methods, 372*, 109556.

Turkeltaub, P. E., Eden, G. F., Jones, K. M., & Zeffiro, T. A. (2002). Meta-analysis of the functional neuroanatomy of single-word reading: Method and validation. *NeuroImage, 16*, 765–780.

Turkeltaub, P. E., Eickhoff, S. B., Laird, A. R., Fox, M., Wiener, M., & Fox, P. (2012). Minimizing within-experiment and within-group effects in activation likelihood estimation meta-analyses. *Human Brain Mapping, 33*, 1–13.

Vatakis, A., Balcı, F., Luca, M. D., & Correa, Á. (2018). *Timing and time perception: Procedures measures, and applications*. BRILL. https://doi.org/10.1163/9789004280205

Wager, T. D., Jonides, J., & Reading, S. (2004). Neuroimaging studies of shifting attention: A meta-analysis. *NeuroImage, 22*, 1679–1693.

Wager, T. D., Lindquist, M. A., Nichols, T. E., Kober, H., & Van, S. J. X. (2009). Evaluating the consistency and specificity of neuroimaging data using meta-analysis. *NeuroImage, 45*, S210–S221.

Wiener, M., Turkeltaub, P., & Coslett, H. B. (2010a). The image of time: A voxel-wise meta-analysis. *NeuroImage, 49*, 1728–1740.

Wiener, M., Turkeltaub, P. E., & Coslett, H. B. (2010b). Implicit timing activates the left inferior parietal cortex. *Neuropsychologia, 48*, 3967–3971.

Wiener, M., Matell, M. S., & Coslett, H. B. (2011). Multiple mechanisms for temporal processing. *Frontiers in Integrative Neuroscience, 5*, 31.

Wittmann, M. (2013). The inner sense of time: How the brain creates a representation of duration. *Nature Reviews. Neuroscience, 14*, 217–223.

Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van, E. D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods, 8*, 665–670.

Probing Beat Perception with Event-Related Potentials (ERPs) in Human Adults, Newborns, and Nonhuman Primates

Fleur L. Bouwer, Gábor P. Háden,
and Henkjan Honing

Abstract

The aim of this chapter is to give an overview of how the perception of rhythmic temporal regularity such as a regular beat in music can be studied in human adults, human newborns, and nonhuman primates using event-related brain potentials (ERPs). First, we discuss different aspects of temporal structure in general, and musical rhythm in particular, and we discuss the possible mechanisms underlying the perception of regularity (e.g., a beat) in rhythm. Additionally, we highlight the importance of dissociating beat perception from the perception of other types of structure in rhythm, such as predictable sequences of temporal intervals, ordinal structure, and rhythmic grouping. In the second section of the chapter, we start with a discussion of auditory ERPs elicited by infrequent and frequent sounds: ERP responses to regularity violations, such as mismatch negativity (MMN), N2b, and P3, as well as early sensory responses to sounds, such as P1 and N1, have been shown to be instrumental in probing beat perception. Subsequently, we discuss how beat perception can be probed by comparing ERP responses to sounds in regular and irregular sequences, and by comparing ERP responses to sounds in different metrical positions in a rhythm, such as on and off the beat or on strong and weak beats. Finally, we will discuss previous research that has used the aforementioned ERPs and paradigms to study beat perception in human adults, human newborns, and nonhuman primates. In doing so, we consider the possible pitfalls and prospects of the technique, as well as future perspectives.

F. L. Bouwer (✉)
Cognitive Psychology Unit, Institute of Psychology, Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands

Department of Psychology, Brain & Cognition, University of Amsterdam,
Amsterdam, The Netherlands
e-mail: f.l.bouwer@fsw.leidenuniv.nl

G. P. Háden
Institute of Cognitive Neuroscience and Psychology, Budapest, Hungary

Department of Telecommunications and Media Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, Budapest, Hungary

H. Honing
Music Cognition group (MCG), Institute for Logic, Language and Computation (ILLC), Amsterdam Brain and Cognition (ABC), University of Amsterdam, Amsterdam, The Netherlands

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_13

Keywords

Auditory perception · Music · Rhythm · Beat · Meter · Event-related potentials (ERPs) · Mismatch negativity (MMN) · Temporal expectations · Prediction · EEG

Introduction

In music, as in several other domains, events occur over time. The way events are structured in time, both in music and in other domains, allows the brain to anticipate the timing of events and, in doing so, to optimize processing of events that occur at expected moments in time (Nobre & van Ede, 2018). In addition, temporal expectations (e.g., predicting "when" an event will happen) can guide our movement. This is of particular interest when considering musical rhythm, where temporal expectations allow us to dance and make music together (Honing, 2012; Leow & Grahn, 2014), and may play a role in our enjoyment of music (Fiveash et al., 2023). Temporal expectations can be formed based on different information in the environment, such as the contingency between a cue and a temporal interval, and the passage of time itself (Nobre & van Ede, 2018). Such interval-based predictions, as well as foreperiod effects, are discussed in depth elsewhere (Buhusi & Meck, 2005; Ng & Penney, 2014). Here, we focus on temporal expectations as present in *rhythm*, which denotes the temporal structure of a sequence of multiple events.

Figure 1 shows a schematic overview of rhythmic structure. A rhythmic sequence of seven sounds is depicted, with temporal intervals of various lengths separating the sounds, forming a rhythmic pattern of shorter and longer temporal intervals. In music, importantly, in addition to structure in the form of a rhythmic pattern (Fig. 1B), rhythm often induces the perception of a regular pulse or *beat* (Bouwer et al., 2021; Nobre & van Ede, 2018). The beat (Fig. 1C) is a perceived regularly recurring salient moment in time (Cooper & Meyer, 1960) that we can tap and dance to. In musical rhythm, the beat often coincides with an event, but a beat can also coincide with plain silence (see shaded area in Fig. 1): Listeners can perceive a beat even in the absence of cues to this regularity in the rhythmic signal and can persist in perceiving a beat in the pres-

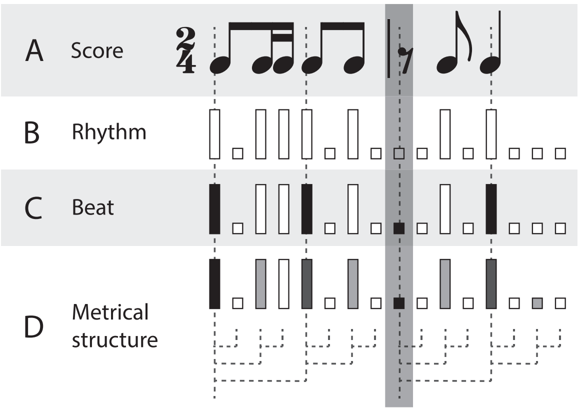


Fig. 1 Schematic overview of structure in rhythm. Rhythm can be conceptualized as a sequence of events in time. Panel A depicts an example rhythm in common music notation. In Panel B, C, and D, sounds are depicted as vertical bars. On top of perceiving the rhythmic pattern formed by the temporal structure of the sounds (e.g., the succession of longer and shorter intervals in time, panel B), we can perceive a regular beat, here depicted as black events (C). Several nested hierarchical levels of regularity make up a metrical structure (D), with differences in salience between strong beats (depicted in black), weak beats (depicted in dark gray), and subdivisions of the beat (in light gray and white). The metrical interpretation is represented as a metrical tree, with the length of the branches representing the theoretical metric salience of a specific position in the sequence. Note that the third beat (shaded area) coincides with silence: this is a "loud rest" or *syncopation*, with a missing event on a perceived beat

ence of conflicting rhythmic information (Honing & Bouwer, 2019; Longuet-Higgins & Lee, 1984). The beat is often embedded in a hierarchical structure of multiple perceived levels of temporal regularity. At a higher level, we can hear regularity in the form of regular stronger and weaker beats (often referred to as *meter*, like in a waltz, which has a strong-weak-weak pattern of beats), and at a lower level, we can perceive regular subdivisions of the beat. Together, these regularities create a hierarchical pattern of saliency known as a metrical structure (Fig. 1D). We can perceive temporal regularity with a period roughly in the timescale of 200 to 2000 ms (London, 2002, 2012). Within this range, we have a clear preference for beats with a period around 600 ms or 100 beats per minute (Fraisse, 1982), and while listeners can to some extent guide the level of regularity they attend to most (Drake et al., 2000), the regularity closest to the preferred rate is often considered most salient (e.g., the beat).

In this chapter, we first discuss the processes underlying the perception of a regular beat, and possible considerations for designing stimuli that induce beat perception. Next, we discuss how beat perception can be studied using event-related potentials (ERPs), and we give an overview of studies probing beat perception with ERPs in human adults, human newborns, and nonhuman animals. The current chapter updates a previous overview on this topic (Honing et al., 2014). Note that we focus on perceptual aspects of beat perception. For a discussion of how beat perception relates to movement, the motor system, and motor entrainment, see overviews on this topic elsewhere (Cannon & Patel, 2021; Damm et al., 2019; Merchant et al., 2015; Repp & Su, 2013).

Mechanisms of Beat Perception

Entrainment as a Mechanism for Beat Perception

The perception of a regular beat and the temporal expectations we form in response to a beat are often explained within the framework of entrainment (Henry & Herrmann, 2014; Obleser & Kayser, 2019): the synchronization of an internal regularity to the regularity in an external stimulus. From a psychological perspective, entrainment has been described by Dynamic Attending Theory (DAT) (Jones, 2009; Large & Jones, 1999). DAT proposes that internal fluctuations in attentional energy, termed attending rhythms, elicit expectations about when future events occur. The internal fluctuations in attentional energy can adapt their phase and period to an external rhythm, leading to alignment of peaks in attentional energy with metrically strong positions (i.e., peaks in attentional energy fall on the beat). At moments of heightened attentional energy, events are expected to occur, and processing of events is enhanced (Haegens & Zion Golumbic, 2018). The attending rhythms are thought to be self-sustaining and can occur at multiple nested levels, tracking events with different periods simultaneously (Drake et al., 2000; Large & Jones, 1999). These features of the Dynamic Attending model correspond respectively to the stability of our metrical percept and the perception of multiple hierarchical levels of regularity (Large, 2008). Behavioral support for DAT comes from studies showing a processing advantage on the beat (e.g., in phase with an external regularity) for perceiving temporal intervals (Large & Jones, 1999), pitch (Jones et al., 2002), intensity changes (Bouwer et al., 2020; Bouwer & Honing, 2015), and phonemes (Quené & Port, 2005). The processing advantage persists after a rhythmic sequence ends, in line with the supposed self-sustaining nature of the attending rhythms (Hickok et al., 2015; Saberi & Hickok, 2022b). However, note that recently, the persistent behavioral facilitation of events in phase with a regularity outlasting rhythmic stimulation could not always be replicated, which spurred discussion on the automaticity and ubiquity of entrainment (Bauer et al., 2015; Bouwer, 2022; Lin et al., 2021; Saberi & Hickok, 2022a, b; Sun et al., 2021). Several explanations for these discrepant findings have been suggested, including the presence of large individual differences in the strength of entrainment (Bauer et al., 2015; Saberi & Hickok, 2022b; Sun et al., 2021), the dependence of entrainment on uncertainty in the auditory input (Saberi & Hickok, 2022b), and the dependence of entrainment on the rate of the rhythmic signal (Pesnot Lerousseau et al., 2021; Saberi & Hickok, 2022b).

At the neural level, entrainment may be implemented by the alignment of low-frequency oscillations (in the delta range; 0.5-4 Hz) to external regularity (Haegens & Zion Golumbic, 2018; Henry & Herrmann, 2014; Obleser & Kayser, 2019; Rimmele et al., 2018), leading to heightened neural sensitivity at expected time points (Haegens & Zion Golumbic, 2018; Henry & Herrmann, 2014), akin to the peaks in attentional energy described by DAT. In line with the self-sustaining nature of the attending rhythms described by DAT, neural oscillations can also retain their alignment to a rhythmic sequence after sensory stimulation is stopped (Bouwer et al., 2023; Kösem et al., 2018; van Bree et al., 2021). Entrainment has mainly been studied in the context of regular, periodic stimulation, to explain the prediction of regular, isochronous beats (e.g., predictions that are equally spaced in time). Recently however, models of entrainment have also been used to explain predictions for non-isochronous rhythmic patterns (e.g., the predictions of successions of short and long intervals, that are not necessarily of equal length), in the context of irregular meters as found in Balkan music (Tichko & Large, 2019). For the purposes of this chapter, importantly, entrainment theories, both at the psychological and neural level, predict that processing is enhanced for events that are in phase with the entraining signal (Haegens & Zion Golumbic, 2018).

Predictive Processing as a Mechanism for Beat Perception

The perception of a beat is a bidirectional process: not only can a varying musical rhythm induce the perception of a regular beat (hence also referred to as "beat induction" (Honing, 2012)), but a regular beat can also influence the perception of the very same rhythm that induces it. Hence beat perception can be seen as an interaction between bottom-up and top-down sensory and cognitive processes (Desain & Honing, 1999), and as such fits well within the framework of predictive processing (Koelsch et al., 2019; Vuust & Witek, 2014). Within this framework, the perceived metrical structure provides a representation within which incoming sounds are interpreted. This representation is constantly updated based on the incoming sensory information. The relation between the events in the music and the perceived metrical structure thus is a flexible one, in which the perceived metrical structure is both inferred from the music and has an influence on how we perceive the music (Desain & Honing, 2003; Grube & Griffiths, 2009). Of importance to the current chapter, within predictive processing models, it is often assumed that sensory processing for expected events is attenuated (Friston, 2005). Thus, entrainment and predictive processing accounts of rhythm perception make somewhat different predictions about the underlying mechanisms of beat perception (Bouwer & Honing, 2015; Palmer & Demos, 2022), be it synchronization of an internal regularity with an external one, or creation of a hierarchical mental representation of the beat regularity. Now that we have considered the possible mechanisms underlying the perception of a beat, in the next section we consider aspects of rhythmic stimuli that may induce a perceived beat.

Beat from the Bottom Up: Considerations for Stimulus Design

Inducing a Beat from a Rhythmic Sequence

The simplest rhythmic stimulus that may induce the perception of a regular beat is an isochronous sequence (e.g., a sequence with identical durations between tones, like a metronome, see Fig. 2, example 1A). To probe beat perception, responses to sounds in such sequences have been compared to responses to sounds in sequences with irregular, jittered timing (Fig. 2, example 1B), with the premise that while an isochronous sequence can elicit a perceived beat, a jittered sequence cannot.

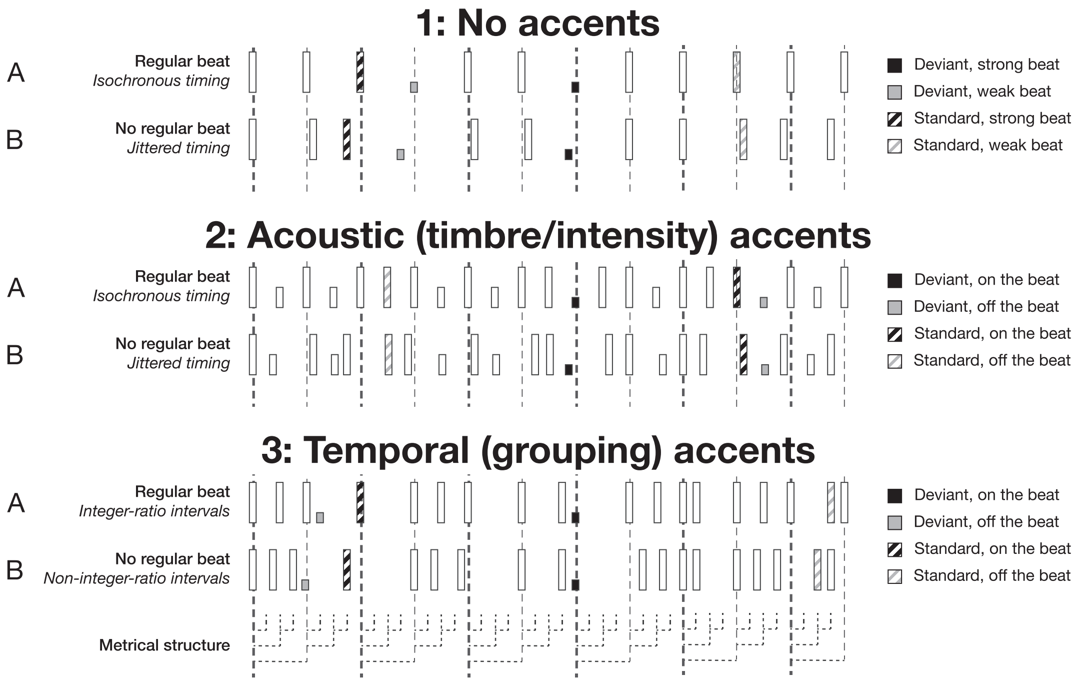


Fig. 2 Examples of rhythmic sequences used to study beat perception. Rhythms consist of sound events represented here by vertical bars. Dashed vertical lines represent the perceived beats. Long vertical bars and bars of intermediate length (for example 2) represent (frequent) standard sounds. The shortest vertical bars represent (infrequent) deviant sounds, such as an unexpected decrease in loudness, which are used to elicit a specific series of ERPs (see *ERPs in response to expectancy violations*). The tree structure underneath the example rhythms depicts the (theoretical) perceived metrical structure for the rhythms that induce a beat (example A). 1) The simplest stimulus to study beat perception is arguably an isochronous sequence (1A), with a rate within the range of human preferred tempo. Responses to deviant and standard events in such a sequence can be compared to responses to the same events in jittered sequences (1B, see for example (Schwartze et al., 2011, 2013; Teki et al., 2011)). Some studies have also compared responses to deviants in odd (black) and even (gray) positions in isochronous sequences (1A), to study subjective accenting (Brochard et al., 2003; Potter et al., 2009). 2) A rhythm with alternating loud (long white vertical bars) and soft (intermediate white vertical bars) sounds is thought to induce a regular duple beat when it has isochronous timing (2A). Beat perception can be probed by comparing responses to events on the beat (black) and off the beat (gray), either for deviant or standard sounds. Care must be taken to compare events that are acoustically identical and occur in an acoustically identical context. To control for sequential learning, the difference between responses on and off the beat in an isochronously timed sequence (2A) can be compared to the same contrast in a sequence with jittered timing and the same statistical structure (2B), which is thought to induce sequential learning, but not a beat (Bouwer et al., 2016; Háden et al., 2024; Honing et al., 2018). Note that here, it is also possible to (subjectively) perceive strong and weak beats (e.g., perceive yet another level of regularity). For the sake of simplicity, this is not depicted in this Fig. 3) When a rhythm has non-isochronous timing, a beat can be induced by temporal grouping accents. When an (accented) event mostly occurs with regular intervals, listeners will infer a regular beat (3A). Here, to control for grouping, like in example 2, responses in a sequence with regular accents and integer-ratio durations (3A) can be contrasted with responses in a sequence with the same grouping structure, but irregular accents and non-integer-ratio durations (3B), which is thought not to induce a perceived beat (Bouwer et al., 2020; Grahn & Brett, 2007). As for example 2, for the sake of simplicity, only one level of regularity (the beat) is depicted here

Thus, a difference in responses to events in such regular and jittered sequences may be ascribed to the presence of a perceived beat in the former but not the latter. In addition to the regularity at the beat level, we can also perceive metrical structure in isochronous sequences, even if all sounds are identical (i.e., an equitone sequence). It was shown that listeners perceive events in odd positions (Fig. 2, example 1A, black shades) as more salient than events in even positions (Fig. 2, example 1A, gray shades), in line with odd positions representing metrically accented and even positions representing metrically unaccented events (Brochard et al., 2003; Potter et al., 2009). This phenomenon is termed subjective accenting and is reminiscent of perceiving "tick-tock" when listening to a clock instead of "tick-tick" (e.g., we hear a stronger and weaker "tick" and "tock" even if all "ticks" are in fact physically identical). Note that subjective accenting may depend on the rate of the sequence, with listeners shifting the number of notes perceived as one group (or one beat) depending on the tempo. Interestingly, while humans prefer a beat at a rate around 600 ms, subjective accenting seems to favor rates that are slower, suggesting that it is akin to the alternation of strong and weak beats (e.g., meter), rather than events on and off the beat (Bååth, 2015; Poudrier, 2020). In addition, subjective accenting is a highly variable effect that does not always occur in all listeners (Criscuolo et al., 2023).

While listeners can thus perceive beat and meter in isochronous, equitone sequences, in natural rhythm, the beat usually needs to be inferred from a varying rhythmic signal. Moreover, the regularity listeners perceive need not even be apparent from a rhythmic signal, as in fact, rhythm sometimes does not contain regularity at the beat rate at all (Tal et al., 2017). To infer a metrical structure from music with a varying rhythmic structure, we often make use of accents. In a sequence of events, an accent is a more salient event because it differs from other, non-accented events along some auditory dimension (Ellis & Jones, 2009). When accents exhibit regularity in time, we can induce a regular beat from them. Accented tones are then usually perceived as on the beat or, on a higher level, as coinciding with a strong rather than a weak beat (Lerdahl & Jackendoff, 1983). Loudness accents may be used to allow listeners to infer a beat from a rhythm (Bouwer et al., 2018), and pitch accents also have been shown to play a role in perceiving the beat (Ellis & Jones, 2009; Hannon et al., 2004). Indeed, spectral information may even be more informative for the brain to entrain to than the sound envelope of a rhythm (e.g., changes in loudness or onsets) (Weineck et al., 2022). It is very likely that in natural music, many sound features can contribute to an accent structure and our perception of the beat, including not only loudness and pitch but also timbre. In line with this, the use of ecologically valid stimuli may enhance the perception of a beat (Bolger et al., 2013; Tierney & Kraus, 2013). Example 2A in Fig. 2 depicts a rhythm which mostly consists of alternating loud (long vertical white bars) and softer (intermediate vertical white bars) tones. Such a pattern would induce a duple beat through loudness accents, with some events falling on the beat (black shades) and some events falling off the beat (gray shades).

Accents can also arise from the perceptual grouping of rhythmic events in time, even when sounds are acoustically identical. When an onset is isolated in time relative to other onsets, it sounds like an accent. Second, when two onsets are grouped together, the second onset sounds accented. Finally, for groups of three or more onsets, the first and/or last tone of the group will be perceived as an accent (Povel & Essens, 1985). Such temporal accents may drive the perception of a beat in a bottom-up manner. Recordings from midbrain neurons in rodents have shown increased firing rate for events on the beat compared to events off the beat in rhythms with purely temporal accents, consistent with the idea that increased responses to tones that are salient based on temporal grouping may drive human beat perception (Rajendran et al., 2017, 2020). Example 3A in Fig. 2 shows a rhythm in which the beat is elicited by temporal accents. Here, a beat can be perceived through temporal accents that are regularly spaced, with an (accented) event always coinciding with perceived beat times. Note that to perceive a beat in this type of rhythm, not only regular spacing of accents but also the presence of intervals with integer-ratio durations is of importance (Grahn & Brett, 2007; Jacoby & McDermott, 2017). For example, if the regularity of the beat is present with a period of 600 ms (i.e., 100 beats per minute), it is beneficial to the perception of the beat if a rhythm contains temporal intervals of 150, 300, and 450 ms, which are all related to the beat interval at integer ratios (in this case ratios of 4/1, 2/1, and 4/3). Both for example 2A and example 3A, we would expect differential responses to events on and off the beat, as events on the beat are more expected if a beat is perceived.

In addition to bottom-up influences on a perceived beat from accents and the temporal structure of a rhythm, listeners can impose different metrical structures on rhythmic sequences if instructed to do so (Iversen et al., 2009; Nozaradan et al., 2011), and cultural background and experience may affect the beat we perceive (Gerry et al., 2010; Hannon & Trehub, 2005; Jacoby & McDermott, 2017; Kaplan et al., 2022; Yates et al., 2016).

Dissociating Beat Perception from Duration-Based Temporal Expectations

Importantly, one challenge in beat perception research is to dissociate responses to a regular beat from responses to other types of structure in the rhythm, such as duration-based temporal structure, ordinal structure, low-level acoustic differences, and temporal grouping. First, we can, in addition to hearing a beat, perceive temporal structure in predictable single durations, and predictable rhythmic patterns, be it by learning the contingency between a cue and a specific temporal duration, by learning a sequence of absolute intervals (e.g., the time intervals between two events), or by learning a rhythmic pattern in the form of relative durations (e.g., the ratios between consecutive inter-onset intervals) (Bouwer et al., 2020, 2023; Breska & Deouell, 2017; Morillon et al., 2016; Nobre & van Ede, 2018). Neuroimaging work suggests that specific networks are dedicated to perceiving absolute and relative durations respectively. While a network comprising the cerebellum and the inferior olive is involved in absolute duration-based timing, a different network, including the basal ganglia and the SMA, is active for relative or beat-based timing (Teki et al., 2011). It is still unclear how the perception of absolute durations, relative durations, rhythmic patterns, and metrical structure are related, with some suggesting that the underlying mechanism for pattern and beat perception is similar (Cannon, 2021; Cannon & Patel, 2021) and some suggesting separate mechanisms (Bouwer et al., 2020, 2023). Hence, when studying beat perception, it is important to take into account possible overlap between temporal structure based on a beat, and temporal structure based on patterns and absolute durations (Bouwer et al., 2021). This may be a challenge for studies relying on isochrony to study beat perception (e.g., Fig. 2, example 1), as the temporal structure in an isochronous sequence can be described both in terms of its regularity, and in terms of the repetition of a single interval (Bouwer et al., 2021; Keele et al., 1989). To account for this, the use of more complex stimuli, with at least one level of hierarchy (e.g., some events on the beat and some events off the beat, like in examples 2 and 3) may be instrumental.

Dissociating Beat Perception from Ordinal Structure

When a beat is elicited by accents in otherwise isochronous sequences, to induce two levels in a metrical hierarchy (Fig. 2, example 2A), one challenge that arises in probing beat perception is that in strongly beat inducing sequences, the accents themselves also introduce ordinal structure. For instance, in example 2A, a listener may infer that a soft sound is always followed by a louder sound and that a loud sound is followed by a soft sound in most cases (e.g., loud and soft sounds mostly alternate). Thus, listeners may learn the ordinal, statistical structure of a sequence (Conway & Christiansen, 2001), something humans are capable of at a young age (Saffran et al., 1999). To account for such ordinal structure in probing beat perception, one approach is to not just compare the difference between responses to sounds in metrically strong and weak positions (e.g., on and off the beat) in isochronous sequences (Fig. 2, example 2A), in which this difference is affected by statistical learning and the perceived beat, but to contrast this difference with the difference between responses to sounds in metrically strong and weak positions in jittered sequences (Fig. 2, example 2B), in which the difference is only affected by statistical learning (Bouwer et al., 2016).

Dissociating Beat Perception from Low-Level Acoustics and Grouping

In natural music, a beat is often induced by creating accents on the beat (similar to example 2A in Fig. 2). Because accented sounds by definition need to stand out from non-accented sounds, this often means that tones on the beat have a different sound than tones that are not on the beat. Similarly, the acoustic context (e.g., the tone preceding the tone of interest) of weak and strong metrical positions is not identical. Such acoustic differences may lead to differences in low-level perceptual features like masking, and may affect sensory responses (Bouwer et al., 2014; Honing et al., 2014; Winkler et al., 2013). To account for this, stimuli must ideally be controlled to be able to probe different metrical positions with identical acoustic properties. To this end, example 2A contains occasional events on offbeat positions that have a loud sound, making them identical to sounds on the beat. This makes it possible to probe sounds on and off the beat with identical acoustic properties and context (Bouwer et al., 2016; Háden et al., 2024; Honing et al., 2018).

Finally, temporal grouping may be instrumental in inducing a beat in non-isochronous rhythms (Fig. 2, example 3A). That is, grouping of events may lead to perceived accents, which may then be used by a listener to abstract a beat structure from a non-isochronous sequence (Povel & Essens, 1985). However, it must be considered that differences in salience due to perceptual grouping may lead to differences in neural responses, regardless of the presence of a beat (Andreou et al., 2015). To account for this, a similar strategy as described above for ordinal structure may be followed, whereby responses to events in a non-isochronous rhythm with regularly spaced accents, which is thought to induce a beat (e.g., Fig. 2, example 3A), are compared to responses to events in a rhythm without regularly spaced accents, but with an identical grouping structure (e.g., Fig. 2, example 3B), which is thought to not induce a beat (Bouwer et al., 2020).

To summarize, in musical rhythm, humans often perceive nested, hierarchical levels of regularity known as a metrical structure, with the most salient level of regularity representing the beat. The perception of beat and meter has been explained by entrainment theories and theories of predictive processing, which make slightly different predictions for how the perceived metrical structure affects the processing of events in a rhythm (e.g., entrainment should lead to enhanced processing of events in strong metrical positions, while predictions should lead to attenuation of events that are expected) (Bouwer & Honing, 2015; Lange, 2013). A metrical structure can be inferred from a rhythm through accents in various forms. Importantly, in studying beat perception, the perception of metrical structure needs to be dissociated from other types of structure present in rhythm, such as duration-based temporal structure, ordinal structure, low-level acoustic variability, and grouping. Hence, a simple comparison of responses on and off the beat is often not enough to infer something about beat perception, as events on and off the beat often differ in many more characteristics than just their metrical position.

Measuring Beat Perception with Event-Related Potentials (ERPs)

Some of the main questions regarding beat perception are concerned with whether beat perception is innate (or spontaneously developing) and/or species-specific (Honing, 2018; Honing et al., 2014). Testing human newborns and nonhuman animals to answer these questions requires a method that is noninvasive and does not require an overt response from the participant. EEG is well suited for this task and has the temporal resolution to track the perception of a beat over time. Several different approaches exist in probing beat perception with EEG. Analyses of EEG responses in the frequency domain may directly probe the entrainment of low-frequency neural oscillations to an external regularity (Nozaradan, 2014; Tal et al., 2017), but also need to account for possible methodological pitfalls (Novembre & Iannetti, 2018; Zoefel et al., 2018). Here, we focus on the well-studied approach of analyzing event-related potentials (ERPs), and we discuss several recent studies that have used ERPs to probe beat perception in human adults, newborns, and nonhuman primates.

Auditory ERPs

ERPs are hypothesized to reflect the sensory and cognitive processing in the central nervous system associated with particular (auditory) events (Luck, 2005). ERPs are isolated from the EEG signal by averaging the signal in response to many trials containing the event of interest. Through this averaging procedure, any activity that is not time-locked to the event is averaged out, leaving the response specific to the event of interest: the ERP. While ERPs do not provide a direct functional association with the underlying neural processes, there are several advantages to the technique, such as the ability to record temporally fine-grained and covert responses not observable in behavior. Also, several ERP components have been well studied and documented, not only in human adults but also in newborns and nonhuman animals. Some of these components, used in testing beat perception, are elicited with an oddball paradigm.

ERPs in Response to Expectancy Violations

An auditory oddball paradigm consists of a frequently recurring sequence of stimuli (standards), in which infrequently a stimulus is changed (deviant) in some feature (e.g., pitch, intensity, and timing). The deviant stimulus thus violates the expectations that are established by the standard stimuli. Depending on the task of the subject a deviant stimulus elicits a series of ERP components reflecting different stages and mechanisms of processing. The mismatch negativity (MMN) is a negative ERP component elicited between 100 and 200 ms after the deviant stimulus. MMN is thought to reflect automatic deviance detection through a memory-template matching process (Kujala et al., 2007; Näätänen et al., 2007), and can be elicited by expectancy violations in sound features such as pitch, duration, or timbre (Winkler, 2007; Winkler & Czigler, 2012), abstract rules (Paavilainen et al., 2007), or stimulus omissions (Yabe et al., 1997). The N2b is a component similar to the MMN in latency, polarity, and function, but it is only elicited when the deviant is attended and relevant to the task (Schröger & Wolff, 1998). At around 300 ms after the deviant stimulus, a positive component can occur, known as the P3a, which reflects attention switching and orientation toward the deviant stimulus. For task-relevant deviants, this component can overlap with the slightly later P3b, reflecting match/mismatch with a working memory representation (S. H. Patel & Azzam, 2005; Polich, 2007). The latency and amplitude of the MMN, N2b, P3a, and P3b are sensitive to the relative magnitude of the expectancy violation (Comerchero & Polich, 1999; Fitzgerald & Picton, 1983; Rinne et al., 2006; Schröger & Winkler, 1995) and correspond to discrimination performance in behavioral tasks (Novitski et al., 2004). These properties are exploited when probing beat perception with ERPs. Moreover, ERP responses to expectancy violations, most notably the MMN, have been recorded in comatose patients (Näätänen et al., 2007), sleeping newborns (Alho et al., 1992), and anesthetized animals (Csépe et al., 1987), making ERP research an ideal instrument for interspecies comparisons and for testing the innateness of beat perception.

ERPs in Response to Frequent Stimuli

While the abovementioned ERPs are elicited by expectancy violations, any sound will elicit a succession of obligatory responses, regardless of whether a sound is frequent or infrequent. Hence, in addition to using responses to expectancy violations to probe beat perception, we can also compare responses to frequent sounds (standards). In the current chapter, we focus on two early sensory responses (as studied in humans): the P1 and the N1. The auditory P1 (sometimes termed P50, as it typically peaks at about 50 ms post-stimulus onset) and N1 (sometimes termed N100, as it typically peaks around 100 ms post-stimulus onset) components are thought to be generated in auditory cortices, and are sensitive to stimulus features, like loudness and pitch change, and presentation rate (Näätänen & Picton, 1987; Picton et al., 1974; Winkler et al., 2013). In addition, N1 has been shown to be affected by both attention and expectations, including temporal expectations (Lange, 2013; Picton & Hillyard, 1974), making it a potentially informative component to study in the context of musical rhythm.

Using ERPs to Probe Beat Perception

The general idea of using ERPs to probe beat perception is that an event on the beat is perceived differently from an event occurring not on the beat due to the metrical expectations of the listener, and thus that two physically identical events in different metrical positions should yield different brain responses. More specifically, ERP responses elicited by expectancy violations (e.g., MMN, N2b, P3a, P3b) are typically larger for more unexpected events. If a beat is perceived, we form strong expectations for events to occur on the beat (Honing & Bouwer, 2019). Hence, ERPs in response to expectancy violations that interfere with the perceived beat (e.g., a deviant softer sound on the beat, depicted by the short vertical bars in black shades in Fig. 2, examples 2A and 3A) should be larger than ERPs in response to violations that do not interfere with a perceived beat, either because they are in line with the metrical structure (e.g., a deviant softer sound in an offbeat position, depicted by the short vertical bars in gray shades in Fig. 2, examples 2A and 3A) or because no beat is perceived (e.g., a deviant softer sound in a jittered sequence, depicted by the short vertical bars in Fig. 2, examples 2B and 3B). In addition, several components of the obligatory auditory-evoked potential (e.g., the P1 and N1 responses) are smaller for expected than unexpected sounds, in line with predictive processing accounts that predict the silencing of the predicted sensory input (Lange, 2013). Hence, in the presence of a perceived metrical structure, events in weak metrical positions (e.g., standard sounds off the beat, depicted by the long vertical bars in gray shades in Fig. 2, examples 2A and 3A) are less expected than events in strong metrical positions (e.g., standard sounds on the beat, depicted by the long vertical bars in black shades in Fig. 2, examples 2A and 3A) and may therefore elicit stronger responses.

ERP responses to expectancy violations and P1 and N1 responses can also be affected by attention. The N2b and P3b only occur when a stimulus is task-relevant (Polich, 2007; Schröger & Wolff, 1998), while the MMN can be modulated by attention (Haroush et al., 2010), and can even be completely eliminated when deviations in attended and unattended auditory streams vie for feature-specific processing resources (Sussman, 2007). Since dynamic attending theory predicts enhanced processing in metrically strong positions due to a peak in attentional energy, we may expect that ERP components in response to expectancy violations are affected by metrical position due to differences in attention, with larger responses to events that coincide with peaks in attention (e.g., in strong metrical positions). At the same time, N1 has been shown to be enhanced by attention, hence the response to events in strong metrical positions may be larger than the response to events in weak metrical positions (Haegens & Zion Golumbic, 2018).

Note that several mechanisms may thus affect ERPs to rhythm in different ways, and sometimes even in opposite directions, with larger responses to events in strong metrical positions due to attention effects, and smaller responses due to the effects of expectations (Lange, 2013). Also, in most cases, an implicit assumption made by studies using oddball designs is that expectations for when a sound will occur are coupled with expectations for the sound itself (“what”). In other words, in the studies below, when an expectation is violated, it is almost always the expectations for a certain sound (“what”) that is violated, and not the expectation for sound timing itself. Whether expectations for timing can be formed at all without any expectation for sound identity is a subject for debate, and outside of the scope of this chapter (Clarke, 2005; Gibson, 1975; Morillon et al., 2016).

Importantly, when using ERPs to probe beat perception, the ERPs are not a direct index of the processes involved in the perception of a beat.

Rather, ERPs are used that have been extensively studied over the years and are known to be affected by attention and expectations. Since the main mechanisms underlying beat perception have been associated with the processes of attending and expectancy (see *Mechanisms of beat perception*), ERPs can be used to index the strength of beat perception by indexing the strength of attention and expectations. The ERPs themselves thus do not reflect the perceived beat, but rather, are modulated by it. We will now turn to a discussion of research that has used ERPs to probe beat perception in human adults, newborns, and nonhuman primates.

Probing Beat Perception in Human Adults with ERP Responses to Expectancy Violations

Comparing Isochronous to Jittered Sequences

As described above, the simplest way of probing beat perception is by comparing responses to infrequent sounds embedded within an isochronous, presumably beat inducing sequence (Fig. 2, example 1A) with responses to infrequent sounds within a sequence with jittered timing (Fig. 2, example 1B). Typically, P3 responses are larger, and sometimes earlier, for deviants in isochronous than jittered sequences (Lange, 2009; Rimmele et al., 2011; Schmidt-Kassow et al., 2009; Schwartze et al., 2011), in line with stronger expectations being formed about the occurrence of events in isochronous sequences. This effect is somewhat attenuated in cerebellar patients (Kotz et al., 2014) and children with developmental coordination disorder (Chang et al., 2021), and has been related to movement, both in healthy adults and Parkinson patients (Conradi et al., 2016; Lei et al., 2019), confirming a role for motor networks in the formation of temporal expectations. Results for earlier responses are somewhat mixed, but larger N2b responses to deviants in isochronous than jittered sequences have been observed (Kotz et al., 2014; Rimmele et al., 2011). The effect seems to be attention-dependent, though numerically, the same effect can be seen for the MMN in unattended conditions (Schwartze et al., 2011).

However, of note, because an isochronous stimulus is used, it is unclear whether these results are due to beat perception, or rather, differences in learning single intervals. Interestingly, one study in the visual domain found no differentiation between the effects of temporal expectations on the P3 for isochronous sequences and cue-based expectations (Breska & Deouell, 2017). Also, similar P3 effects can be observed for sequences with grouping structure, but not temporal regularity (Schmidt-Kassow et al., 2009). Thus, the contrast between responses to isochronous and jittered sequences likely contains a combination of beat perception and the perception of other types of regularity.

Comparing Responses to Strong and Weak Beats

To account for the presence of duration-based temporal processing, one option is to add an extra hierarchical level to the metrical structure and examine differences between metrical positions. One example of how deviant detection can show the presence of metrical perception comes from studies examining subjective rhythmization (Brochard et al., 2003; Potter et al., 2009). In these studies, participants were presented with an isochronous series of tones. They were hypothesized to perceive the tones in odd positions as stronger than tones in even positions, due to an imposed duple metrical structure. Infrequently, a softer tone was introduced, either in odd or in even positions (Fig. 2, example 1A). These deviants elicited an N2b and a P3b. The P3b to deviants in odd positions had a larger amplitude than the P3b to deviants in even positions, showing that the deviants were indeed detected better—or perceived as more violating—on a strong beat than on a weak beat (Brochard et al., 2003; Potter et al., 2009). In a related study, physical accents in the form of tones with longer durations were used to induce a duple or triple meter, and similar to the subjective rhythmization studies, the P3 response to softer target tones was larger in strong than weak metrical positions. Here, a similar effect was found for the earlier N2b components, albeit only in the duple meter condition (Abecasis et al., 2005).

Note that in these studies, the isochronous sequence on which a structure was imposed was at a rate close to the preferred tempo for humans, and as such, the difference between odd and even positions can be interpreted more as meter (e.g., strong and weak beats) than as beat (e.g., on the beat and off the beat). It is unclear whether these results are based on listeners imposing the temporal structure of regularity at the level of the meter or on listeners imposing a hierarchical grouping structure, with groups of two or three events. To examine this, one strategy could be to contrast the difference between responses to strong and weak beats in an isochronous sequence (e.g., example 1A in Fig. 2) with the same difference in a jittered sequence (e.g., example 1B in Fig. 2).

Of note, in the studies looking at subjective rhythmization, the rhythmic sequences were always task-relevant, and the ERP components of interest were the attention-dependent P3b and N2b. One other study examined meter processing under unattended conditions by using drum rhythms with occasionally omitted sounds on strong and weak beats, and found a latency difference for the MMN dependent on meter (e.g., shorter latency for strong than weak beat violations (Ladinig et al., 2009, 2011)). However, these findings could not be replicated in a bigger sample (Bouwer et al., 2014), suggesting that meter processing may require attention. Additionally, meter processing may be affected by musical training (Nave-Blodgett et al., 2021).

Comparing Responses on and off the Beat

At the level of the beat, several studies have used oddball paradigms to study the difference in responses on and off the beat. For drum rhythms with infrequent omissions, MMN responses were larger for omissions on the beat than off the beat, even when the sequences were not attended (Bouwer et al., 2014). Similarly, MMN was larger for intensity decrements in odd then even positions for isochronous sequences at a rate that corresponded to twice the preferred rate of humans (e.g., the isochronous sequence was at the level of subdivisions of a beat, with odd events on the beat and even events off the beat (Bouwer & Honing, 2015)). Interestingly, for isochronous sequences without acoustic cues to the hierarchical metrical structure, this effect was larger for Western listeners than for bicultural listeners who are familiar with sub-Saharan African music (Haumann et al., 2018), indicative of an effect of experience. Indeed, another study found that deviance responses to omissions on and off the beat were related to musical training (Silva & Castro, 2019), and differences may also be due to innate variability in strategies for temporal processing (Snyder et al., 2010).

Note that for the abovementioned studies, deviants consisted of softer sounds or omissions. Both entrainment and predictive processing accounts of beat perception would predict these deviants to be more salient in strong metrical positions, either since more processing resources are focused on those points in time, or because listeners form strong expectations for louder sounds on the beat (Bouwer & Honing, 2015). Results from studies using intensity increments as deviants may be more in line with the latter explanation, as these consistently found larger ERP responses to unexpected increments off the beat then on the beat (Abecasis et al., 2009; Bouwer & Honing, 2015; Geiser et al., 2010). In these studies, however, no jittered control sequences (e.g., examples B in Fig. 2) were used, leaving the possibility open that the rhythmic stimuli may have induced a temporal grouping structure. Of note, behaviorally, listeners show the effects of grouping even for non-isochronous rhythmic sequences with a timing structure that does not induce a beat easily (Bouwer et al., 2020).

Importantly, the stimuli used by (Bouwer et al., 2014) contained multiple types of structure in addition to the beat. This study used drum rhythms with alternating bass drum, snare drum, and hihat sounds. While omissions on the beat always followed a hihat sound, omissions off the beat followed a bass drum sound—an order of events that was overall more likely in the sequences. Hence, the observed effects could be due to statistical learning (i.e. learning the transitional probabilities between consecutive sounds), rather than beat perception (Bouwer et al., 2016). In a follow-up study, this was accounted for by using jittered sequences as a control condition (e.g., Fig. 2, example 2B). Here, the difference in ERP responses (MMN, N2b, and P3a) to intensity decrements on and off the beat was larger in the isochronous sequences (Fig. 2, example 2A) than in the jittered sequences (Fig. 2, example 2B), which the authors took as evidence for beat perception. This effect came on top of the statistical learning that was evident from the difference between responses to deviants on and off the beat in the jittered sequences, and the effect was present regardless of attention to the sequences (Bouwer et al., 2016). The results from this study are depicted in Fig. 3A. One open question is whether the differences in responses could potentially be due to better statistical learning in isochronous than jittered sequences, as temporal expectations have been shown to affect statistical learning (Tsogli et al., 2022).

Finally, several studies have found bigger ERP responses for deviations from the rhythmic surface structure than deviations from the hierarchical metrical (and ordinal) structure, both for rhythms consistent of drum sounds indicating the metrical structure (Vuust et al., 2005, 2009) and rhythms with temporal accents (Edalati et al., 2021; Geiser et al., 2009). This shows that absolute temporal expectations can greatly influence ERP responses to rhythm, and highlights the importance of controlling for differences in the surface structure of rhythm.

To summarize, a large collection of studies has now shown the presence of differences in responses to deviant sounds on and off the beat for several ERP components, including the MMN, N2b, P3a, and P3b, often without attention directed to a rhythmic stimulus, and with musically untrained listeners. However, it remains a challenge to design stimuli that can readily ascribe this effect to beat perception.

Probing Beat Perception in Human Adults by Looking at the Auditory P1 and N1 Response

Comparing Isochronous to Jittered Sequences

A large body of research has shown smaller sensory responses to sounds in isochronous than jittered sequences, both for the N1 component (Foldal et al., 2020; Kotz et al., 2014; Lange, 2009, 2010; Makov & Zion Golumbic, 2020; Schwartze et al., 2013; Schwartze & Kotz, 2015; van Atteveldt et al., 2015) and the P1 component (Brinkmann et al., 2021; Rimmele et al., 2011; Schwartze et al., 2013, 2015; Schwartze & Kotz, 2015). This is in line with the attenuation of expected sounds as predicted by predictive processing accounts of temporal expectations. This effect was shown to be largely independent of attention (Makov & Zion Golumbic, 2020; Schwartze et al., 2013). While the use of isochronous sequences without hierarchical structure prohibits strong conclusions about the involvement of beat-based perception, of note, this effect was diminished in patients with basal ganglia lesions (Schwartze et al., 2015), but not in patients with cerebellar lesions (Kotz et al., 2014). As the basal ganglia, but not the cerebellum, is specifically involved in beat-based perception (Grahn, 2009; Merchant et al., 2015), this may suggest that for isochronous sequences, temporal expectations rely at least to some extent on beat perception.

Comparing Responses in Different Metrical Positions

While early sensory responses are usually attenuated by the presence of temporal predictability in isochronous sequences, interestingly, studies comparing early sensory responses on and off the beat have found opposite results, with larger responses for events on the beat. This was found for the N1 response for rhythms with temporal accents indicating the beat (Abecasis et al., 2009) and melodies with pitch structure indicating the beat (Fitzroy & Sanders, 2015), and for the P1 response for isochronous sequences at a fast rate (e.g., with odd tones being on the beat and even

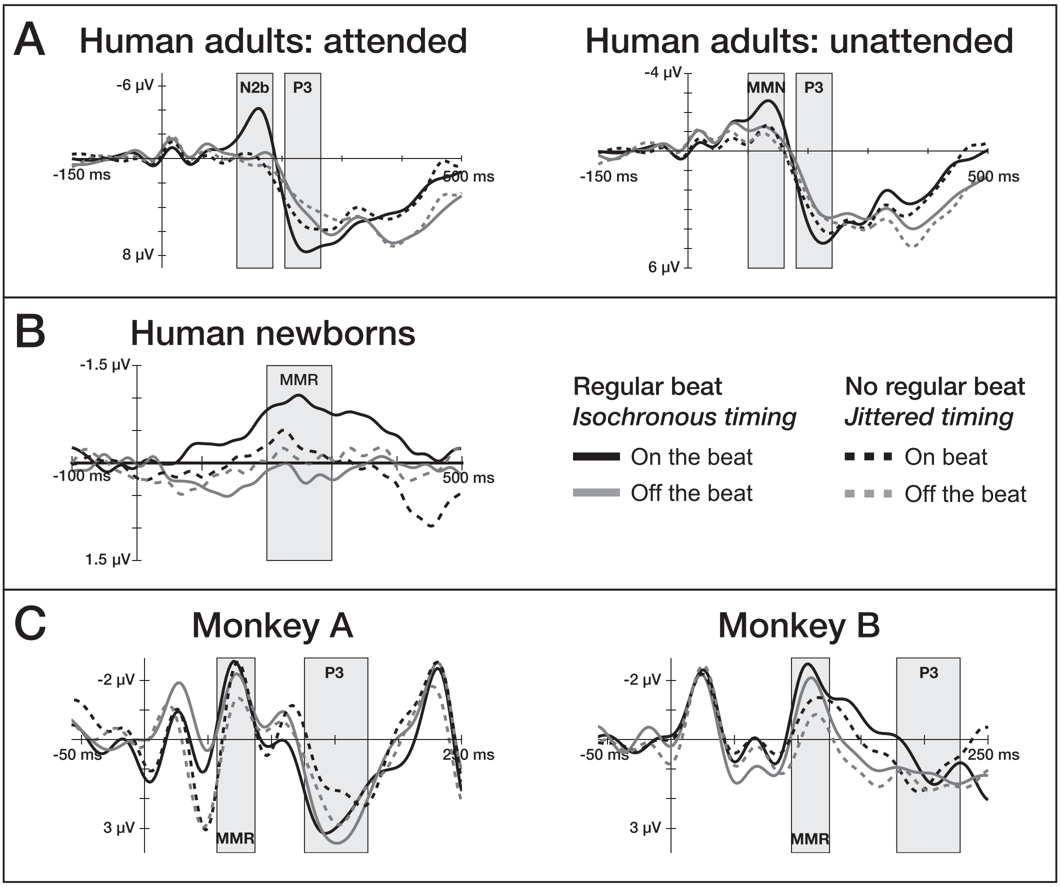


Fig. 3 ERP results from studies probing beat perception in human adults, newborns, and nonhuman primates. (**a**) Difference waves (e.g., difference between ERPs to deviant and standard sounds) for infrequent intensity decrements presented within isochronous and jittered sequences, either on the beat or off the beat (Fig. 2, example 2). For human adults, N2b, MMN, and P3 responses were larger on the beat (black) than off the beat (gray), and this difference was more pronounced in isochronous (solid lines) than jittered sequences (dashed lines), suggestive of beat perception (Bouwer et al., 2016). (**b**) In newborns, similar to adults, the MMR was largest for deviants on the beat in isochronous sequences, providing evidence for the presence of beat processing (Háden et al., 2024). Note that the latency and morphology of newborn MMR are very different from the MMN found in adults. (**c**) In two nonhuman primates presented with the same paradigm (Fig. 2, example 2), the MMR was larger for deviants presented within isochronous sequences (solid lines) than for deviants presented within jittered sequences (dashed lines). However, here, the difference between the responses to deviants on and off the beat was not larger in the isochronous than jittered condition, suggesting that while the animals were capable of perceiving the temporal regularity of the isochronous sequences, they did not represent the full metrical structure including the beat (Honing et al., 2018). Note that like for newborns, the morphology of the ERPs and the latency of the MMR are different from that commonly found in human adults (see also Table 1), and highly variable between individuals

tones off the beat (Bouwer & Honing, 2015)), and real music (Tierney & Kraus, 2013). Similarly, the N1 response was found to be larger for events on a strong beat than for events on a weak beat in two studies with isochronous sequences on which listeners were instructed to impose a duple, triple, or quadruple meter (Fitzroy & Sanders, 2021; Schaefer et al., 2010). Thus, while the putative beat perception in isochronous sequences leads to attenuated responses as compared to in jittered sequences without a beat, at the same time, beat perception seems to enhance responses on the beat when compared to responses off the beat.

There are several explanations for this discrepancy. First, the effects of attention and prediction on early sensory processing are thought to be opposite, with the former leading to enhancement and the latter to attenuation (Lange, 2013). Possibly, the balance in the extent to which attentional processes and predictive processes related to beat perception are present differs depending on the type of sequence used. Another possibility is that whereas the contrast between isochronous and jittered sequences taps into processes associated with temporal expectations, the contrast between different metrical positions taps into process associated with hierarchical perception and grouping. Evidence for this idea comes from two studies that manipulated beat perception (e.g., in the temporal domain, the temporal regularity of the signal) while controlling for the grouping structure of non-isochronous rhythms (akin to Fig. 2, sequence 3A and 3B). In both these studies, sensory responses were attenuated for events on the beat in sequences with regularly spaced accents (e.g., with a beat, Fig. 2, example 3A) as compared to in sequences with irregular accents (e.g., without a beat, Fig. 2, example 3B (Bouwer et al., 2020; Schirmer et al., 2021)), even in the absence of attention (Bouwer et al., 2020). In one of these studies, ERPs and behavioral responses were measured separately for events on the beat and off the beat, both in the sequences with and without a beat. Of note, while the ERPs yielded no significant difference between events on and off the beat, behaviorally, there was an advantage for events on the beat, even for sequences without a regular beat, indicative of possible grouping effects (Bouwer et al., 2020).

To summarize, temporal expectations generally seem to lead to attenuation of the P1 and N1 component of the auditory-evoked potential, irrespective of the task relevance of a rhythmic sequence. In contrast, metrical structure may lead to enhancement of these components for metrically strong as compared to weak positions. This discrepancy may be explained by dissociating between temporal expectations, including beat perception, and expectations based on grouping and hierarchical structure.

In general, both studies examining responses to infrequent sounds (e.g., probing expectancy violations with oddball paradigms) and studies examining early sensory responses to frequent sounds have found consistent differences in ERP responses dependent on the presence of a regular beat. In many studies, these effects were shown to be independent of task relevance, and the effects were present in participants without specific musical training. These properties make ERPs an interesting candidate to probe beat perception in human newborns and nonhuman primates (Honing et al., 2014), which we will turn to in the next sections.

Measuring ERPs in Human Newborns

Birth is a special moment for research as it is the first time that the infants' nervous system is easily accessible to electrophysiological measurements, and a starting point of development with unfiltered auditory input (Lecanuet, 1996). However, the auditory system develops from the second trimester during pregnancy (Moore & Linthicum, 2007), and shows signs of discrimination of sounds even in utero (Huotilainen et al., 2005). Hence, birth cannot be taken as a sharp boundary between innate and learned abilities, albeit there is some evidence separating these abilities in preterm infants (Mahmoudzadeh et al., 2017). Due to the extremely rapid development of the auditory system during the first year, recordings from newborns are not only noisier than recordings from adults but also qualitatively different, lacking adult-obligatory components such as the P1 and N1 (Eggermont & Ponton, 2003). MMN-like ERP responses in newborns were first measured by Alho et al. (Alho et al., 1992). It is not yet clear whether the infants' responses are identical or only analogous to the adult MMN responses (Háden et al., 2016). Based on the ERP correlates of deviant-standard discrimination we can assume that auditory information that leads to discrimination in adults is also processed in the infants' brains. However, further processing steps are unclear. ERPs both negative and positive in polarity and within a wide variety of latency ranges from about 80 ms up to 500 ms were found in response to oddball designs (Virtala et al., 2022). With these caveats in mind, in the discussion below we will refer to these ERP responses found in newborns and young infants as mismatch responses (MMR).

Measuring ERPs in newborns is a technical and analytical challenge, not only because of the inherent noisiness of the signal but also due to the limited recording time usually available, the altered state of the infants that are mostly sleeping throughout the recording, and the limited number of channels used in newborn recordings. Fortunately, the use of high-density (64+) electrode nets became widespread, several preprocessing pipelines aim to address noise in recordings (Debnath et al., 2020; Fló et al., 2022; Gabard-Durnam et al., 2018; Kumaravel et al., 2022), and templates for more accurate source reconstruction have become available (O'Reilly et al., 2021). These advances allow for more fine-grained analyses of infantile auditory processing and better comparison with adult results. Such advances can also motivate the replication of basic results in the field.

Several abilities that underlie music perception seem to be functioning already at birth. Newborns are able to separate two sound streams based on sound frequency (Winkler et al., 2003) and detect pattern repetitions which they incorporate into their model of the auditory scene (Stefanics et al., 2007). Most important to beat perception is the ability to process temporal relations. The presentation of a stimulus earlier or later than expected in an isochronous sequence elicits an MMR in 10-month-old infants (Brannon et al., 2004), at least for large time intervals (500–1500 ms). Newborns are also sensitive to shorter changes (60–100 ms) in stimulus length (Čeponiené et al., 2002; Cheour et al., 2002), and 6-month-old infants detect even shorter gaps (4–16 ms) inserted in tones (Trainor et al., 2001, 2003), showing the remarkable temporal resolution of the auditory system. Háden et al. showed that newborns are sensitive to changes in the presentation rate of the stimulation, can detect the beginning of sound trains, and react to the omission of expected stimuli (Háden et al., 2012). Furthermore, there are indications that newborns can learn hierarchical rules (Moser et al., 2020), and can integrate contextual information in their predictions about future events over both shorter (Háden et al., 2015) and longer time periods (Todd et al., 2022). Some of the abilities that reflect the general organization of temporal pattern processing in the brain may be present even before term birth. Preterm newborn infants were shown to exhibit an MMR to earlier than expected tones in a non-isochronous rhythmic pattern in duple meter (Edalati et al., 2022). Thus, the infant brain, even preterm, can detect rhythmic pattern violations. Dynamic causal modeling (DCM) of the MMR revealed extensive top-down and bottom-up connections between the auditory cortices and temporal structures on both sides, and right frontal areas (Edalati et al., 2022), similar to the network found in adults (Phillips et al., 2015). Taken together, these results indicate that investigating phenomena reliant on temporal processing (e.g., beat and meter perception) is viable in (newborn) infants.

Using MMR to Probe Beat Perception in Human Newborns

Several studies to date have examined beat perception in newborns using MMR as an index of temporal expectations. One study examined processing of unexpected sounds within natural language that was presented either spoken, sung, or rhythmically recited to a strong beat at about 2 Hz, as intended for a nursery rhyme (Suppanen et al., 2019). Deviants in the form of changes in words, vowels, sound intensity, or pitch were introduced on stressed syllables (e.g., on the beat). MMR to vowel and word changes was only elicited in the rhythmic nursery rhyme condition. The enhancement of MMR in a rhythmic context is reminiscent of the larger oddball responses to deviants in isochronous than jittered sequences found in adults (Schwartze et al., 2011). Interestingly, only vowel and word changes elicited an MMR, but not intensity and pitch changes. This may have been due to the collation of responses for all intensity and pitch changes, including intensity increments as well as decrements, which in adults can lead to opposite results (Bouwer & Honing, 2015). However, these results may also underline the importance of context, in this case linguistic, on the processing of acoustic deviants, and raises the question whether processing of linguistic stimuli may be privileged even at birth (Thiede et al., 2019).

Two studies have looked at differences in MMR on and off the beat in newborns. First, Winkler et al. tested whether newborns can extract a regular beat from a varying rhythmic stimulus (Winkler et al., 2009), using a paradigm previously used in adults to probe meter perception (Ladinig et al., 2009, 2011). Newborns were presented with a drum pattern in duple meter, in which sounds on the first beat (e.g., the strongest metrical position in the pattern) were occasionally omitted. The response to these omissions was compared to the response to omissions off the beat, and to the response to omissions in a control sequence consisting of patterns in which sounds on the first beat were always omitted. The ERP responses to the omissions on the beat differed significantly from responses to patterns without omission, omissions off the beat, and omissions in the control sequence. The results were interpreted as evidence for the presence of ability to detect a beat in newborns. However, the omission on the beat differed from the omissions off the beat in multiple ways, including differences in acoustic context, and differences in the transitional probabilities of the omitted sounds. Therefore, the results of this study could have been biased by indexing not just beat perception but also low-level acoustic differences between conditions, and sequential learning (Bouwer et al., 2014).

To control for these possible confounds, a subsequent study (Háden et al., 2024) used a paradigm previously used to probe beat perception in human adults (Bouwer et al., 2016) and nonhuman primates (Honing et al., 2018). Newborns were presented with a drum rhythm with alternating accented and unaccented sounds that induce a beat (or duple meter) when presented with isochronous timing, but not when presented with randomly jittered timing (Fig. 2, example 2). Infrequently, softer sounds were introduced as deviants, falling either on the beat or off the beat. Deviants were always preceded and followed by identical sounds, to control for the effects of acoustic context on ERPs (see Fig. 2, example 2). Results showed a clear difference in MMR amplitude between metrical positions in the isochronous sequence, but not in the equivalent jittered sequence (Fig. 3b). However, the current paradigm could not show effects of statistical learning (e.g., a difference in responses on and off the beat for the jittered sequences), despite previous evidence for this ability working in newborns (Bosseler et al., 2016), and the presence of this effect in adults using the same paradigm (Bouwer et al., 2016). Despite the qualitative differences between an adult MMN and the newborn MMR, these results provide converging evidence that beat processing is present in newborns infants, even when controlling for acoustic context and statistical learning.

Measuring ERPs in Nonhuman Primates

There is quite some discussion on whether beat perception is species-specific (Fitch, 2015; Ravignani, 2018; Wilson & Cook, 2016). Evidence in support of beat perception in a select number of species comes from experiments that test motor entrainment to a beat through overt behavior (A. D. Patel, 2021). However, if the production of synchronized movement to sound or music is not observed in a species, this is no evidence for the absence of beat perception. It could well be that certain animals are simply not able to synchronize their movements to a varying rhythm, while they can perceive a beat. Also, with behavioral methods that rely on overt motoric responses, it is difficult to separate between the contribution of perception and action. Electrophysiological measures, such as ERPs, do not require an overt response, and as such provide an attractive alternative to probe beat perception in animals (Honing et al., 2018).

While most animal studies have used implanted electrodes to record electroencephalograms (EEG) (Javitt et al., 1994; Laughlin et al., 1999; Pincze et al., 2001), noninvasive electrophysiological techniques such as scalp-recorded evoked potentials (EP) and event-related potentials (ERP) are considered an attractive alternative. Next to being a mandatory requirement for studying some nonhuman primates such as chimpanzees (Fukushima et al., 2010; Hirata et al., 2013), these methods allow for a direct comparison between human and nonhuman primates. As such they have contributed to establishing animal models of the human brain and human brain disorders (Gil-da-Costa et al., 2013; Godlove et al., 2011), a better understanding of the neural mechanisms underlying the generation of human evoked EP/ERP components (Fishman & Steinschneider, 2012), as well as delineating cross-species commonalities and differences in brain functions, including rhythm perception and cognition (Fukushima et al., 2010; Hirata et al., 2013; Itoh et al., 2015; Reinhart et al., 2012; Ueno et al., 2008, 2009). The most relevant ERP components for comparative primate studies of rhythm perception are summarized in Table 1.

Since the discovery of the MMN component, researchers have tried to find analogous processes in animal models (Shiramatsu & Takahashi, 2021; Woodman, 2011) and to integrate deviance detection and predictive processing into a general framework of auditory perception (Näätänen et al., 2010). A wide range of electrophysiological methods from scalp electrodes to single-cell recordings have been used on animal models. These methods highlight different phenomena of varying spatial and temporal resolution. The most vital difference is that scalp and epidural recordings may yield components similar to the human MMN (i.e. electric responses generated by large brain areas), whereas local field potential, multiunit activity, and single-cell recordings work on a lower spatial scale and reflect stimulus-specific adaptation (Nelken & Ulanovsky, 2007). SSA has many common properties with MMN; both can be observed in similar paradigms, and it is still debated whether SSA reflects the cellular level activity underlying MMN. However, this discussion is beyond the scope of the current chapter.

Using epidural recording, MMN-like responses (from here on referred to as MMR) have been shown in different species including rats (Nakamura et al., 2011), cats (Csépe et al., 1987; Pincze et al., 2001, 2002) and macaque monkeys (Javitt et al., 1992, 1994). In most of these studies, frequency and amplitude violations were used. In rats, deviance detection was shown for both a temporal feature, sound duration (Nakamura et al., 2011), and to an abstract feature, namely melodic contour (Ruusuvirta et al., 2007). Recordings from scalp electrodes showed MMR in mice (Umbricht et al., 2005), and in a single chimpanzee (Ueno et al., 2008). While not all attempts at recording MMR from animals were successful, it seems that an MMR can be reliably elicited in some animal models (Harms et al., 2016; Schall et al., 2015; Shiramatsu & Takahashi, 2021) and thus can be used to study auditory processing in nonhuman primates.

Table 1 Homologies between rhesus monkey, chimpanzee, and human cortical auditory-evoked potentials (ERPs). Time range in ms; alternative naming in square brackets

	Human *scalp* (Picton et al., 1974)	Ape *scalp* (Ueno et al., 2008)	Monkey *scalp* (Honing et al. 2012)	Monkey *scalp* (Gil-da-Costa et al. 2013)	Monkey *scalp* (Itoh et al. 2015)	Monkey *scalp* (Honing et al. 2018)	Monkey *cranial* (Teichert, 2016)	Monkey *epidural* (Javitt et al., 2000)
P1	50–60		–	–	25–30 [mP1]	20–40-	45–65 [P55]	5–40
N1	75–100		–	–	45–65 [mN1]	40–60	70–105 [N85]	40–120
MMR	100–200	125–180	60–110	48–120	–	60–125	–	–
P3a	200–250		–	100–250	–	125–225	–	–

Using MMR to Probe Beat Perception in Nonhuman Primates

Honing et al. (Honing et al., 2012) demonstrated, for the first time, that an MMR can be recorded from the scalp in rhesus monkeys *(Macaca mulatta)*, both for pitch deviants and unexpected omissions. Ueno et al. (Ueno et al., 2008) used a similar method in a chimpanzee (*Pan troglodytes*) and Gil-da-Costa et al. (Gil-da-Costa et al., 2013) made a comparison between measuring an MMR in humans and macaques (*Macaca fascicularis*). Together these results provide support for the idea that a mismatch response can be used as an index of the detection of expectancy violations in an auditory signal in both humans and nonhuman primates. A follow-up study, using stimuli and an experimental paradigm identical to those used to study beat perception in human adults (Bouwer et al., 2016) and newborns (Háden et al., 2024), confirmed that rhesus monkeys are sensitive to the isochrony of a rhythmic sequence, but not to its induced beat (Honing et al., 2018). Results from the two monkeys in this study are depicted in Fig. 3c. These findings are in line with the hypothesis that beat perception is somewhat species-specific. Note that while rhesus monkeys continue to be an important animal model for the human brain, and results in monkeys and humans are often compared (Gil-da-Costa et al., 2013), we have to be cautious in directly comparing ERP signals from humans and nonhuman animals because of obvious differences in neural architecture.

Behaviorally, contrary to what was previously thought (Zarco et al., 2009), macaques do show the ability to predictively tap to a metronome, and to modify their tapping tempo to tempo changes in the entraining stimulus, when provided with sufficient feedback and reward (Gámez et al., 2018). In addition, and consistent with these behavioral results, it was shown that during isochronous tapping, the medial premotor cortex in monkeys indexes time intervals in a relative and predictive manner (Betancourt et al., 2023; Gámez et al., 2019). But note that processing isochrony is not the same as beat perception, and may be subserved by a different mechanism (Bouwer et al., 2021; Honing et al., 2018). For an overview of time encoding in the primate medial premotor cortex, see Merchant et al., this volume.

Overall, the observed differences between humans and monkeys provide support for the gradual audiomotor evolution (GAE) hypothesis (Honing et al., 2018; Honing & Merchant, 2014; Merchant & Honing, 2014). This hypothesis suggests beat-based timing to be more developed in humans as opposed to apes and monkeys, and that it evolved through a gradual chain of anatomical and functional changes to the interval-based mechanism to generate an additional beat-based mechanism. More specifically, the integration of sensorimotor information throughout the mCBGT circuit and other brain areas during the perception or execution of single intervals is similar in human and nonhuman primates, but different in the processing of multiple intervals (Merchant & Honing, 2014). While the mCBGT circuit was shown to be also involved in beat-based mechanisms in brain imaging studies (e.g., (Teki et al., 2011)), direct projections from the medial premotor cortex (MPC) to the primary auditory cortex (A1) via the inferior parietal lobe (IPL) that is involved in sensory and cognitive functions such as attention and spatial sense, may be the underpinning of beat-based timing as found in humans, and possibly apes (Merchant & Honing, 2014; Proksch et al., 2020).

Probing beat perception and isochrony perception in animals is still in its infancy (Bouwer et al., 2021; Henry et al., 2021; Wilson & Cook, 2016). But it appears, at least within the primate lineage, that beat perception has evolved gradually, peaking in humans and present only with limitations in chimpanzees (Hattori & Tomonaga, 2020), bonobos (Large & Gray, 2015), macaques (Honing et al., 2018), and other nonhuman primates (Raimondi et al., 2023).

While beat perception can be argued to be fundamental to the capacity for music (i.e. musicality (Honing, 2012)), it continues to be difficult to trace back this skill in the animal world. In the few species that are studied, it appears to be mostly vocal learners that are sensitive to a regular pulse (the beat) in a varying rhythmic stimulus such as music. Seminal examples are a sulphur-crested cockatoo (A. D. Patel et al., 2009) and a gray parrot (Schachner et al., 2009) that are capable of synchronizing to the beat of human music and, importantly, maintaining synchrony when the same music is played at a different tempo. The observation that this behavior was initially only shown in vocal learning species gave rise to the vocal learning and rhythmic synchronization (VLRS) hypothesis (A. D. Patel, 2006, 2021), suggesting our ability to move in time with an auditory beat in a precise, predictive, and tempo-flexible manner originated in the neural circuitry for complex vocal learning. This hypothesis is an alternative to the GAE hypothesis discussed earlier.

However, the gradual audiomotor evolution (GAE) and vocal learning (VLRS) hypotheses differ in several ways (see also Proksch et al., 2020). First, the GAE hypothesis does not claim that the neural circuit that is engaged in rhythmic entrainment is deeply linked to vocal perception, production, and learning, even if some overlap between the circuits exists. Furthermore, since the *cortico-basal ganglia-thalamic* circuit (CBGT) has been involved in beat-based mechanisms in imaging studies, we suggest that the reverberant flow of audiomotor information that loops across the anterior prefrontal CBGT circuits maybe the underpinning of human rhythmic entrainment. Lastly, the GAE hypothesis suggests that the integration of sensorimotor information throughout the mCBGT circuit and other brain areas during the perception or execution of single intervals is similar in human and nonhuman primates.

In addition, a recent counterexample to the VLRS hypothesis is a California sea lion (*Zalophus californianus*; not considered a vocal learner) that is able to synchronize head movements to a variety of musical fragments, as well as showing generalization over different tempi (Cook et al., 2013; Rouse et al., 2016). Overall, it seems that perceiving a beat in a complex stimulus (i.e. music) and being able to synchronize to it is not restricted to humans, might well be more widespread than previously thought, and not restricted to vocal learners per se (Bouwer et al., 2021; ten Cate & Honing, 2023; Wilson & Cook, 2016).

Discussion and Conclusion

In this chapter, we have shown how ERPs can be used to probe the perception of a regular beat in rhythm. Measuring ERPs is relatively straightforward, it can be realized in populations that are difficult to study behaviorally (like infants and monkeys), and it is a well-researched method. However, several challenges remain, for beat perception research in general, and for ERP studies in particular.

First, as we have stressed throughout this chapter, musical rhythm contains many types of structure, including not only temporal structure, both in terms of a regular beat and absolute temporal intervals, but also grouping, ordinal structure, and hierarchy. The beat can be considered the most prominent periodicity in a rhythmic signal (Fiveash et al., 2022), and beat perception has been considered as the ability to flexibly extract a regular temporal structure from rhythm (Penhune & Zatorre, 2019). Such definitions of the beat clearly involve the temporal aspect of rhythm, and specifically the temporal periodicity associated with beat-based perception. For many studies targeting beat perception with ERPs, it is not completely clear whether influences of absolute timing, grouping, ordinal structure, and hierarchical structure can be ruled out, as these structural aspects of rhythm often covary with the temporal regularity that is the beat, and are often even necessary to induce a beat.

Related to this, some have suggested that the perception of hierarchical metrical structure is different from the perception of a beat or pulse as temporal regularity (Fitch, 2013; Silva & Castro, 2019). The idea that meter processing is indeed more about hierarchical structure, or the alternation of stressed and unstressed events, than about temporal regularity is in line with models of meter in language, where the meter does not necessarily adhere to a temporal regularity. In language, learning the alternation of stressed (e.g., salient) and unstressed sounds is vital to processing (Henrich et al., 2014; Henrich & Scharinger, 2022; Magne et al., 2016), and the hierarchical structure that arises from such nontemporal structure is often termed meter. This is, however, at odds with models of beat perception that consider beat and meter to be interrelated, with meter perception relying on similar (oscillatory) mechanisms as beat perception (e.g., meter in this interpretation is just another level of regularity within a structure of multiple nested levels of regularity) (Drake et al., 2000; Large, 2008). The relationship between the different aspects of rhythm perception, and specifically the relationship between beat perception and hierarchy perception, remains an important topic for future studies.

One disadvantage of using ERPs to study beat perception is that with ERPs, what is probed is not the mechanism of beat perception itself, but rather the effect a perceived beat has on the sensory processing of incoming information, be it expected or unexpected tones. Combining results from ERP studies with results from studies that directly probe the underlying mechanisms of beat perception, for example, by examining low-frequency neural oscillations in response to rhythm (Lenc et al., 2021), will provide more insight in this regard. Also, the studies discussed in this chapter mostly deal with purely perceptual effects of beat perception. While some studies have used ERPs in studying motor synchronization to a beat (Andrea-Penna et al., 2020; Conradi et al., 2016; Lei et al., 2019; Mathias et al., 2020; Schwartze & Kotz, 2015), given the tight coupling between beat perception and movement, this remains an interesting topic for future work. Ultimately, combining different methods and paradigms will allow us to get a more coherent picture of the perception of beat and meter, and address its apparent innateness, and domain and species specificity. All in all, this research will contribute to a better understanding of the fundamental role that beat and meter perception play in music.

Funding FB is supported by Veni Grant VI. Veni.201G.066 awarded by the Dutch Research Council NWO. GPH is supported by the Hungarian National Research Development and Innovation Office (FK139135), the János Bolyai Research Grant (BO/00523/21/2) and the New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation (ÚNKP-23-5-BME-429). HH is supported by an Open Competition Grant (406.20.CW.002) awarded by the Dutch Research Council NWO.

References

Abecasis, D., Brochard, R., Del Río, D., Dufour, A., & Ortiz, T. (2009). Brain lateralization of metrical accenting in musicians. *Annals of the New York Academy of Sciences, 1169*, 74–78. https://doi.org/10.1111/j.1749-6632.2009.04766.x

Abecasis, D., Brochard, R., Granot, R., & Drake, C. (2005). Differential brain response to metrical accents in isochronous auditory sequences. *Music Perception, 22*(3), 549–562. https://doi.org/10.1525/mp.2005.22.3.549

Alho, K., Woods, D. L., Algazi, A., & Näätänen, R. N. (1992). Intermodal selective attention. II. Effects of attentional load on processing of auditory and visual stimuli in central space. *Electroencephalography and Clinical Neurophysiology, 82*(5), 356–368. https://doi.org/10.1016/0013-4694(92)90005-3

Andrea-Penna, G. M. D., Iversen, J. R., Chiba, A. A., Khalil, A. K., & Minces, V. H. (2020). One tap at a time: Correlating sensorimotor synchronization with brain signatures of temporal processing. *Cerebral Cortex Communications, 1*(1), 1–8. https://doi.org/10.1093/texcom/tgaa036

Andreou, L.-V., Griffiths, T. D., & Chait, M. (2015). Sensitivity to the temporal structure of rapid sound sequences—An MEG study. *NeuroImage, 110*, 194–204. https://doi.org/10.1016/j.neuroimage.2015.01.052

Bååth, R. (2015). Subjective rhythmization: A replication and an assessment of two theoretical explanations. *Music Perception, 33*(2), 244–254. https://doi.org/10.1525/mp.2015.33.2.244

Bauer, A.-K. R., Jaeger, M., Thorne, J. D., Bendixen, A., & Debener, S. (2015). The auditory dynamic attending theory revisited: A closer look at the pitch comparison task. *Brain Research, 1626*, 198–210. https://doi.org/10.1016/j.brainres.2015.04.032

Betancourt, A., Pérez, O., Gámez, J., Mendoza, G., & Merchant, H. (2023). Amodal population clock in the primate medial premotor system for rhythmic tapping. *Cell Reports, 42*(10), 113234. https://doi.org/10.1016/j.celrep.2023.113234

Bolger, D., Trost, W. J., & Schön, D. (2013). Rhythm implicitly affects temporal orienting of attention across modalities. *Acta Psychologica, 142*(2), 238–244. https://doi.org/10.1016/j.actpsy.2012.11.012

Bosseler, A. N., Teinonen, T., Tervaniemi, M., & Huotilainen, M. (2016). Infant directed speech enhances statistical learning in newborn infants: An ERP study. *PLoS One, 11*(9), e0162177. https://doi.org/10.1371/journal.pone.0162177

Bouwer, F. L. (2022). Neural entrainment to auditory rhythms: Automatic or top-down driven? *Journal of Neuroscience, 42*(11), 2146–2148. https://doi.org/10.1523/JNEUROSCI.2305-21.2022

Bouwer, F. L., Burgoyne, J. A., Odijk, D., Honing, H., & Grahn, J. A. (2018). What makes a rhythm complex? The influence of musical training and accent type on beat perception. *PLoS One, 13*(1), e0190322. https://doi.org/10.1371/journal.pone.0190322

Bouwer, F. L., Fahrenfort, J. J., Millard, S. K., Kloosterman, N. A., & Slagter, H. A. (2023). A silent disco: Differential effects of beat-based and pattern-based temporal expectations on persistent entrainment of low-frequency neural oscillations. *Journal of Cognitive Neuroscience, 35*(6), 990–1020. https://doi.org/10.1162/jocn_a_01985

Bouwer, F. L., & Honing, H. (2015). Temporal attending and prediction influence the perception of metrical rhythm: Evidence from reaction times and ERPs. *Frontiers in Psychology, 6*(July), 1094. https://doi.org/10.3389/fpsyg.2015.01094

Bouwer, F. L., Honing, H., & Slagter, H. A. (2020). Beat-based and memory-based temporal expectations in rhythm: Similar perceptual effects, different underlying mechanisms. *Journal of Cognitive Neuroscience, 32*(7), 1221–1241. https://doi.org/10.1162/jocn_a_01529

Bouwer, F. L., Nityananda, V., Rouse, A. A., & ten Cate, C. (2021). Rhythmic abilities in humans and non-human animals: A review and recommendations from a methodological perspective. *Philosophical Transactions of the Royal Society B, 376*, 20200335. https://doi.org/10.1098/rstb.2020.0335

Bouwer, F. L., Van Zuijen, T. L., & Honing, H. (2014). Beat processing is pre-attentive for metrically simple rhythms with clear accents: An ERP study. *PLoS One, 9*(5), e97467. https://doi.org/10.1371/journal.pone.0097467

Bouwer, F. L., Werner, C. M., Knetemann, M., & Honing, H. (2016). Disentangling beat perception from sequential learning and examining the influence of attention and musical abilities on ERP responses to rhythm. *Neuropsychologia, 85*, 80–90. https://doi.org/10.1016/j.neuropsychologia.2016.02.018

Brannon, E. M., Roussel, L. W., Meck, W. H., & Woldorff, M. (2004). Timing in the baby brain. *Cognitive Brain Research, 21*(2), 227–233. https://doi.org/10.1016/j.cogbrainres.2004.04.007

Breska, A., & Deouell, L. Y. (2017). Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLoS Biology, 15*(2), e2001665. https://doi.org/10.1371/journal.pbio.2001665

Brinkmann, P., Rigoulot, S., Kadi, M., Schwartze, M., Kotz, S. A., & Bella, S. D. (2021). About time: Ageing influences neural markers of temporal predictability. *Biological Psychology, 163*, 108135. https://doi.org/10.1016/j.biopsycho.2021.108135

Brochard, R., Abecasis, D., Potter, D. D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science, 14*(4), 362–366. https://doi.org/10.1111/1467-9280.24441

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience, 6*(10), 755–765. https://doi.org/10.1038/nrn1764

Cannon, J. J. (2021). Expectancy-based rhythmic entrainment as continuous Bayesian inference. *PLoS Computational Biology, 17*(6), e1009025. https://doi.org/10.1371/journal.pcbi.1009025

Cannon, J. J., & Patel, A. D. (2021). How beat perception co-opts motor neurophysiology. *Trends in Cognitive Sciences, 25*(2), 137–150. https://doi.org/10.1016/j.tics.2020.11.002

Čeponiené, R., Kushnerenko, E., Fellman, V., Renlund, M., Suominen, K., & Näätänen, R. N. (2002). Event-related potential features indexing central auditory discrimination by newborns. *Cognitive Brain Research, 13*(1), 101–113. https://doi.org/10.1016/S0926-6410(01)00093-3

Chang, A., Li, Y. C., Chan, J. F., Dotov, D. G., Cairney, J., & Trainor, L. J. (2021). Inferior auditory time perception in children with motor difficulties. *Child Development, 92*(5), 1–17. https://doi.org/10.1111/cdev.13537

Cheour, M., Čèponiené, R., Leppänen, P., Alho, K., Kujala, T., Renlund, M., Fellman, V., & Näätänen, R. N. (2002). The auditory sensory memory trace decays rapidly in newborns. *Scandinavian Journal of Psychology, 43*, 33–39. https://doi.org/10.1111/1467-9450.00266

Clarke, E. F. (2005). *Ways of listening: An ecological approach to the perception of musical meaning*. Oxford University Press.

Comerchero, M. D., & Polich, J. (1999). P3a and P3b from typical auditory and visual stimuli. *Clinical Neurophysiology, 110*(1), 24–30. https://doi.org/10.1016/S0168-5597(98)00033-1

Conradi, N., Abel, C., Frisch, S., Kell, C. A., Kaiser, J., & Schmidt-Kassow, M. (2016). Actively but not passively synchronized motor activity amplifies predictive timing. *NeuroImage, 139*, 211–217. https://doi.org/10.1016/j.neuroimage.2016.06.033

Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human primates. *Trends in Cognitive Sciences, 5*(12), 539–546.

Cook, P. F., Rouse, A., Wilson, M., & Reichmuth, C. (2013). A California sea lion (Zalophus californianus) can keep the beat: Motor entrainment to rhythmic auditory stimuli in a non vocal mimic. *Journal of Comparative Psychology, 127*(2), 412–427. https://doi.org/10.1037/a0032345

Cooper, G., & Meyer, L. B. (1960). *The rhythmic structure of music*. University of Chicago Press.

Criscuolo, A., Schwartze, M., Henry, M. J., Obermeier, C., & Kotz, S. A. (2023). Individual neurophysiological signatures of spontaneous rhythm processing. *NeuroImage, 273*, 120090. https://doi.org/10.1016/j.neuroimage.2023.120090

Csépe, V., Karmos, G., & Molnár, M. (1987). Evoked potential correlates of stimulus deviance during wakefulness and sleep in cat -animal model of mismatch negativity. *Electroencephalography and Clinical Neurophysiology, 66*(6), 571–578. https://doi.org/10.1016/0013-4694(87)90103-9

Damm, L., Varoqui, D., De Cock, V. C., Bella, S. D., & Bardy, B. (2019). Why do we move to the beat?

A multi-scale approach, from physical principles to brain dynamics. *Neuroscience & Biobehavioral Reviews, 112*, 553–584. https://doi.org/10.1016/J.NEUBIOREV.2019.12.024

Debnath, R., Buzzell, G. A., Morales, S., Bowers, M. E., Leach, S. C., & Fox, N. A. (2020). The Maryland analysis of developmental EEG (MADE) pipeline. *Psychophysiology, 57*(6), e13580. https://doi.org/10.1111/psyp.13580

Desain, P., & Honing, H. (1999). Computational models of beat induction: The rule-based approach. *Journal of New Music Research, 28*(1), 29–42. https://doi.org/10.1076/jnmr.28.1.29.3123

Desain, P., & Honing, H. (2003). The formation of rhythmic categories and metric priming. *Perception, 32*(3), 341–365. https://doi.org/10.1068/p3370

Drake, C., Jones, M. R., & Baruch, C. (2000). The development of rhythmic attending in auditory sequences: Attunement, referent period, focal attending. *Cognition, 77*(3), 251–288. https://doi.org/10.1016/S0010-0277(00)00106-2

Edalati, M., Mahmoudzadeh, M., Ghostine, G., Kongolo, G., Safaie, J., Wallois, F., & Moghimi, S. (2022). Preterm neonates distinguish rhythm violation through a hierarchy of cortical processing. *Developmental Cognitive Neuroscience, 58*, 101168. https://doi.org/10.1016/j.dcn.2022.101168

Edalati, M., Mahmoudzadeh, M., Safaie, J., Wallois, F., & Moghimi, S. (2021). Violation of rhythmic expectancies can elicit late frontal gamma activity nested in theta oscillations. *Psychophysiology, 58*(11), 1–17. https://doi.org/10.1111/psyp.13909

Eggermont, J. J., & Ponton, C. W. (2003). Auditory-evoked potential studies of cortical maturation in normal hearing and implanted children: Correlations with changes in structure and speech perception. *Acta Oto-Laryngologica, 123*(2), 249–252. https://doi.org/10.1080/0036554021000028098

Ellis, R. J., & Jones, M. R. (2009). The role of accent salience and joint accent structure in meter perception. *Journal of Experimental Psychology: Human Perception and Performance, 35*(1), 264–280. https://doi.org/10.1037/a0013482

Fishman, Y. I., & Steinschneider, M. (2012). Searching for the mismatch negativity in primary auditory cortex of the awake monkey: Deviance detection or stimulus specific adaptation? *Journal of Neuroscience, 32*(45), 15747–15758. https://doi.org/10.1523/JNEUROSCI.2835-12.2012

Fitch, W. T. (2013). Rhythmic cognition in humans and animals: Distinguishing meter and pulse perception. *Frontiers in Systems Neuroscience, 7*(October), 68. https://doi.org/10.3389/fnsys.2013.00068

Fitch, W. T. (2015). Four principles of bio-musicology. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*(1664), 20140091. https://doi.org/10.1098/rstb.2014.0091

Fitzgerald, P., & Picton, T. (1983). Event-related potentials recorded during the discrimination of improbable stimuli. *Biological Psychology, 17*(4), 241–276. https://doi.org/10.1016/0301-0511(83)90003-0

Fitzroy, A. B., & Sanders, L. D. (2015). Musical meter modulates the allocation of attention across time. *Journal of Cognitive Neuroscience, 27*(12), 2339–2351. https://doi.org/10.1162/jocn

Fitzroy, A. B., & Sanders, L. D. (2021). Subjective metric organization directs the allocation of attention across time. *Auditory Perception & Cognition, 3*(1), 1–26. https://doi.org/10.1080/25742442.2021.1898924

Fiveash, A., Bella, S. D., Bigand, E., Gordon, R. L., & Tillmann, B. (2022). You got rhythm , or more: The multidimensionality of rhythmic abilities. *Attention, Perception & Psychophysics, 84*, 1370–1392.

Fiveash, A., Ferreri, L., Bouwer, F. L., Kösem, A., Moghimi, S., Ravignani, A., Keller, P. E., & Tillmann, B. (2023). Can rhythm-mediated reward boost learning, memory, and social connection? Perspectives for future research. *Neuroscience & Biobehavioral Reviews, 149*, 105153. https://doi.org/10.1016/j.neubiorev.2023.105153

Fló, A., Gennari, G., Benjamin, L., & Dehaene-Lambertz, G. (2022). Automated pipeline for infants continuous EEG (APICE): A flexible pipeline for developmental cognitive studies. *Developmental Cognitive Neuroscience, 54*, 101077. https://doi.org/10.1016/j.dcn.2022.101077

Foldal, M. D., Blenkmann, A. O., Llorens, A., Knight, R. T., Solbakk, A.-K., & Endestad, T. (2020). The brain tracks auditory rhythm predictability independent of selective attention. *Scientific Reports, 10*(1), 7975. https://doi.org/10.1038/s41598-020-64758-y

Fraisse, P. (1982). Rhythm and tempo. In D. Deutsch (Ed.), *Psychology of Music* (pp. 149–180). Academic Press.

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

Fukushima, H., Hirata, S., Ueno, A., Matsuda, G., Fuwa, K., Sugama, K., Kusunoki, K., Hirai, M., Hiraki, K., Tomonaga, M., & Hasegawa, T. (2010). Neural correlates of face and object perception in an awake chimpanzee (pan troglodytes) examined by scalp-surface event-related potentials. *PLoS One, 5*(10), e13366. https://doi.org/10.1371/journal.pone.0013366

Gabard-Durnam, L. J., Mendez Leal, A. S., Wilkinson, C. L., & Levin, A. R. (2018). The Harvard automated processing pipeline for electroencephalography (HAPPE): Standardized processing software for developmental and high-artifact data. *Frontiers in Neuroscience, 12*, 97. https://doi.org/10.3389/fnins.2018.00097

Gámez, J., Mendoza, G., Prado, L., Betancourt, A., & Merchant, H. (2019). The amplitude in periodic neural state trajectories underlies the tempo of rhythmic tapping. *PLoS Biology, 17*(4), e3000054. https://doi.org/10.1371/journal.pbio.3000054

Gámez, J., Yc, K., Ayala, Y. A., Dotov, D., Prado, L., & Merchant, H. (2018). Predictive rhythmic tapping to isochronous and tempo changing metronomes in the nonhuman primate. *Annals of the New York Academy of Sciences, 1423*, 396–414. https://doi.org/10.1111/nyas.13671

Geiser, E., Sandmann, P., Jäncke, L., & Meyer, M. (2010). Refinement of metre perception—Training increases hierarchical metre processing. *The European Journal of Neuroscience, 32*(11), 1979–1985. https://doi.org/10.1111/j.1460-9568.2010.07462.x

Geiser, E., Ziegler, E., Jancke, L., Meyer, M., Jäncke, L., Meyer, M., Jancke, L., Meyer, M., Jäncke, L., & Meyer, M. (2009). Early electrophysiological correlates of meter and rhythm processing in music perception. *Cortex, 45*(1), 93–102. https://doi.org/10.1016/j.cortex.2007.09.010

Gerry, D. W., Faux, A. L., & Trainor, L. J. (2010). Effects of Kindermusik training on infants' rhythmic enculturation. *Developmental Science, 13*(3), 545–551. https://doi.org/10.1111/j.1467-7687.2009.00912.x

Gibson, J. J. (1975). Events are perceivable but time is not. In J. T. Fraser & N. M. Lawrence (Eds.), *The study of time II* (pp. 295–301). Springer Verlag.

Gil-da-Costa, R., Stoner, G. R., Fung, R., & Albright, T. D. (2013). Nonhuman primate model of schizophrenia using a noninvasive EEG method. *Proceedings of the National Academy of Sciences, 110*(38), 15425–15430. https://doi.org/10.1073/pnas.1312264110

Godlove, D. C., Emeric, E. E., Segovis, C. M., Young, M. S., Schall, J. D., & Woodman, G. F. (2011). Event-related potentials elicited by errors during the stop-signal task I. Macaque monkeys. *Journal of Neuroscience, 31*(44), 15640–15649. https://doi.org/10.1523/JNEUROSCI.3349-11.2011

Grahn, J. A. (2009). The role of the basal ganglia in beat perception: Neuroimaging and neuropsychological investigations. *Annals of the New York Academy of Sciences, 1169*, 35–45. https://doi.org/10.1111/j.1749-6632.2009.04553.x

Grahn, J. A., & Brett, M. (2007). Rhythm and beat perception in motor areas of the brain. *Journal of Cognitive Neuroscience, 19*(5), 893–906. https://doi.org/10.1162/jocn.2007.19.5.893

Grube, M., & Griffiths, T. D. (2009). Metricality-enhanced temporal encoding and the subjective perception of rhythmic sequences. *Cortex, 45*(1), 72–79. https://doi.org/10.1016/j.cortex.2008.01.006

Háden, G. P., Bouwer, F. L., Honing, H., & Winkler, I. (2024). Beat processing in newborn infants cannot be explained by statistical learning based on transition probabilities. *Cognition, 243*, 105670. https://doi.org/10.1016/j.cognition.2023.105670

Háden, G. P., Honing, H., & Winkler, I. (2012). Newborn infants are sensitive to sound timing. In *12th International conference on music perception and cognition*, Thessaloniki, Greece, pp. 378–379.

Háden, G. P., Németh, R., Török, M., & Winkler, I. (2015). Predictive processing of pitch trends in newborn infants. *Brain Research, 1626*, 14–20. https://doi.org/10.1016/j.brainres.2015.02.048

Háden, G. P., Németh, R., Török, M., & Winkler, I. (2016). Mismatch response (MMR) in neonates: Beyond refractoriness. *Biological Psychology, 117*, 26–31. https://doi.org/10.1016/j.biopsycho.2016.02.004

Haegens, S., & Zion Golumbic, E. (2018). Rhythmic facilitation of sensory processing: A critical review. *Neuroscience and Biobehavioral Reviews, 86*(March), 150–165. https://doi.org/10.1016/j.neubiorev.2017.12.002

Hannon, E. E., Snyder, J. S., Eerola, T., & Krumhansl, C. L. (2004). The role of melodic and temporal cues in perceiving musical meter. *Journal of Experimental Psychology: Human Perception and Performance, 30*(5), 956–974. https://doi.org/10.1037/0096-1523.30.5.956

Hannon, E. E., & Trehub, S. E. (2005). Metrical categories in infancy and adulthood. *Psychological Science, 16*(1), 48–55. https://doi.org/10.1111/j.0956-7976.2005.00779.x

Harms, L., Michie, P. T., & Näätänen, R. N. (2016). Criteria for determining whether mismatch responses exist in animal models: Focus on rodents. *Biological Psychology, 116*, 28–35. https://doi.org/10.1016/j.biopsycho.2015.07.006

Haroush, K., Hochstein, S., & Deouell, L. Y. (2010). Momentary fluctuations in allocation of attention: Cross-modal effects of visual task load on auditory discrimination. *Journal of Cognitive Neuroscience, 22*(7), 1440–1451. https://doi.org/10.1162/jocn.2009.21284

Hattori, Y., & Tomonaga, M. (2020). Rhythmic swaying induced by sound in chimpanzees (Pan troglodytes). *Proceedings of the National Academy of Sciences of the United States of America, 117*(2), 936–942. https://doi.org/10.1073/pnas.1910318116

Haumann, N. T., Vuust, P., Bertelsen, F., & Garza-Villarreal, E. A. (2018). Influence of musical enculturation on brain responses to metric deviants. *Frontiers in Neuroscience, 12*, 218. https://doi.org/10.3389/fnins.2018.00218

Henrich, K., Alter, K., Wiese, R., & Domahs, U. (2014). The relevance of rhythmical alternation in language processing: An ERP study on English compounds. *Brain and Language, 136*, 19–30. https://doi.org/10.1016/j.bandl.2014.07.003

Henrich, K., & Scharinger, M. (2022). Predictive processing in poetic language: Event-related potentials data on rhythmic omissions in metered speech. *Frontiers in Psychology, 12*, 782765. https://doi.org/10.3389/fpsyg.2021.782765

Henry, M. J., Cook, P., de Reus, K., Nityananda, V., Rouse, A. A., & Kotz, S. A. (2021). An ecological approach to measuring synchronization abilities across the animal kingdom. Philosophical Transactions of the Royal Society B. 376(1835):20200336.

Henry, M. J., & Herrmann, B. (2014). Low-frequency neural oscillations support dynamic attending in temporal context. *Timing & Time Perception, 2*(1), 62–86. https://doi.org/10.1163/22134468-00002011

Hickok, G., Farahbod, H., & Saberi, K. (2015). The rhythm of perception: Entrainment to acoustic rhythms induces subsequent perceptual oscillation. *Psychological Science, 26*(7), 1006–1013. https://doi.org/10.1177/0956797615576533

Hirata, S., Matsuda, G., Ueno, A., Fukushima, H., Fuwa, K., Sugama, K., Kusunoki, K., Tomonaga, M., Hiraki, K., & Hasegawa, T. (2013). Brain response to affective pictures in the chimpanzee. *Scientific Reports, 3*(1), 1342. Article 1. https://doi.org/10.1038/srep01342

Honing, H. (2012). Without it no music: Beat induction as a fundamental musical trait. *Annals of the New York Academy of Sciences, 1252*(1), 85–91. https://doi.org/10.1111/j.1749-6632.2011.06402.x

Honing, H. (2018). In H. Honing (Ed.), *The origins of musicality*. MIT Press.

Honing, H., & Bouwer, F. L. (2019). Rhythm. In P. J. Rentfrow & D. Levitin (Eds.), *Foundations in music psychology: Theory and research* (pp. 33–69). MIT Press.

Honing, H., Bouwer, F. L., & Háden, G. P. (2014). Perceiving temporal regularity in music: The role of auditory event-related potentials (ERPs) in probing beat perception. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of interval timing* (Vol. 829, pp. 305–323). Springer. https://doi.org/10.1007/978-1-4939-1782-2

Honing, H., Bouwer, F. L., Prado, L., & Merchant, H. (2018). Rhesus monkeys (Macaca mulatta) sense isochrony in rhythm, but not the beat: Additional support for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience, 12*, 475. https://doi.org/10.3389/fnins.2018.00475

Honing, H., & Merchant, H. (2014). Differences in auditory timing between human and non-human primates. *Behavioral and Brain Sciences, 27*(6), 557–558.

Honing, H., Merchant, H., Háden, G. P., Prado, L., & Bartolo, R. (2012). Rhesus monkeys (Macaca mulatta) detect rhythmic groups in music, but not the beat. *PLoS One, 7*(12), e51369. https://doi.org/10.1371/journal.pone.0051369

Huotilainen, M., Kujala, A., Hotakainen, M., Parkkonen, L., Taulu, S., Simola, J., Nenonen, J., Karjalainen, M., & Näätänen, R. N. (2005). Short-term memory functions of the human fetus recorded with magnetoencephalography. *Neuroreport, 16*(1), 81.

Itoh, K., Nejime, M., Konoike, N., Nakada, T., & Nakamura, K. (2015). Noninvasive scalp recording of cortical auditory evoked potentials in the alert macaque monkey. *Hearing Research, 327*, 117–125. https://doi.org/10.1016/j.heares.2015.05.007

Iversen, J. R., Repp, B. H., & Patel, A. D. (2009). Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences, 1169*, 58–73. https://doi.org/10.1111/j.1749-6632.2009.04579.x

Jacoby, N., & McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology, 27*(3), 359–370. https://doi.org/10.1016/j.cub.2016.12.031

Javitt, D. C., Jayachandra, M., Lindsley, R. W., Specht, C. M., & Schroeder, C. E. (2000). Schizophrenia-like deficits in auditory P1 and N1 refractoriness induced by the psychomimetic agent phencyclidine (PCP). *Clinical Neurophysiology, 111*(5), 833–836. https://doi.org/10.1016/S1388-2457(99)00313-2

Javitt, D. C., Schroeder, C. E., Steinschneider, M., Arezzo, J. C., & Vaughan, H. G. (1992). Demonstration of mismatch negativity in the monkey. *Electroencephalography and Clinical Neurophysiology, 83*(1), 87–90.

Javitt, D. C., Steinschneider, M., Schroeder, C. E., Vaughan, H. G., & Arezzo, J. C. (1994). Detection of stimulus deviance within primate primary auditory cortex: Intracortical mechanisms of mismatch negativity (MMN) generation. *Brain Research, 667*(2), 192–200.

Jones, M. R. (2009). Musical time. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 81–92). Oxford University Press.

Jones, M. R., Moynihan, H., MacKenzie, N., & Puente, J. (2002). Temporal aspects of stimulus-driven attending in dynamic arrays. *Psychological Science, 13*(4), 313–319. https://doi.org/10.1111/1467-9280.00458

Kaplan, T., Cannon, J., Jamone, L., & Pearce, M. (2022). Modeling enculturated bias in entrainment to rhythmic patterns. *PLoS Computational Biology, 18*(9), e1010579. https://doi.org/10.1371/journal.pcbi.1010579

Keele, S. W., Nicoletti, R., Ivry, R. B., & Pokorny, R. A. (1989). Mechanisms of perceptual timing: Beat-based or interval-based judgements? *Psychological Research, 50*, 251–256. https://doi.org/10.1007/BF00309261

Koelsch, S., Vuust, P., & Friston, K. J. (2019). Predictive processes and the peculiar case of music. *Trends in Cognitive Sciences, 23*(1), 63–77. https://doi.org/10.1016/J.TICS.2018.10.006

Kösem, A., Bosker, H. R., Takashima, A., Meyer, A., Jensen, O., & Hagoort, P. (2018). Neural entrainment determines the words we hear. *Current Biology, 28*(18), 2867–2875.e3. https://doi.org/10.1016/j.cub.2018.07.023

Kotz, S. A., Stockert, A., & Schwartze, M. (2014). Cerebellum, temporal predictability and the updating of a mental model. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 369*(1658), 20130403. https://doi.org/10.1098/rstb.2013.0403

Kujala, T., Tervaniemi, M., & Schröger, E. (2007). The mismatch negativity in cognitive and clinical neuroscience: Theoretical and methodological considerations. *Biological Psychology, 74*(1), 1–19. https://doi.org/10.1016/j.biopsycho.2006.06.001

Kumaravel, V. P., Farella, E., Parise, E., & Buiatti, M. (2022). NEAR: An artifact removal pipeline for human newborn EEG data. *Developmental Cognitive Neuroscience, 54*, 101068. https://doi.org/10.1016/j.dcn.2022.101068

Ladinig, O., Honing, H., Háden, G. P., & Winkler, I. (2009). Probing attentive and preattentive emergent meter in adult listeners without extensive music training. *Music Perception, 26*(4), 377–386. https://doi.org/10.1525/mp.2009.26.4.377

Ladinig, O., Honing, H., Háden, G. P., & Winkler, I. (2011). Erratum to Probing attentive and pre-attentive emergent meter in adult listeners with no extensive music training. *Music Perception, 26*(4), 444. https://doi.org/10.1525/mp.2011.28.4.444

Lange, K. (2009). Brain correlates of early auditory processing are attenuated by expectations for time and pitch. *Brain and Cognition, 69*(1), 127–137. https://doi.org/10.1016/j.bandc.2008.06.004

Lange, K. (2010). Can a regular context induce temporal orienting to a target sound? *International Journal of Psychophysiology, 78*(3), 231–238. https://doi.org/10.1016/j.ijpsycho.2010.08.003

Lange, K. (2013). The ups and downs of temporal orienting: A review of auditory temporal orienting studies and a model associating the heterogeneous findings on the auditory N1 with opposite effects of attention and prediction. *Frontiers in Human Neuroscience, 7*(June), 263. https://doi.org/10.3389/fnhum.2013.00263

Large, E. W. (2008). Resonating to musical rhythm: Theory and experiment. In S. Grondin (Ed.), *Psychology of time* (pp. 189–231). Emerald Group Publishing. https://doi.org/10.1016/B978-0-08046-977-5.00006-5

Large, E. W., & Gray, P. M. (2015). Spontaneous tempo and rhythmic entrainment in a bonobo (Pan paniscus). *Journal of Comparative Psychology, 129*(4), 317–328. https://doi.org/10.1037/com0000011

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*(1), 119–159. https://doi.org/10.1037/0033-295X.106.1.119

Laughlin, N. K., Hartup, B. K., Lasky, R. E., Meier, M. M., & Hecox, K. E. (1999). The development of auditory event related potentials in the rhesus monkey (Macaca mulatta). *Developmental Psychobiology, 34*(1), 37–56.

Lecanuet, J. P. (1996). Prenatal auditory experience. In *Musical beginnings: Origins and development of musical competence* (pp. 3–36). Oxford University Press.

Lei, J., Conradi, N., Abel, C., Frisch, S., Brodski-Guerniero, A., Hildner, M., Kell, C. A., Kaiser, J., & Schmidt-Kassow, M. (2019). Cognitive effects of rhythmic auditory stimulation in Parkinson's disease: A P300 study. *Brain Research, 1716*, 70–79. https://doi.org/10.1016/j.brainres.2018.05.016

Lenc, T., Merchant, H., Keller, P. E., Honing, H., Varlet, M., & Nozaradan, S. (2021). Mapping between sound, brain and behaviour: Four-level framework for understanding rhythm processing in humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences, 376*, 20200325. https://doi.org/10.1098/rstb.2020.0325

Leow, L., & Grahn, J. A. (2014). Neural mechanisms of rhythm perception: Present findings and future directions. In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of interval timing* (pp. 325–338). Springer. https://doi.org/10.1007/978-1-4939-1782-2

Lerdahl, F., & Jackendoff, R. (1983). An overview of hierarchical structure in music. *Music Perception, 1*(2), 229–252.

Lin, W. M., Oetringer, D. A., Bakker-Marshall, I., Emmerzaal, J., Wilsch, A., ElShafei, H. A., Rassi, E., & Haegens, S. (2021). No behavioural evidence for rhythmic facilitation of perceptual discrimination. *European Journal of Neuroscience, 55*(11-12), 3352–3364. https://doi.org/10.1111/ejn.15208

London, J. (2002). Cognitive constraints on metric systems: Some observations and hypotheses. *Music Perception, 19*(4), 529–550. https://doi.org/10.1525/mp.2002.19.4.529

London, J. (2012). *Hearing in time: Psychological aspects of musical meter*. Oxford University Press.

Longuet-Higgins, H. C., & Lee, C. S. (1984). The rhythmic interpretation of monophonic music. *Music Perception, 1*(4), 424–441.

Luck, S. (2005). *An introduction to the event-related potential technique*. The MIT Press.

Magne, C., Jordan, D. K., & Gordon, R. L. (2016). Speech rhythm sensitivity and musical aptitude: ERPs and individual differences. *Brain and Language, 153–154*, 13–19. https://doi.org/10.1016/j.bandl.2016.01.001

Mahmoudzadeh, M., Wallois, F., Kongolo, G., Goudjil, S., & Dehaene-Lambertz, G. (2017). Functional maps at the onset of auditory inputs in very early preterm human neonates. *Cerebral Cortex, 27*(4), 2500–2512. https://doi.org/10.1093/cercor/bhw103

Makov, S., & Zion Golumbic, E. (2020). Irrelevant predictions: Distractor rhythmicity modulates neural encoding in auditory cortex. *Cerebral Cortex, 30*(11), 5792–5805. https://doi.org/10.1093/cercor/bhaa153

Mathias, B., Zamm, A., Gianferrara, P. G., Ross, B., & Palmer, C. (2020). Rhythm complexity modulates behavioral and neural dynamics during auditory–motor synchronization. *Journal of Cognitive Neuroscience, 32*(10), 1864–1880. https://doi.org/10.1162/jocn_a_01601

Merchant, H., Grahn, J. A., Trainor, L. J., Rohrmeier, M. A., & Fitch, W. T. (2015). Finding the beat: A neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B, 370*, 20140093. https://doi.org/10.1098/rstb.2014.0093

Merchant, H., & Honing, H. (2014). Are non-human primates capable of rhythmic entrainment? Evidence for the gradual audiomotor evolution hypothesis. *Frontiers in Neuroscience, 7*(January), 274. https://doi.org/10.3389/fnins.2013.00274

Moore, J. K., & Linthicum, F. H. (2007). The human auditory system: A timeline of development. *International Journal of Audiology, 46*(9), 460–478. https://doi.org/10.1080/14992020701383019

Morillon, B., Schroeder, C. E., Wyart, V., & Arnal, L. H. (2016). Temporal prediction in lieu of periodic stimulation. *The Journal of Neuroscience, 36*(8), 2342–2347. https://doi.org/10.1523/JNEUROSCI.0836-15.2016

Moser, J., Schleger, F., Weiss, M., Sippel, K., Dehaene-Lambertz, G., & Preissl, H. (2020). Magnetoencephalographic signatures of hierarchical rule learning in newborns. *Developmental Cognitive Neuroscience, 46*, 100871. https://doi.org/10.1016/j.dcn.2020.100871

Näätänen, R. N., Kujala, T., & Winkler, I. (2010). Auditory processing that leads to conscious perception: A unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology, 48*, 4–22. https://doi.org/10.1111/j.1469-8986.2010.01114.x

Näätänen, R. N., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*(12), 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026

Näätänen, R. N., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*(4), 375–425. https://doi.org/10.1111/j.1469-8986.1987.tb00311.x

Nakamura, T., Michie, P. T., Fulham, W. R., Todd, J., Budd, T. W., Schall, U., Hunter, M., & Hodgson, D. M. (2011). Epidural auditory event-related potentials in the rat to frequency and duration deviants: Evidence of mismatch negativity? *Frontiers in Psychology, 2*, 1–17. https://doi.org/10.3389/fpsyg.2011.00367

Nave-Blodgett, J. E., Hannon, E., & Snyder, J. S. (2021). Hierarchical beat perception develops throughout childhood and adolescence and is enhanced in those with musical training. *Journal of Experimental Psychology: General, 150*(2), 314–339. https://doi.org/10.1037/xge0000903

Nelken, I., & Ulanovsky, N. (2007). Mismatch negativity and stimulus-specific adaptation in animal models. *Journal of Psychophysiology, 21*(3), 214–223. https://doi.org/10.1027/0269-8803.21.34.214

Ng, K. K., & Penney, T. B. (2014). Probing interval timing with scalp-recorded electroencephalography (EEG). In H. Merchant & V. de Lafuente (Eds.), *Neurobiology of Interval Timing* (pp. 187–207). Springer. https://doi.org/10.1007/978-1-4939-1782-2_11

Nobre, A. C., & van Ede, F. (2018). Anticipated moments: Temporal structure in attention. *Nature Reviews Neuroscience, 19*, 34–48. https://doi.org/10.1038/nrn.2017.141

Novembre, G., & Iannetti, G. D. (2018). Tagging the musical beat: Neural entrainment or event-related potentials? *Proceedings of the National Academy of Sciences, 115*(47), E11002–E11003. https://doi.org/10.1073/pnas.1815311115

Novitski, N., Tervaniemi, M., Huotilainen, M., & Näätänen, R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research, 20*(1), 26–36. https://doi.org/10.1016/j.cogbrainres.2003.12.011

Nozaradan, S. (2014). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 369*(1658), 20130393. https://doi.org/10.1098/rstb.2013.0393

Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *The Journal of Neuroscience, 31*(28), 10234–10240. https://doi.org/10.1523/JNEUROSCI.0411-11.2011

Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences, 23*(11), 913–926. https://doi.org/10.1016/J.TICS.2019.08.004

O'Reilly, C., Larson, E., Richards, J. E., & Elsabbagh, M. (2021). Structural templates for imaging EEG cortical sources in infants. *NeuroImage, 227*, 117682. https://doi.org/10.1016/j.neuroimage.2020.117682

Paavilainen, P., Arajärvi, P., & Takegata, R. (2007). Preattentive detection of nonsalient contingencies between auditory features. *Neuroreport, 18*(2), 159–163. https://doi.org/10.1097/WNR.0b013e328010e2ac

Palmer, C., & Demos, A. P. (2022). Are we in time? How predictive coding and dynamical systems explain musical synchrony. *Current Directions in Psychological Science, 31*(2), 147–153. https://doi.org/10.1177/09637214211053635

Patel, A. D. (2006). Musical rhythm, linguistic rhythm, and human evolution. *Music Perception, 24*(1), 99–104. https://doi.org/10.1525/mp.2006.24.1.99

Patel, A. D. (2021). Vocal learning as a preadaptation for the evolution of human beat perception and synchronization. *Philosophical Transactions of the Royal Society B: Biological Sciences, 376*(1835), 20200326. https://doi.org/10.1098/rstb.2020.0326

Patel, A. D., Iversen, J. R., Bregman, M. R., & Schulz, I. (2009). Experimental evidence for synchronization to a musical beat in a nonhuman animal. *Current Biology, 19*(10), 827–830. https://doi.org/10.1016/j.cub.2009.03.038

Patel, S. H., & Azzam, P. N. (2005). Characterization of N200 and P300: Selected studies of the Event-Related Potential. *International Journal of Medical Sciences, 2*(4), 147–154. https://doi.org/10.7150/ijms.2.147

Penhune, V. B., & Zatorre, R. J. (2019). Rhythm and time in the premotor cortex. *PLoS Biology, 17*(6), e3000293. https://doi.org/10.1371/journal.pbio.3000293

Pesnot Lerousseau, J., Trébuchon, A., Morillon, B., & Schön, D. (2021). Frequency selectivity of persistent cortical oscillatory responses to auditory rhythmic stimulation. *The Journal of Neuroscience, 41*(38), 7991–8006. https://doi.org/10.1523/JNEUROSCI.0213-21.2021

Phillips, H. N., Blenkmann, A., Hughes, L. E., Bekinschtein, T. A., & Rowe, J. B. (2015). Hierarchical organization of frontotemporal networks for the prediction of stimuli across multiple dimensions. *Journal of Neuroscience, 35*(25), 9255–9264. https://doi.org/10.1523/JNEUROSCI.5095-14.2015

Picton, T. W., & Hillyard, S. A. (1974). Human auditory evoked potentials. II: Effects of attention. *Electroencephalography and Clinical Neurophysiology, 36*, 191–199.

Picton, T. W., Hillyard, S. A., Krausz, H. I., & Galambos, R. (1974). Human auditory evoked potentials. I: Evaluation of components. *Electroencephalography* *and Clinical Neurophysiology, 36*, 179–190. https://doi.org/10.1016/0013-4694(74)90155-2

Pincze, Z., Lakatos, P., Rajkai, C., Ulbert, I., & Karmos, G. (2001). Separation of mismatch negativity and the N1 wave in the auditory cortex of the cat: A topographic study. *Clinical Neurophysiology, 112*(5), 778–784. https://doi.org/10.1016/S1388-2457(01)00509-0

Pincze, Z., Lakatos, P., Rajkai, C., Ulbert, I., & Karmos, G. (2002). Effect of deviant probability and interstimulus/interdeviant interval on the auditory N1 and mismatch negativity in the cat auditory cortex. *Cognitive Brain Research, 13*(2), 249–253. https://doi.org/10.1016/S0926-6410(01)00105-7

Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Potter, D. D., Fenwick, M., Abecasis, D., & Brochard, R. (2009). Perceiving rhythm where none exists: Event-related potential (ERP) correlates of subjective accenting. *Cortex, 45*(1), 103–109. https://doi.org/10.1016/j.cortex.2008.01.004

Poudrier, È. (2020). The influence of rate and accentuation on subjective rhythmization. *Music Perception, 38*, 27–45. https://doi.org/10.1525/MP.2020.38.1.27

Povel, D.-J., & Essens, P. (1985). Perception of temporal patterns. *Music Perception, 2*(4), 411–440. https://doi.org/10.2307/40285311

Proksch, S., Comstock, D. C., Médé, B., Pabst, A., & Balasubramaniam, R. (2020). Motor and predictive processes in auditory beat and rhythm perception. *Frontiers in Human Neuroscience, 14*(September), 578546. https://doi.org/10.3389/fnhum.2020.578546

Quené, H., & Port, R. F. (2005). Effects of timing regularity and metrical expectancy on spoken-word perception. *Phonetica, 62*(1), 1–13. https://doi.org/10.1159/000087222

Raimondi, T., Di Panfilo, G., Pasquali, M., Zarantonello, M., Favaro, L., Savini, T., Gamba, M., & Ravignani, A. (2023). Isochrony and rhythmic interaction in ape duetting. *Proceedings of the Royal Society B: Biological Sciences, 290*(1990), 20222244. https://doi.org/10.1098/rspb.2022.2244

Rajendran, V. G., Harper, N. S., Garcia-Lazaro, J. A., Lesica, N. A., & Schnupp, J. W. H. (2017). Midbrain adaptation may set the stage for the perception of musical beat. *Proceedings of the Royal Society B: Biological Sciences, 284*(1866), 20171455. https://doi.org/10.1098/rspb.2017.1455

Rajendran, V. G., Harper, N. S., & Schnupp, J. W. H. (2020). Auditory cortical representation of music favours the perceived beat. *Royal Society Open Science, 7*(3), 191194. https://doi.org/10.1098/rsos.191194

Ravignani, A. (2018). Rhythm and synchrony in animal movement and communication. *Current Zoology, 65*(1), 77–81. https://doi.org/10.1093/cz/zoy087

Reinhart, R. M. G., Heitz, R. P., Purcell, B. A., Weigand, P. K., Schall, J. D., & Woodman, G. F. (2012). Homologous mechanisms of visuospatial working memory maintenance in macaque and human: Properties and sources. *Journal of Neuroscience, 32*(22), 7711–7722. https://doi.org/10.1523/JNEUROSCI.0215-12.2012

Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: A review of recent research (2006-2012). *Psychonomic Bulletin and Review, 20*(3), 403–452. https://doi.org/10.3758/s13423-012-0371-2

Rimmele, J. M., Jolsvai, H., & Sussman, E. S. (2011). Auditory target detection is affected by implicit temporal and spatial expectations. *Journal of Cognitive Neuroscience, 23*(5), 1136–1147. https://doi.org/10.1162/jocn.2010.21437

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive sensing of periodic and aperiodic auditory patterns. *Trends in Cognitive Sciences, 22*(10), 870–882. https://doi.org/10.1016/J.TICS.2018.08.003

Rinne, T., Särkkä, A., Degerman, A., Schröger, E., & Alho, K. (2006). Two separate mechanisms underlie auditory change detection and involuntary control of attention. *Brain Research, 1077*(1), 135–143. https://doi.org/10.1016/j.brainres.2006.01.043

Rouse, A. A., Cook, P. F., Large, E. W., & Reichmuth, C. (2016). Beat keeping in a Sea Lion as coupled oscillation: Implications for comparative understanding of human rhythm. *Frontiers in Neuroscience, 10*, 257. https://doi.org/10.3389/fnins.2016.00257

Ruusuvirta, T., Koivisto, K., Wikgren, J., & Astikainen, P. (2007). Processing of melodic contours in urethane-anaesthetized rats. *The European Journal of Neuroscience, 26*(3), 701–703. https://doi.org/10.1111/j.1460-9568.2007.05687.x

Saberi, K., & Hickok, G. (2022a). A critical analysis of Lin et al.'s (2021) failure to observe forward entrainment in pitch discrimination. *European Journal of Neuroscience, 56*(8), 5191–5200. https://doi.org/10.1111/ejn.15778

Saberi, K., & Hickok, G. (2022b). Forward entrainment: Psychophysics, neural correlates, and function. *Psychonomic Bulletin & Review, 30*(3), 803–821. https://doi.org/10.3758/s13423-022-02220-y

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70*(1), 27–52. https://doi.org/10.1016/S0010-0277(98)00075-4

Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology, 19*(10), 831–836. https://doi.org/10.1016/j.cub.2009.03.061

Schaefer, R. S., Vlek, R. J., & Desain, P. (2010). Decomposing rhythm processing: Electroencephalography of perceived and self-imposed rhythmic patterns. *Psychological Research, 75*, 95–106. https://doi.org/10.1007/s00426-010-0293-4

Schall, U., Müller, B. W., Kärgel, C., & Güntürkün, O. (2015). Electrophysiological mismatch response recorded in awake pigeons from the avian functional equivalent of the primary auditory cortex. *Neuroreport, 26*(5), 239. https://doi.org/10.1097/WNR.0000000000000323

Schirmer, A., Wijaya, M., Chiu, M. H., Maess, B., & Gunter, T. C. (2021). Musical rhythm effects on visual attention are non-rhythmical: Evidence against metrical entrainment. *Social Cognitive and Affective Neuroscience, 16*(1–2), 58–71. https://doi.org/10.1093/scan/nsaa077

Schmidt-Kassow, M., Schubotz, R. I., & Kotz, S. A. (2009). Attention and entrainment: P3b varies as a function of temporal predictability. *Neuroreport, 20*(1), 31–36. https://doi.org/10.1097/WNR.0b013e32831b4287

Schröger, E., & Winkler, I. (1995). Presentation rate and magnitude of stimulus deviance effects on human pre-attentive change detection. *Neuroscience Letters, 193*, 185–188. https://doi.org/10.1016/0304-3940(95)11696-T

Schröger, E., & Wolff, C. (1998). Attentional orienting and reorienting is indicated by human event-related brain potentials. *Neuroreport, 9*(15), 3355–3358.

Schwartze, M., Farrugia, N., & Kotz, S. A. (2013). Dissociation of formal and temporal predictability in early auditory evoked potentials. *Neuropsychologia, 51*(2), 320–325. https://doi.org/10.1016/j.neuropsychologia.2012.09.037

Schwartze, M., & Kotz, S. A. (2015). The timing of regular sequences: Production, perception, and covariation. *Journal of Cognitive Neuroscience, 27*(9), 1697–1707. https://doi.org/10.1162/jocn_a_00805

Schwartze, M., Rothermich, K., Schmidt-Kassow, M., & Kotz, S. A. (2011). Temporal regularity effects on pre-attentive and attentive processing of deviance. *Biological Psychology, 87*(1), 146–151. https://doi.org/10.1016/j.biopsycho.2011.02.021

Schwartze, M., Stockert, A., & Kotz, S. A. (2015). Striatal contributions to sensory timing: Voxel-based lesion mapping of electrophysiological markers. *Cortex, 71*, 332–340. https://doi.org/10.1016/j.cortex.2015.07.016

Shiramatsu, T. I., & Takahashi, H. (2021). Mismatch-negativity (MMN) in animal models: Homology of human MMN? *Hearing Research, 399*, 107936. https://doi.org/10.1016/j.heares.2020.107936

Silva, S., & Castro, S. L. (2019). Structural meter perception is pre-attentive. *Neuropsychologia, 133*(June), 107184. https://doi.org/10.1016/j.neuropsychologia.2019.107184

Snyder, J. S., Pasinski, A. C., & McAuley, J. D. (2010). Listening strategy for auditory rhythms modulates neural correlates of expectancy and cognitive processing. *Psychophysiology, 48*(2), 198–207. https://doi.org/10.1111/j.1469-8986.2010.01053.x

Stefanics, G., Háden, G. P., Huotilainen, M., Balázs, L., Sziller, I., Beke, A., Fellman, V., & Winkler, I. (2007). Auditory temporal grouping in newborn infants. *Psychophysiology, 44*(5), 697–702. https://doi.org/10.1111/j.1469-8986.2007.00540.x

Sun, Y., Michalareas, G., & Poeppel, D. (2021). The impact of phase entrainment on auditory detection is highly variable: Revisiting a key finding. *European Journal of Neuroscience, 55*(11–12), 1–18. https://doi.org/10.1111/ejn.15367

Suppanen, E., Huotilainen, M., & Ylinen, S. (2019). Rhythmic structure facilitates learning from auditory input in newborn infants. *Infant Behavior and Development, 57*, 101346. https://doi.org/10.1016/J.INFBEH.2019.101346

Sussman, E. S. (2007). A new view on the MMN and attention debate. *Journal of Psychophysiology, 21*(3), 164–175. https://doi.org/10.1027/0269-8803.21.34.164

Tal, I., Large, E. W., Rabinovitch, E., Wei, Y., Schroeder, C. E., Poeppel, D., & Zion Golumbic, E. (2017). Neural entrainment to the beat: The "Missing-Pulse" phenomenon. *The Journal of Neuroscience, 37*(26), 6331–6341. https://doi.org/10.1523/JNEUROSCI.2500-16.2017

Teichert, T. (2016). Tonal frequency affects amplitude but not topography of rhesus monkey cranial EEG components. *Hearing Research, 336*, 29–43. https://doi.org/10.1016/j.heares.2016.04.001

Teki, S., Grube, M., Kumar, S., & Griffiths, T. D. (2011). Distinct neural substrates of duration-based and beat-based auditory timing. *The Journal of Neuroscience, 31*(10), 3805–3812. https://doi.org/10.1523/JNEUROSCI.5561-10.2011

ten Cate, C., & Honing, H. (2023). Precursors of Music and Language in Animals. In D. Sammler (Ed.), *Oxford handbook of language and music*. Oxford University Press.

Thiede, A., Virtala, P., Ala-Kurikka, I., Partanen, E., Huotilainen, M., Mikkola, K., Leppänen, P. H. T., & Kujala, T. (2019). An extensive pattern of atypical neural speech-sound discrimination in newborns at risk of dyslexia. *Clinical Neurophysiology, 130*(5), 634–646. https://doi.org/10.1016/j.clinph.2019.01.019

Tichko, P., & Large, E. W. (2019). Modeling infants' perceptual narrowing to musical rhythms: Neural oscillation and Hebbian plasticity. *Annals of the New York Academy of Sciences, 1453*(April), 125–139. https://doi.org/10.1111/nyas.14050

Tierney, A. T., & Kraus, N. (2013). Neural responses to sounds presented on and off the beat of ecologically valid music. *Frontiers in Systems Neuroscience, 7*(May), 14. https://doi.org/10.3389/fnsys.2013.00014

Todd, J., Háden, G. P., & Winkler, I. (2022). Relevance to the higher order structure may govern auditory statistical learning in neonates. *Scientific Reports, 12*(1), 5905. Article 1. https://doi.org/10.1038/s41598-022-09994-0

Trainor, L. J., McFadden, M., Hodgson, L., Darragh, L., Barlow, J., Matsos, L., & Sonnadara, R. (2003). Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *International Journal of Psychophysiology, 51*, 5–15. https://doi.org/10.1016/S0167-8760(03)00148-X

Trainor, L. J., Samuel, S. S., Desjardins, R. N., & Sonnadara, R. (2001). Measuring temporal resolution in infants using mismatch negativity. *Neuroreport, 12*(11), 2443–2448. https://doi.org/10.1097/00001756-200108080-00031

Tsogli, V., Jentschke, S., & Koelsch, S. (2022). Unpredictability of the "when" influences prediction error processing of the "what" and "where.". *PLoS One, 17*(2 February), 1–19. https://doi.org/10.1371/journal.pone.0263373

Ueno, A., Hirata, S., Fuwa, K., Sugama, K., Kusunoki, K., Matsuda, G., Fukushima, H. H., Hiraki, K., Tomonaga, M., & Hasegawa, T. (2008). Auditory ERPs to stimulus deviance in an awake chimpanzee (Pan troglodytes): Towards hominid cognitive neurosciences. *PLoS One, 3*(1), e1442. https://doi.org/10.1371/journal.pone.0001442

Ueno, A., Hirata, S., Fuwa, K., Sugama, K., Kusunoki, K., Matsuda, G., Fukushima, H., Hiraki, K., Tomonaga, M., & Hasegawa, T. (2009). Brain activity in an awake chimpanzee in response to the sound of her own name. *Biology Letters, 6*(3), 311–313. https://doi.org/10.1098/rsbl.2009.0864

Umbricht, D., Vyssotki, D., Latanov, A., Nitsch, R., & Lipp, H.-P. (2005). Deviance-related electrophysiological activity in mice: Is there mismatch negativity in mice? *Clinical Neurophysiology, 116*(2), 353–363. https://doi.org/10.1016/j.clinph.2004.08.015

van Atteveldt, N., Musacchia, G., Zion-Golumbic, E., Sehatpour, P., Javitt, D. C., & Schroeder, C. E. (2015). Complementary fMRI and EEG evidence for more efficient neural processing of rhythmic vs. Unpredictably timed sounds. *Frontiers in Psychology, 6*(Oct), 1–11. https://doi.org/10.3389/fpsyg.2015.01663

van Bree, S., Sohoglu, E., Davis, M. H., & Zoefel, B. (2021). Sustained neural rhythms reveal endogenous oscillations supporting speech perception. *PLoS Biology, 19*(2), e3001142. https://doi.org/10.1371/journal.pbio.3001142

Virtala, P., Putkinen, V., Kailaheimo-Lönnqvist, L., Thiede, A., Partanen, E., & Kujala, T. (2022). Infancy and early childhood maturation of neural auditory change detection and its associations to familial dyslexia risk. *Clinical Neurophysiology, 137*, 159–176. https://doi.org/10.1016/j.clinph.2022.03.005

Vuust, P., Ostergaard, L., Pallesen, K. J., Bailey, C., & Roepstorff, A. (2009). Predictive coding of music—Brain responses to rhythmic incongruity. *Cortex; a Journal Devoted to the Study of the Nervous System and Behavior, 45*(1), 80–92. https://doi.org/10.1016/j.cortex.2008.05.014

Vuust, P., Pallesen, K. J., Bailey, C., Van Zuijen, T. L., Gjedde, A., Roepstorff, A., & Østergaard, L. (2005). To musicians, the message is in the meter: Pre-attentive neuronal responses to incongruent rhythm are left-lateralized in musicians. *NeuroImage, 24*(2), 560–564. https://doi.org/10.1016/j.neuroimage.2004.08.039

Vuust, P., & Witek, M. A. G. (2014). Rhythmic complexity and predictive coding: A novel approach to modeling rhythm and meter perception in music. *Frontiers in Psychology, 5*(October), 1111. https://doi.org/10.3389/fpsyg.2014.01111

Weineck, K., Wen, O. X., & Henry, M. J. (2022). Neural synchronization is strongest to the spectral flux of slow music and depends on familiarity and beat salience. *eLife, 11*, e75515. https://doi.org/10.7554/eLife.75515

Wilson, M., & 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. *Psychonomic Bulletin & Review, 23*(6), 1647–1659. https://doi.org/10.3758/s13423-016-1013-x

Winkler, I. (2007). Interpreting the mismatch negativity. *Journal of Psychophysiology, 21*(3), 147–163. https://doi.org/10.1027/0269-8803.21.34.147

Winkler, I., & Czigler, I. (2012). Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology, 83*(2), 132–143. https://doi.org/10.1016/j.ijpsycho.2011.10.001

Winkler, I., Denham, S., & Escera, C. (2013). Auditory event-related potentials. In D. Jaeger & R. Jung (Eds.), *Encyclopedia of computational neuroscience* (pp. 1–29). Springer. https://doi.org/10.1007/978-1-4614-7320-6_99-1

Winkler, I., Háden, G. P., Ladinig, O., Sziller, I., & Honing, H. (2009). Newborn infants detect the beat in music. *Proceedings of the National Academy of Sciences, 106*(7), 2468–2471. https://doi.org/10.1073/pnas.0809035106

Winkler, I., Kushnerenko, E., Horváth, J., Ceponiene, R., Fellman, V., Huotilainen, M., Näätänen, R. N., & Sussman, E. S. (2003). Newborn infants can organize the auditory world. *Proceedings of the National Academy of Sciences of the United States of America, 100*(20), 11812–11815. https://doi.org/10.1073/pnas.2031891100

Woodman, G. F. (2011). Homologues of human ERP components in nonhuman primates. In E. S. Kappenman & S. J. Luck (Eds.), *The Oxford handbook of event-related potential components*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195374148.013.0271

Yabe, H., Tervaniemi, M., Reinikainen, K., & Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *Neuroreport, 8*(8), 1971–1974.

Yates, C. M., Justus, T., Atalay, N. B., Mert, N., & Trehub, S. E. (2016). Effects of musical training and culture on meter perception. *Psychology of Music, 45*(2), 231–245. https://doi.org/10.1177/0305735616657407

Zarco, W., Merchant, H., Prado, L., & Mendez, J. C. (2009). Subsecond timing in primates: Comparison of interval production between human subjects and rhesus monkeys. *Journal of Neurophysiology, 102*(6), 3191–3202. https://doi.org/10.1152/jn.00066.2009

Zoefel, B., ten Oever, S., & Sack, A. T. (2018). The involvement of endogenous neural oscillations in the processing of rhythmic input: More than a regular repetition of evoked neural responses. *Frontiers in Neuroscience, 12*, 95. https://doi.org/10.3389/fnins.2018.00095

Rhythms in Speech

M. Florencia Assaneo and Joan Orpella

Abstract

Speech can be defined as the human ability to communicate through a sequence of vocal sounds. Consequently, speech requires an emitter (the speaker) capable of generating the acoustic signal and a receiver (the listener) able to successfully decode the sounds produced by the emitter (i.e., the acoustic signal). Time plays a central role at both ends of this interaction. On the one hand, speech production requires precise and rapid coordination, typically within the order of milliseconds, of the upper vocal tract articulators (i.e., tongue, jaw, lips, and velum), their composite movements, and the activation of the vocal folds. On the other hand, the generated acoustic signal unfolds in time, carrying information at different timescales. This information must be parsed and integrated by the receiver for the correct transmission of meaning. This chapter describes the temporal patterns that characterize the speech signal and reviews research that explores the neural mechanisms underlying the generation of these patterns and the role they play in speech comprehension.

Keywords

Syllabic rhythm · Speech production · Speech perception · Brain oscillations

Time has long been recognized as a critical dimension for speech perception and for auditory perception more generally. Time is, in fact, so inherent to sound that, without time, sounds would not exist. This property of sound sets audition aside from other sensory domains, such as vision or olfaction, in which the relevant objects (a picture, a scent) can, in principle, exist in their static form. In this chapter, we will dive deep into the role played by time and, particularly, temporal regularity in the speech process (see Fig. 1). First, we will describe the physical properties of the speech acoustic signal, with a particular emphasis on its rhythmic structure. Next, we will move to the listener's perspective to show that the perceptual system capitalizes on this rhythmic structure to decode the speech signal successfully. Then, we will explore how the speaker generates this rhythm and how it relates to the biophysical properties of the speech-motor system. To conclude the chapter, we will merge listener and speaker dimensions and present

---

M. F. Assaneo (✉)
Instituto de Neurobiología, Universidad Autónoma de México, Santiago de Querétaro, Mexico
e-mail: fassaneo@inb.unam.mx

J. Orpella
Department of Neuroscience, Georgetown University Medical Center, Washington, DC, USA
e-mail: joan.orpella@georgetown.edu

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_14

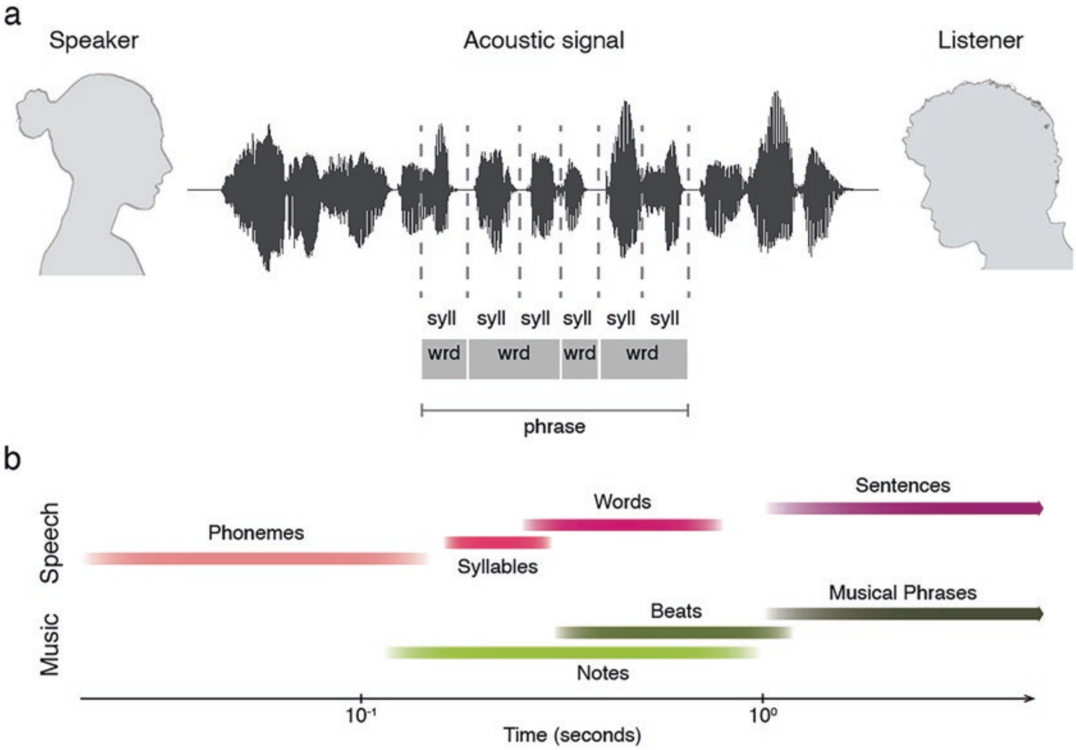


Fig. 1 Schematic representation of the speech process and its temporal scales. (**a**) The left silhouette represents a female Spanish speaker articulating the sentence "Cierra siempre la puerta de casa con llave" (Always lock the house door). The central panel shows the acoustic signal generated by the speaker. This signal carries information at different timescales. Syllables, for example, correspond to short segments lasting between 125 ms and 330 ms. One or more syllables form words and several words produce phrases. The silhouette on the right represents a conversational partner, who receives and decodes the generated acoustic signal extracting from it the information sent by the speaker. (**b**) There are important similarities between the speech process and music; both are inherently human and involve acoustic signals that unfold in time. The figure shows a comparison between the temporal scales of commonly defined units of speech and music

hypotheses about the benefits of having perception and production systems operate optimally through the same rhythmic patterns.

The Acoustic Signal

Acoustic signals can be decomposed into two main components: a carrier signal and a modulatory wave. The carrier defines the content at different frequencies which compose the speech signal. The modulatory wave represents the amplitude of the signal. From the perspective of perception, the carrier defines the tone of the acoustic signal and the modulatory wave its volume. When it comes to speech, the acoustic signal exhibits rich dynamics in both components (Elliott & Theunissen, 2009). On the one hand, it shows rapid changes in the frequency content of the carrier signal, known as the fine structure of the speech signal. On the other, it shows continuous variations in its amplitude, defined as the speech envelope. The speech envelope can be easily identified in plots of the acoustic signal in the time domain (e.g., Fig. 2a). To examine the fine structure (the carrier), it is necessary to represent the signal in frequency space (i.e., decompose the signal into frequency bands). Specifically, the speech fine structure becomes visible in the spectrogram of the acoustic signal (a representation of the signal that displays the time evolution of its frequency content; Fig. 2b). Phonemes,

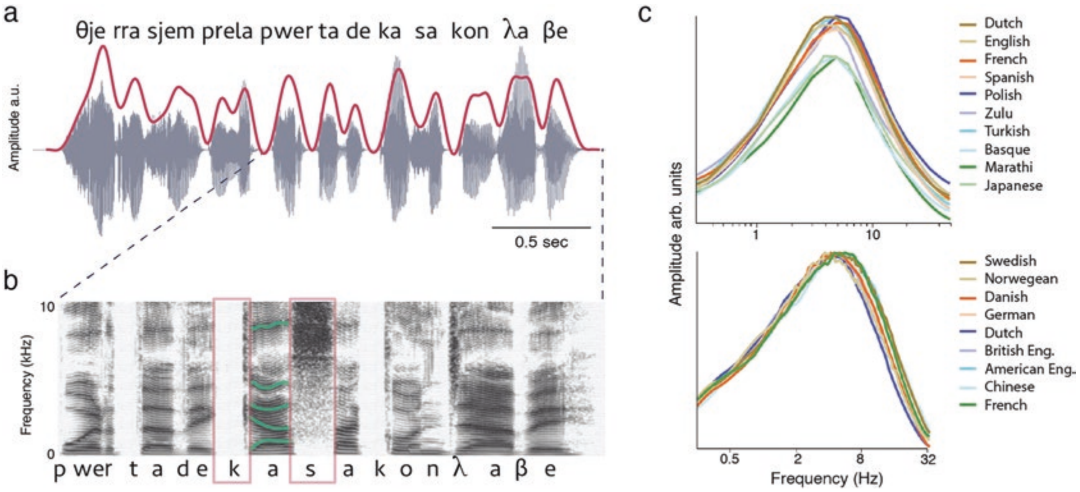


Fig. 2 The acoustic signal. (**a**) Time domain. Time evolution of the acoustic signal corresponding to the same sentence presented in Fig. 1, with its envelope superimposed as a magenta trace. On top, the corresponding phonetic transcription (i.e., symbols identifying the spoken sounds; for more detail, see (International Phonetic Association, 1999)) aligned with the acoustic signal. (**b**) Frequency domain. Spectrogram corresponding to the acoustic signal delimited by the dashed lines ("puerta de casa con llave"). Below the phonetic transcription corresponding to each segment of the spectrogram. As highlighted by the pink rectangles, consonants can be defined by a silence followed by a burst of noise (like "*k*") or by a high-frequency narrowband noise (like "*s*"). Vowels, instead, are characterized by a set of emphasized frequencies, known as formants, highlighted in green for one example case. The trajectories followed by the formants when transitioning from a consonant to the stable part of the vowel define the characteristic signature of the specific consonant-vowel combination. (**c**) Modulation spectrum. Each line shows the average of the envelopes' spectra computed from different audio samples of different speakers of a given language. (Upper panel adapted from Varnet et al. (2017). Lower panel adapted from Ding et al. (2017))

defined as the vocal sounds composing a language, can be identified by their characteristic signatures in the spectrogram (Liberman, 1996). For example, emphasized frequency bands, known as formants, determine the identity of vowels, while formant transitions following noise bursts identify consonants (see green lines in Fig. 2b) (Stevens, 1999). Given that phonemes can be determined by the signal's fine structure, the most common approach to studying the acoustic signal of speech has focused on its spectrogram, irrespective of the envelope. However, a set of studies that examined this dimension of the signal (the envelope) revealed its critical role for perception. Specifically, these studies showed that, if the envelope is preserved, the speech signal remains intelligible even when the information carried by the fine structure is largely degraded. That is, the fine structure may be replaced by different carrier signals altogether (what is known as noise-vocoded speech), but as long as the envelope is preserved, comprehension is still possible (Faulkner et al., 2000; Shannon et al., 1995). This line of research shows that a large amount of the information conveyed by the speech acoustic signal is embedded in its envelope.

The speech acoustic signal is rhythmically structured, as can be readily inferred by looking at its waveform and the relatively regular sequence of increments and decrements in its amplitude (see Fig. 2a). Historically, however, the rhythmic structure of speech was first examined through the lens of linguists, who attempted to measure this rhythmical structure in the succession of arbitrary linguistic units (mostly syllables) rather than by examining the acoustic signal directly (i.e., without committing to any predetermined units). This line of investigation proposed that languages come in two rhythmic types: syllable-timed and stress-timed (Abercrombie, 2022). In syllable-timed languages, the sequence of syllables composing a speech stream have equal durations. For stress-timed languages, it is the time interval between stressed syllables (accented/prominent syllables in the speech) that remains rhythmically stable. Studies that define languages using this nomenclature are still common. Spanish and English, for example, are typically categorized as syllable-timed and stress-timed, respectively. However, efforts to find a nonarbitrary measurement to automatically classify languages as belonging to one or the other group have been unsuccessful (Nolan & Jeon, 2014) and this categorization of languages is now considered a mere perceptual phenomenon.

A more principled approach to examining the rhythmic properties of speech is to study the acoustic signal directly, without commitment to preconceived linguistic units. In two comprehensive studies (Ding et al., 2017; Varnet et al., 2017), researchers explored the spectrum of the speech envelope (i.e., the frequencies that characterize the speech amplitude modulations) across different speakers, languages, and conditions (e.g., isolated sentences, audiobooks, conversational speech). To do this, the authors first extracted the speech envelope from each audio sample (see Fig. 2a). Next, they obtained the spectrum of each of the computed envelopes through an analytical procedure known as the fast Fourier transform. Finally, they averaged these modulation spectra across speakers and conditions within the language to arrive at the average envelope spectrum for that language. The results of both studies using this analysis approach reveal striking consistencies across languages, with average envelope spectra across languages in the range of 2–8 Hz, peaking around 5 Hz (see Fig. 2c).

Given that the mean syllable duration across languages is roughly 200 ms (about 5 syllables per second), these results align with previous linguistic considerations. Moreover, it has been proposed that the breadth of the reported peak reflects the variability in syllable duration (Greenberg et al., 2003), which may explain why linguists failed to measure the syllables' isochrony. In sum, although the sequence of linguistic units (i.e., syllables) that compose speech are not perfectly isochronous, there is still a universal quasi-rhythmic speech structure where the syllabic rate across languages is restricted to a range from 2 to 8 syllables per second.

The Listener

As mentioned, the speech signal carries critical information over short timescales, in the range of tens of milliseconds, corresponding to its fine structure. Important information is also carried by the slow energy modulations characterized by the speech envelope (see Fig. 2), which roughly correspond to the spoken syllabic rate. Stress and tonal contrasts, as well as linguistic (e.g., signaling a statement vs. a question) and affective (e.g., signaling sadness) prosody, may be carried at the syllable level (Rosen, 1992). But beyond this, it seems that the rhythmic nature of these slow modulations is critical to the listener for the decoding of the fine detail itself, that is, for speech intelligibility.

The auditory system appears to be well attuned to the slow modulations in the signal, particularly in frequencies of the order of 2–8 Hz. Early psychophysical studies highlighted this sensitivity, which is not specific to speech (e.g., Houtgast & Steeneken, 1985; Viemeister, 1979). In these studies, the threshold for the perception of amplitude-modulated signals at different frequencies is understood as a measure of the temporal resolution of the system. A consistent finding is that the perceptual threshold is lowest (i.e., perception is best) at modulation frequencies between 2 and 8 Hz, increasing (worst perception) as we move away from this range, suggesting that the 2–8 Hz frequency range is privileged in auditory perception in general. Moreover, these slow rhythms bear a close link to speech intelligibility. Indeed, intelligibility is highest when the rhythmic structure of the signal (the envelope's spectrum) falls within this specific range of frequencies (2–8 Hz), which is also well-aligned with the average modulation rates observed across languages (see Fig. 2c).

A compelling demonstration of the relevance of slow rhythms for speech intelligibility comes from studies using time-compressed speech (e.g., Ahissar et al., 2001; Ghitza & Greenberg, 2009; Versfeld & Dreschler, 2002). A common finding is that intelligibility is maintained up to speeds of about 8 syllables per second and declines rapidly after that. Preserving the natural syllabic rhythm may thus be critical for speech comprehension. In an elegant experiment, Ghitza and Greenberg (2009) used naturally spoken sentences with unpredictable semantic content (e.g., *The vast trade dealt the task*) to measure the intelligibility of time-compressed speech. In the lab, intelligibility is typically measured by presenting participants with a word or sentence (often manipulated in some way, e.g., time-compressed, masked by noise) and asking them to type it back. As expected, intelligibility was poor (with error rates of about 50%) when sentences were compressed by a factor of 3 (i.e., at unnaturally fast syllabic rates). What was more interesting was that, when silent gaps were inserted between the successive chunks of compressed speech (corresponding to acoustic segments of 40 ms), a manipulation that recovered the overall temporal structure of the original signal (Fig. 3a), intelligibility was considerably restored. Since the length of the acoustic segments was the same in all cases (i.e., there was no additional acoustic information in the silent gaps condition), this “perceptual restoration” was hypothesized to result from the reestablishing of the natural syllabic rate, affording the system the necessary time for decoding the acoustic information.

So far, we have seen that the speech signal exhibits a remarkable rhythmicity within a range of relatively low frequencies, that the auditory system seems to be most sensitive to modulations within this range, and that this range of frequencies also defines the limits for speech intelligibility. Can we also find these rhythms in the listener’s brain activity? And, given their relevance for intelligibility, how does the perceptual system use them?

Given the aforementioned findings, it is reasonable to predict that the auditory cortex will show a particular sensitivity to temporal modulation rates within the same restricted range of 2–8 Hz, that is, within the theta frequency band. Several neuroimaging studies have confirmed this prediction (e.g., Boemio et al., 2005; Fuglsang et al., 2022; Giraud et al., 2007). For example, when testing a wide range of frequencies (4–256 Hz), primary and secondary auditory cortical regions are selective to amplitude modulations (sinusoidally modulated white noise) at frequencies between 4 and 8 Hz (i.e., theta band) (Giraud et al., 2000).

It is interesting to note that spontaneous fluctuations in the theta band are also present in auditory cortical regions (e.g., Keitel & Gross, 2016). Giraud et al. (2007) observed spontaneous oscillatory activity in the theta range during resting state (no task) in the auditory cortex using concurrent electroencephalography and functional magnetic resonance imaging. Thus, the theta rhythm appears to be endogenously generated in these regions, in addition to other rhythms such as gamma (>30 Hz), with which it may interact (see below). This observation places spontaneous cortical activity in auditory regions in close alignment with the natural rhythms of produced speech.

Is the proximity in frequencies between the rhythms of speech and spontaneous auditory activity a coincidence? Possibly not. Data show that, beyond its spontaneous activity and sensitivity to amplitude modulations in the theta range, the auditory cortex closely tracks these amplitude modulations in the signal, that is, the speech envelope (see Fig. 3b). This tracking, sometimes called *entrainment*, occurs regardless of whether the acoustic signal contains intelligible speech (Ahissar et al., 2001; Luo & Poeppel, 2007), backward speech (Howard & Poeppel, 2010), a foreign language (Peña & Melloni, 2012), or white noise (Henry & Obleser, 2012). Because of this, auditory entrainment is often understood as an automatic coupling of the intrinsic oscillatory activity of the auditory system to the acoustic input (Doelling et al., 2019).

Given the robustness of this tracking or entrainment phenomenon, researchers have also asked whether the auditory entrainment to the speech envelope has consequences for

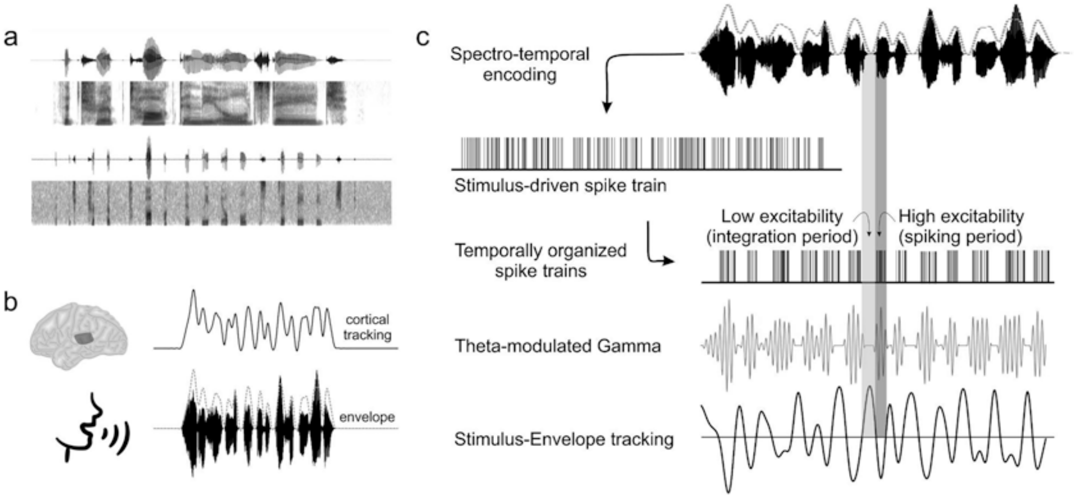


Fig. 3 A listener's perspective. (**a**) Example stimuli used in Ghitza and Greenberg (2009). Acoustic signal and spectrogram of a natural speech sentence (first two rows) and the 40 ms snippets of sound from this sentence with 80 ms interleaved silences used to recover the sentence's original rhythmic structure (acoustic signal and its corresponding spectrogram, in row 3 and 4, respectively) (adapted from Ghitza & Greenberg, 2009). (**b**) Auditory entrainment to speech. The upper panel displays a schematic representation of the auditory cortex activity elicited during the perception of the sentence displayed underneath. During speech perception, auditory brain regions entrain their activity to the speech envelope. (**c**) Schematic representation of the theory of oscillation-based operations in the perception of speech by Giraud and Poeppel (2012). A stimulus-driven spike train is assumed to represent the speech signal in auditory brain regions (top). These regions align their theta oscillatory activity to the speech envelope (bottom), which modulates gamma power. In turn, modulations in gamma power regulate the excitability of neurons in charge of forwarding the signal to higher-order cortical regions, in terms of alternating periods of integration and transmission (spiking). The result of these operations is not only the alignment of cortical activity to the more informative parts of the speech signal but also the packaging of speech information in chunks of a size appropriate for subsequent computations. (Adapted from Giraud and Poeppel (2012))

speech intelligibility. Neurophysiological data show that the degree of entrainment (specifically in the theta band) correlates with intelligibility (Doelling et al., 2014; Park et al., 2015; Peelle & Davis, 2012). For instance, using magnetoencephalography (MEG), Doelling and colleagues (2014) showed that attenuating the temporal fluctuations in the speech envelope linked to the syllabic rate reduced cortical tracking as well as intelligibility. Critically, tracking and intelligibility were both recovered when the temporal cues at the syllable rate were artificially reinstated by introducing "clicks" at the time points of the original temporal modulation peaks. Relatedly, auditory entrainment to the perceived speech envelope of poor readers (Abrams et al., 2009) and children with dyslexia (Goswami, 2011) shows abnormal patterns, which may explain the observed deficits.

An important question is whether cortical tracking in the critical frequency band characterized by the envelope is functionally relevant for intelligibility. In other words, the fact that envelope tracking correlates with intelligibility does not necessarily imply that tracking plays a causal role in intelligibility. Conclusions regarding causality can be drawn from studies showing that intelligibility is compromised when auditory entrainment to the speech envelope is disrupted via noninvasive electrical cortical stimulation (Wilsch et al., 2018; Zoefel et al., 2018). In a recent study, Riecke and colleagues (2018) showed that the degree of intelligibility of a speech signal, from which envelope fluctuations had been removed, scales with the degree to which intrinsic oscillations in temporal areas can be biased (again, via electrical stimulation) toward the times of those original envelope fluctuations. These studies suggest that entrainment correlates with intelligibility and is necessary for speech to be perceived as intelligible.

There are also reasons to believe that the theta rhythm is a hardwired property of the system. Adults naturally tend to modify their speech when it is directed to infants, enhancing its rhythmic structure by producing more even-timed syllables (Lee et al., 2014). This spontaneous adjustment of speech may actually facilitate language acquisition, since infant-directed speech increases infants' attention to the linguistic input and enhances their cortical tracking of the speech envelope (Kalashnikova et al., 2018).

Why the syllable rhythm and corresponding tracking by auditory areas is so important for speech perception is not fully understood. Several explanations have been advanced. Ghitza (2011) proposed a model (called *Tempo*) in which the auditory input is sequentially processed by a *parsing* path followed by a *decoding* path. While the details of this model are beyond the scope of this chapter, it suffices to say that tracking of the speech input in the theta band is understood as a *syllabic parsing* mechanism. The parsing of the speech signal is critical for its subsequent *decoding*, when segmented chunks of speech are mapped onto memory units, determining speech comprehension. In the case of syllables, for example, syllable-like chunks segmented from the continuous speech during the parsing step are mapped onto so-called *syllable neurons* at the end of each theta cycle.

A related possibility, suggested by Giraud and Poeppel (2012), is that envelope tracking may play two fundamentally different but equally crucial roles, depending on the hemisphere. While theta tracking by the right auditory cortex may be used to integrate the spectral properties of the input for speaker identification and prosodic analysis (Zatorre et al., 2002), left-lateralized theta tracking may be used for the purpose of sampling and discretizing the speech input (see Fig. 3c). In particular, speech envelope tracking in the theta band by the left auditory cortex may be used to modulate gamma activity to discretize the speech input arriving in early auditory cortical areas. The speech input can thus be transformed from continuous neuronal spike trains to packages of spike trains with alternating periods of information integration and subsequent transmission to the next hierarchical level. This mechanism would thus not only represent a useful means for sampling and parsing the speech input by focusing on its most informative parts (e.g., syllables onsets or maximum energy points), but also for providing outputs of the appropriate size for subsequent computational stages (e.g., phonological processing and contact with motor representations).

In sum, the auditory system appears to come equipped with spontaneous theta oscillations and shows a remarkable sensitivity to rhythmic inputs in the same frequency range. When faced with speech, these spontaneous oscillations align with the speech envelope, which shows a consistent rhythmicity in theta. This alignment has been shown to be functionally relevant for intelligibility, and several theories suggest that its main purpose is to optimize the parsing of the speech input and provide chunks of the appropriate size for subsequent decoding of the finer detail. Next, we turn to the speaker's perspective and the biophysical properties of the system that gives rise to the speech signal.

The Speaker

As shown in the previous sections, the speech acoustic signal exhibits temporal regularities in its envelope, which are in turn recovered by the perceptual system. Here, we will explore speech rhythmicity from the speaker's perspective by describing the anatomical system responsible for generating speech and the neural activity underpinning the dynamics of such a system.

The human vocal system comprises two main components: the vocal folds, which are the sound source, and the vocal tract, which acts as a filter of that sound (see Fig. 4a). The folds are no more than two membranes emplaced in the larynx with the ability to vibrate and produce sound when

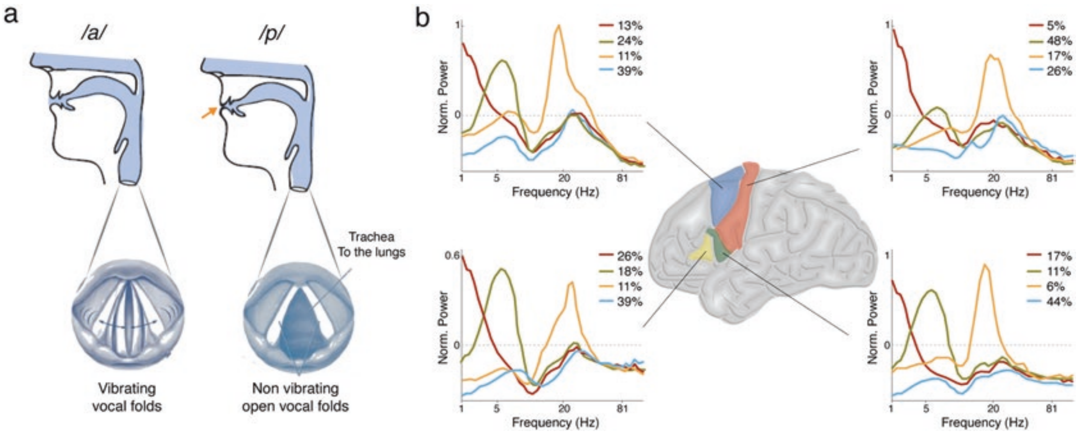


Fig. 4 The speaker. (**a**) Human vocal organ. Schematic representation of the motor gesture corresponding to the production of an */a/*, on the left, and of a */p/* on the right. For */a/*, a voiced vowel, the vocal tract (set of cavities highlighted in blue in the upper panel) is open and the vocal folds are activated (lower panel). For */p/*, an unvoiced consonant, the vocal tract is constricted at the lips (orange arrow) and the folds are open and quiet (lower panel). (**b**) Spectral profile of resting state activity in speech production related regions. Spectral power in comparison to the whole brain. Different colored lines show the set of spectra that best describe the resting state activity in each brain area. These spectra profiles do not necessarily coexist in time. The legends show the percentage of trials (1 second segments) in which each spectrum was presented. Left precentral gyrus (red) composes the motor cortex; supplementary motor area (blue) has been related to speech initiation (Guenther, 2015); left inferior frontal pars opercularis (green) and triangularis (yellow) are related to speech planning (Castellucci et al., 2022) and have been shown to regulate the timing of speech production (Long et al., 2016). (All panels were adapted from Keitel and Gross (2016))

forced by the air pressure coming from the lungs. The vocal tract is constituted by a set of cavities between the glottis and the lips. During the generation of vowels, the sound produced by the folds travels along the tract before reaching the outside of the body (Titze, 1994). The passage along the tract filters the signal, emphasizing some frequencies (i.e., the formants depicted as green lines in Fig. 2b) while suppressing others. This filtering process results in the signatures that characterize the different vowels (Stevens, 1999). Conversely, during the production of unvoiced consonants such as “s” or “k,” the folds remain quiet, and the sound originates as a turbulence in the airflow produced by the constriction of the tract (see Fig. 4a). The degree of constriction (e.g., a complete closure followed by a sudden release for stop consonants like “k” or “p” or a stable narrowing for fricatives likes “s” or “f”) and the point of the tract where the constriction takes place result in the acoustic signatures that characterize the different unvoiced sounds. Crucially, the vocal tract can take many different configurations by the movements of the speech articulators (i.e., tongue, lips, jaw, and velum), a feat that allows humans to produce a wide repertoire of speech sounds (i.e., phonemes).

As stated in the first section of this chapter, the speech acoustic wave can be decomposed into a carrier signal that contains the fine structure and a modulatory signal that defines the envelope of the sound. From a speaker’s perspective, the precise coordination of the displacement of the vocal tract articulators (i.e., jaw, tongue, lips, and velum) assigns a given phoneme its desired acoustic features and gives it its particular signature (e.g., the formant frequencies, the formant transition, or the frequency band of a fricative noise; Stevens, 1999). On the other hand, the amplitude of the sound (envelope) is less dependent on the precise position of the articulators and instead evolves proportionally to the airflow rate along the tract (Titze, 1994). In other words, the speech envelope depends on the degree of overall opening of the vocal tract (i.e., the more closed the vocal tract is, the lower the airflow rate at its exit resulting in a reduced sound volume). This implies that the rhythms of the speech envelope will reflect the temporal patterns of the opening and closing cycles of the vocal tract. A close temporal correspondence between the speech envelope and the area of the mouth opening has been empirically demonstrated, confirming this theoretical deduction (Chandrasekaran et al., 2009).

The fact that the dynamics of the vocal tract depict temporal regularities in the frequency range privileged by the perceptual system (previous section) invites the following question: Do speakers adopt this specific rhythm to satisfy the requirements of the listener? Or does this rhythm in production reflect the natural frequency of the speech motor system? To answer this question, we will review the literature characterizing the oscillatory nature of the brain areas in charge of generating speech.

The frequency band typically associated with motor cortex activity comprises a range between ~12 and 30 Hz (Kilavik et al., 2013). This band, known as the beta frequency band, follows a characteristic dynamic pattern; specifically, it is enhanced at rest (i.e., during what is known as *resting state*) (Congedo et al., 2010), suppressed during motor execution (specifically during speech production; Salmelin et al., 2000), and typically shows a postmovement rebound (Jurkiewicz et al., 2006; Saarinen et al., 2006). In addition to the beta band, oscillatory power in the theta band has also been reported in motor region activity during resting state. Furthermore, focusing on speech production related areas, Giraud and colleagues (2007) showed that a frequency band between 3 and 6 Hz is overrepresented in the spontaneous activity that originates in motor areas associated with mouth (lips and jaw) movements. This result invites the conjecture of a relationship between this natural frequency and the cycles of opening and closing of the vocal tract previously described in this section. Despite these compelling findings, the landscape of frequencies in the speech motor system, as in other brain regions, is likely to be more complex. Of particular interest, Keitel and Gross (2016) demonstrated that brain area activity during rest is not defined by a single oscillatory regime but is instead characterized by complex patterns comprising different frequency bands. In particular, they showed that power spectra of 1-second-long segments of MEG resting state computed on atlas-defined brain areas clustered into different spectral profiles. This implies that, at rest, the activity of each brain region displays a unique pattern of oscillatory activity comprising a variety of frequency bands. These bands do not necessarily coexist in time; instead, it is the amount of time that each frequency is activated that characterizes each specific region. Particularly relevant to this section are the spectral profiles of regions associated with speech production (see Fig. 4b). As shown in the figure, while all regions display activity in a set of frequency bands, in all cases theta is one of these bands that is significantly activated during rest. In short, the brain network responsible for speech production is broad and depicts a complex landscape of different oscillatory regimes during rest. Notwithstanding, it is important to stress that power in a frequency band aligning with the rhythmic patterns of the speech envelope (i.e., from 2 to 8 Hz) has been consistently reported during resting state activity in regions related to speech production.

Another way to explore the oscillatory nature of the speech production system is by assessing vocal production in preverbal infants. At an early stage of development, vocalizations are less culturally shaped, and the white matter pathways connecting motor areas with the rest of the brain are still underdeveloped (Brauer et al., 2013). This immature structural connectivity ensures that motor activity is less influenced by the activation of other brain regions. These conditions are optimal to probe the natural rhythms of the speech motor system. An infant's first vocalizations, known as marginal babbling, begin at 3 months and comprise nonspeech sounds like squealing or cooing. These vocalizations occur occasionally, and each attempt is limited to one of these long-voiced sounds (i.e., there is no rhythmic structure yet). Marginal babbling presumably allows the infant to acquire basic motor commands to activate the vocal folds and produce sounds (Vihman, 2014). Later in development, once these basic commands are acquired, speech-like vocalizations take place. Around 6 months, marginal babbling evolves to canonical babbling, which consists of stereotyped repetitions of a subset of syllables. Strikingly, canonical babbling shows regularities across languages, as does adult speech: the subset of repeated syllables is composed of a stop consonant followed by a vowel (i.e., a complete occlusion and later release of the vocal tract; for example, *ma-ma-ma-…*) and the repetition rate is close to 3 Hz (Dolata et al., 2008). These results indicate that, already early in development, the speech production system prefers rhythmic activity within the theta band.

Yet another means to indirectly assess the privileged frequencies of the speech production system harnesses the fact that motor and auditory brain areas are highly interactive. Anatomically, the arcuate fasciculus directly and indirectly connects auditory temporal regions with frontal motor and premotor regions (Catani et al., 2005). Functionally, speech motor areas are activated during passive listening to speech (Cogan et al., 2014; Du et al., 2014; Wilson et al., 2004). Exploiting this interaction, Assaneo and Poeppel (2018) investigated the *degree of synchrony* between motor and auditory areas during the passive listening of trains of syllables presented at different rates (from 2.5 to 6.5 syll/sec). As expected, they found an entrainment of speech motor areas to the syllable trains. Critically, this entrainment was restricted to a range of frequencies around 4.5 Hz. This result thus replicates previous observations, pointing to frequencies within the theta band as those favored by speech production areas.

Now, the question arises as to which kind of biophysical model can explain these observations implying that speech production areas prefer rhythms close to 4.5 Hz. A good candidate is the neural oscillator, an object widely described in computational neuroscience and shown to be biologically plausible (Wilson & Cowan, 1972). The main features that define a neural oscillator are as follows: (i) the ability to generate a rhythmic output at a given *natural* frequency without any rhythmic input, (ii) the ability to spontaneously synchronize to input signals but only when these signals oscillate at a frequency close to the oscillator's natural frequency, and (iii) the strength of the coupling between the oscillator and input signals determine how close the input frequency needs to be for synchronization to occur (Fig. 5a; Pikovsky et al., 2001). For a more detailed description of what an oscillator is, or is not, refer to (Doelling & Assaneo, 2021). Consistent with the observations so far reviewed in this section, researchers have proposed that (at least some) speech production areas can be modeled as a neural oscillator with a natural frequency close to 4.5 Hz, receiving the auditory activity as input (Fig. 5b; Poeppel & Assaneo, 2020). Such a model explains the enhancement of theta band activity in speech production areas during resting state (i.e., feature (i) of the neural oscillator); and the restricted frequency range of coupling between auditory and motor areas (i.e., feature (ii) of the neural oscillator) reported by Assaneo and Poeppel (2018). In addition, if the coupling between auditory and motor cortices is assumed to be proportional to the white matter pathway connecting them, the model predicts that only individuals with a certain degree of structural connectivity will manage to synchronize their speech production to rhythmic acoustic stimuli. In other words, individuals with weak structural connectivity between the areas will lie on the lower part of the graph in Fig. 5a and, for them, synchronization will occur only if the perceived rhythm is close to their natural rate. For these participants, a small amount of noise (any kind of biological system is influenced by deterministic laws as well as by randomness; Monod, 1971) will also bring them out of the synchronization region (shaded area in Fig. 5a). On the other hand, individuals with strong structural connectivity between areas will lie on the upper part of the graph and display robust synchronization for a wide range of frequencies of the external stimulus. In line with this prediction, data show that,

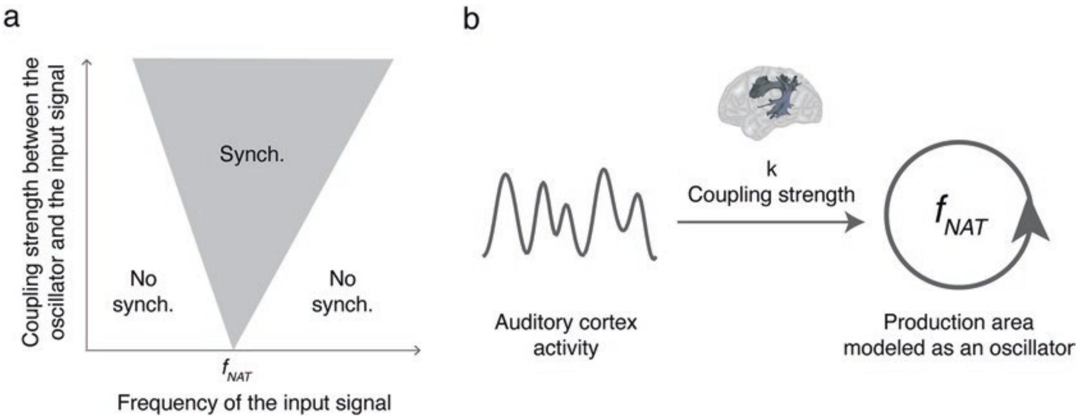


Fig. 5 A plausible biophysical model for speech production areas. (**a**) An oscillator's distinctive behavior. An oscillator synchronizes its activity to an external rhythmic input for combinations of parameters (input frequency, coupling strength) lying on the shaded region. This kind of visualization (i.e., parameters' space depicting areas where synchronization does, or does not, take place) is known as Arnold Tongue (Strogatz, 2018). Note that for very low values of the coupling synchronization, this will take place only if the input signal oscillates virtually at the natural frequency of the oscillator. (**b**) A biophysical model for the speech production system. Schematic representation of the model. Speech production area (right) can be modeled as an oscillator with a natural frequency (f_{NAT}) around 4.5 Hz, which receives as input the activity generated in the auditory cortex (left). The strength of the coupling between areas (k) is proportional to the structural connectivity between them (e.g., white matter volume). According to panel **c**, this model predicts that subjects with high volume in the arcuate fasciculus (i.e., high values of oscillator-input coupling strength) will show motor-to-auditory synchrony for a wide range of perceived syllabic rates. Conversely, a subject with low volume in this white matter structure will synchronize for a more restricted range of syllabic rates. In addition, any noise added to the system will bring this last subject out from the region of synchrony

when participants are instructed to overtly repeat a given syllable while listening to an isochronous syllabic stream, only some of them spontaneously synchronize their speech production to the external rhythm. Crucially, and in line with the theoretical assumption, these *synchronizers* show more volume than the *nonsynchronizers* in the left arcuate fasciculus (Assaneo et al., 2019), the main white matter pathway connecting auditory and motor cortical regions.

Returning to our initial question (i.e., does the speaker adopt a theta rhythm to satisfy the requirements of the listener or does this rhythm reflect the natural frequency of the speech motor system), all the evidence summarized in this section suggest that the stability of the syllabic rate across languages is a consequence of the rhythmic activity that naturally occurs in brain regions responsible for speech production. Remarkably, these speech motor rhythms appear to be evolutionarily preserved, as suggested by research on nonhuman primates. Lip smacking, a stereotyped noisy mouth movement subserving social communication, which has been proposed as a precursor of language, shows temporal regularities around 5 Hz in macaques (Ghazanfar et al., 2012) and 4 Hz in chimpanzees (Pereira et al., 2020). In addition, the vocal folds of marmosets synchronize their activity to the vocal tract movements in the theta band during the utterance of long-distance contact calls (Risueno-Segovia & Hage, 2020).

How do computational models of speech production account for all these observations regarding the theta rhythm? It is probably fair to say that the theta rhythm is not dealt with explicitly in most models. However, a link is traced indirectly by the emphasis placed on the syllable as the primary unit of speech motor control that most theories make, both in linguistics (e.g., Dell, 1986; Levelt, 1989) and in cognitive neuroscience (e.g., Guenther, 2015; Hickok, 2012). Speech motor control refers to the processes involved in the planning and production of speech in terms of, for example, the goals that drive these processes (e.g., hitting syllable-like auditory targets), the nature of the motor plans and commands recruited to achieve these goals (e.g., syllable-like motor chunks), and how all these come into play when speaking. For instance, a recent theory, termed Hierarchical State Feedback Control (Hickok, 2012), explicitly poses (at least) two distinct but interacting hierarchical levels of speech motor control. At the higher level, the motor commands are in charge of the cycles of opening and closing of the vocal tract that will roughly code for syllables (as seen earlier in this and previous sections). At the lower level (i.e., within the higher level), the finer motor commands code for the articulatory detail (roughly phonemes) that is targeted at the end points of each half cycle movement (i.e., the most open and closed positions of the vocal tract movement). Articulatory features at close positions of the vocal tract will mostly pertain to consonants, while articulatory features at open positions will generally pertain to vowels (see Fig. 4a). The theory then goes further to defend that, although the two hierarchical levels necessarily interact, each level likely relates to a different type of sensory target (i.e., the sensory goal driving the motor command). Specifically, the higher level (syllables) involves auditory targets coded in auditory areas. The lower level (phonemes) involves somatosensory targets coded in somatosensory cortical areas. The reason for this is that, because of coarticulation (i.e., the influence of one speech gesture over its neighbors), individual phoneme segments often do not have reliable acoustic realizations and thus cannot be coded in auditory space. However, they do have clear somatosensory consequences as end point vocal tract configurations (e.g., lip closure for a "p" or back tongue raising for a "k"), which would thus define the targets of this level of motor control. The hypothesis, therefore, is that speech acts are defined by higher-level goals, namely hitting syllable-level auditory targets. Within these higher-level motor commands in charge of each opening-closing cycle, there are also subgoals pertaining to the end points of each half-cycle, namely somatosensory targets related to articulatory features that (roughly) define phonemes. This implicitly associates the higher-level motor commands primarily with the generation of the speech envelope, while the lower commands would be responsible for the fine structure (as defined in section "The Acoustic Signal"). Consequently, the theta rhythm would be most relevant in the process of planning and producing speech in the way it relates to its primary units (syllable-like acoustic goals and motor commands).

In sum, the dynamics of the vocal tract exhibit temporal regularities in the theta frequency range, which is also privileged as a natural rhythm of the speech motor neural system and is reflected in the units for speech motor control. In the next (final) section of this chapter, we will explore possible reasons for the relevance of theta rhythm from an emitter-receiver interaction perspective.

The Listener-Speaker Interaction

So far, we characterized the three main agents composing speech, namely the acoustic signal, the listener, and the speaker, and we have shown that all exhibit rhythmic features within the same frequency range. The envelope of the acoustic signal displays a relatively stable temporal pattern across languages, implying that the mean syllabic rate is restricted to a range from 2 to 8 syllables per second (theta band). This quasi-rhythmic pattern is recovered by the perceptual system, which seems to rely on these temporal regularities for comprehension. Moreover, beyond speech perception, the auditory system is tuned to efficiently process rhythmic sounds within the theta band. Finally, from a speaker perspective, we showed that the temporal patterns of speech reflect the oscillatory properties of the brain areas in charge of generating speech. In summary, the speech production and perception systems appear to operate optimally in the same frequency range. In this section, we will review the role oscillations play in other cognitive domains and more generally at the brain level, while trying to elucidate why both systems (i.e., speech perception and production) evolved to favor the same rhythms.

The study of oscillatory regimes on brain activity began in the early twentieth century (Berger, 1930) and is still an active field of study (Giraud, 2020). After many decades of investigation and debate about the amount of frequency bands being enhanced and the functional role played by them, some agreement has been reached. Today, it is widely accepted that brain activity displays temporal regularities comprising a set of frequency bands (Buzsáki, 2006), namely gamma (~40 to 80 Hz), beta (~13 to 35 Hz), alpha (~8 to 12 Hz), theta (~3 to 8 Hz), and delta (~0.5 to 3 Hz). The limits of each band are not monolithically defined, and variability can be found in the literature. However, the hierarchy is preserved (i.e., with delta as the slowest rhythm and gamma the fastest) and the bands are always centered around the same values. Critically, this pattern of oscillatory activity remains stable across species, suggesting that it constitutes a fundamental feature of neural systems (Buzsáki et al., 2013). Although consensus has been reached on the oscillatory nature of brain activity, the functional role played by oscillations and their interaction is still not completely understood (Singer, 2018). In fact, heterogeneous sets of cognitive roles have been assigned to the different frequency bands (e.g., Davis et al., 2012; Lee et al., 2018; Moran & Hong, 2011; Roux & Uhlhaas, 2014; Symons et al., 2016), with scarce efforts to unify the existing lines of research. Of specific interest for this last section is the "parsing role" assigned to theta. It has been suggested that this frequency band serves the sensory systems by aligning the input's temporal dynamics with endogenous rhythm of the brain for a successful later decoding of the input signal by higher-level processing areas (Giraud, 2020; Kayser et al., 2012).

Given the existing literature showing that oscillatory activity is a characteristic feature of the brain and suggesting that it optimally processes rhythmic information, how do perceptual systems deal with the continuous flow of sensory experience? According to the *active sensing* theory, organisms rely on their motor system to sample the continuous sensory environment assigning to the input signal the required temporal granularity (Schroeder et al., 2010). Rhythmic movements within the theta band applied to sampling of the surroundings have been shown across sensory systems and species, thus supporting this proposal. Here we will mention only some examples of these rhythmic sensory-oriented movements. It has been shown that, during visual exploration, both human and nonhuman primates produce rapid and rhythmic eye movements (saccades) ranging from 3 to 8 Hz (Hoffman et al., 2013). Most mammals rely on sniffing (a repetitive sequence of short inhalation-exhalation cycles) to identify odors. Sniffing occurs rhythmically across species, with a characteristic sniffing frequency ranging from 2 to 8 Hz (Spencer et al., 2021). In order to identify the roughness of a surface, a person may move a finger back and forth over it. These tactile movement patterns have been shown to be approximately sinusoidal, with frequencies around 4 Hz (Morley et al., 1983). Rodents rely on licking (a rhythmic pattern of lingual protrusion-retraction) to make gustatory judgments. The licking rhythm also occurs within the theta range (Travers et al., 1997). These examples should suffice to realize that all sensory systems, except the auditory system, appear to be aided by a dynamic peripheral structure able to produce rhythmic movements to sample the environment. In this regard, it has been proposed that the rhythmic structure of speech may be a solution to the lack of a hardware mechanism for the sampling of the auditory environment (i.e., no dedicated movements to sample the continuous sound environment) (Lakatos et al., 2019; Poeppel & Assaneo, 2020). Therefore, during speech perception, no active sensing is necessary given that the temporal regularities required by the perceptual system for optimal decoding of the sensory environment are already present in the signal to be sensed (i.e., the acoustic speech signal). This may explain why speech perception and production systems evolved preferred rates that are aligned in frequency. This phenomenon could thus be seen as an active sensing case featuring two agents, with the motor system parsing the signal on the speaker's side automatically chunking the emitted signal to fulfill the requirements of the listener's sensory system.

This bipartite view of speech production and perception, whereby ongoing oscillations in the auditory system of the listener track or entrain to the rhythms produced by the speaker, is reasonable in an almost idealistic context of speech rhythmicity. However, although speech appears to be rhythmic in the theta range *on average*, a given instance of speech (say, a phrase or sentence) is usually no better than *quasi*-rhythmic. How does the auditory system deal with this variability? A recent hypothesis is that, beyond the automatic coupling of auditory regions to the input, the listener makes use (possibly unconsciously) of predictive temporal cues in the speech stream (e.g., gradual rate fluctuations, syntactic or semantic information) to realign ongoing oscillations with the speech signal (Rimmele et al., 2018; ten Oever & Martin, 2021).

The speech motor system is believed to be the main source of this predictive modulation of auditory activity. Because of this, temporal predictions are often understood as a *covert* form of active sensing (i.e., using the motor system to sample the sensory environment) (Morillon & Schroeder, 2015). This covert form of active sensing could potentially exploit the motor system's ability to generate predictions about the consequences of impending actions, generally used for online motor control and monitoring, to implement the predictive realignment of ongoing oscillations. In line with this, data show that motor and higher-order processing areas predictively modulate speech-entrained, low-frequency oscillations in auditory regions (Keitel et al., 2018; Morillon & Baillet, 2017), enhancing their tracking of the speech signal (Park et al., 2015). Whether oscillatory realignment via the speech motor system involves strictly temporal information or whether it additionally involves predictions about the content of upcoming speech events is still an unresolved question. Whichever the case, this mechanism affords the perceptual auditory system with the necessary robustness and flexibility to deal with the complex signal that constitutes speech.

As a final remark, we would like to acknowledge that, despite this chapter's focus on the oscillatory properties of speech within the theta band, other temporal regularities have been investigated in the speech literature. For example, slower oscillations within the delta (<3 Hz) band are also thought to play a crucial role from a perceptual perspective, subserving the parsing and processing of longer constituents such as phrases (Kösem & van Wassenhove, 2017; Meyer, 2018; Rimmele et al., 2021). However, at the moment, the presence of these frequencies *in the acoustic signal* has been theoretically suggested but not empirically demonstrated. Another example, related to absolute instead of regular interval (i.e., rhythmic) timing perception, is voice onset time (VOT). In phonetics (Cho & Ladefoged, 1999), VOT refers to the time interval between the release of the vocal tract constriction for stop consonants and the onset of voicing. VOT can be measured from the acoustic signal, given that both articulatory events have clear signatures on the spectrogram, and has been shown to be critical for speech perception in allowing the listener to distinguish between different consonants. Researchers have relied on this parameter to characterize language acquisition as well as some speech motor disorders (Englund, 2005; Pascal Auzou et al., 2000; Zlatin & Koenigsknecht, 1976). Time-related phenomena, such as the delta rhythm and VOT, highlight the relevance of temporal regularities in the perception and production of speech other than the theta rhythm. Still other examples can be found in the literature, given that speech is ultimately defined by an acoustic signal unfolding in time (Kotz & Schwartze, 2010). The choice to focus our attention on theta rhythms rests on the facts that they have been consistently reported in all three speech dimensions (i.e., the acoustic signal, the listener, and the speaker) and that their study has allowed researchers to infer biophysical properties of the underlying neural network. Temporal regularities around 4 Hz are present in the acoustic speech signal across languages, are required by the perceptual system to achieve comprehension, and seem to emerge from the timescales characteristic of speech motor brain regions.

Acknowledgments We thank Jessica González Norris for proofreading the manuscript. This work was supported by UNAM-DGAPA-PAPIIT IA200223 (MFA).

References

Abercrombie, D. (2022). Elements of general phonetics. In *Elements of general phonetics*. Edinburgh University Press. https://doi.org/10.1515/9781474463775

Abrams, D. A., Nicol, T., Zecker, S., & Kraus, N. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *Journal of Neuroscience, 29*(24), 7686–7693. https://doi.org/10.1523/JNEUROSCI.5242-08.2009

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences, 98*(23), 13367–13372. https://doi.org/10.1073/pnas.201400998

Assaneo, M. F., & Poeppel, D. (2018). The coupling between auditory and motor cortices is rate-restricted: Evidence for an intrinsic speech-motor rhythm. *Science Advances, 4*(2), eaao3842. https://doi.org/10.1126/sciadv.aao3842

Assaneo, M. F., Ripollés, P., Orpella, J., Lin, W. M., de Diego-Balaguer, R., & Poeppel, D. (2019). Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nature Neuroscience, 22*(4), 627–632. https://doi.org/10.1038/s41593-019-0353-z

Berger, H. (1930). Ueber das Elektrenkephalogramm des Menschen. [Electrocephalography in man.]. *Journal für Psychologie und Neurologie, 40*, 160–179.

Boemio, A., Fromm, S., Braun, A., & Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nature Neuroscience, 8*(3), 389–395. https://doi.org/10.1038/nn1409

Brauer, J., Anwander, A., Perani, D., & Friederici, A. D. (2013). Dorsal and ventral pathways in language development. *Brain and Language, 127*(2), 289–295. https://doi.org/10.1016/j.bandl.2013.03.001

Buzsáki, G. (2006). *Rhythms of the brain*. Oxford University Press.

Buzsáki, G., Logothetis, N., & Singer, W. (2013). Scaling brain size, keeping timing: Evolutionary preservation of brain rhythms. *Neuron, 80*(3), 751–764. https://doi.org/10.1016/j.neuron.2013.10.002

Castellucci, G. A., Kovach, C. K., Howard, M. A., Greenlee, J. D. W., & Long, M. A. (2022). A speech planning network for interactive language use. *Nature, 602*(7895), 117–122. https://doi.org/10.1038/s41586-021-04270-z

Catani, M., Jones, D. K., & Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology, 57*(1), 8–16. https://doi.org/10.1002/ana.20319

Chandrasekaran, C., Trubanova, A., Stillittano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology, 5*(7), e1000436. https://doi.org/10.1371/journal.pcbi.1000436

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*(2), 207–229. https://doi.org/10.1006/jpho.1999.0094

Cogan, G. B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., & Pesaran, B. (2014). Sensory–motor transformations for speech occur bilaterally. *Nature, 507*(7490), 94–98. https://doi.org/10.1038/nature12935

Congedo, M., John, R. E., De Ridder, D., & Prichep, L. (2010). Group independent component analysis of resting state EEG in large normative samples. *International Journal of Psychophysiology, 78*(2), 89–99. https://doi.org/10.1016/j.ijpsycho.2010.06.003

Davis, N. J., Tomlinson, S. P., & Morgan, H. M. (2012). The role of beta-frequency neural oscillations in motor control. *Journal of Neuroscience, 32*(2), 403–404. https://doi.org/10.1523/JNEUROSCI.5106-11.2012

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., & Poeppel, D. (2017). Temporal modulations in speech and music. *Neuroscience & Biobehavioral Reviews, 81*, 181–187. https://doi.org/10.1016/j.neubiorev.2017.02.011

Doelling, K. B., & Assaneo, M. F. (2021). Neural oscillations are a start toward understanding brain activity rather than the end. *PLoS Biology, 19*(5), e3001234. https://doi.org/10.1371/journal.pbio.3001234

Doelling, K. B., Arnal, L. H., Ghitza, O., & Poeppel, D. (2014). Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage, 85*, 761–768. https://doi.org/10.1016/j.neuroimage.2013.06.035

Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B., & Poeppel, D. (2019). An oscillator model better predicts cortical entrainment to music. *Proceedings of the National Academy of Sciences, 116*(20), 10113–10121. https://doi.org/10.1073/pnas.1816414116

Dolata, J. K., Davis, B. L., & MacNeilage, P. F. (2008). Characteristics of the rhythmic organization of vocal babbling: Implications for an amodal linguistic rhythm. *Infant Behavior and Development, 31*(3), 422–431. https://doi.org/10.1016/j.infbeh.2007.12.014

Du, Y., Buchsbaum, B. R., Grady, C. L., & Alain, C. (2014). Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proceedings of the National Academy of Sciences, 111*(19), 7126–7131. https://doi.org/10.1073/pnas.1318738111

Elliott, T. M., & Theunissen, F. E. (2009). The modulation transfer function for speech intelligibility. *PLoS*

Computational Biology, 5(3), e1000302. https://doi.org/10.1371/journal.pcbi.1000302

Englund, K. T. (2005). Voice onset time in infant directed speech over the first six months. *First Language, 25*(2), 219–234. https://doi.org/10.1177/0142723705050286

Faulkner, A., Rosen, S., & Smith, C. (2000). Effects of the salience of pitch and periodicity information on the intelligibility of four-channel vocoded speech: Implications for cochlear implants. *The Journal of the Acoustical Society of America, 108*(4), 1877–1887. https://doi.org/10.1121/1.1310667

Fuglsang, S. A., Madsen, K. H., Puonti, O., Hjortkjær, J., & Siebner, H. R. (2022). Mapping cortico-subcortical sensitivity to 4 Hz amplitude modulation depth in human auditory system with functional MRI. *NeuroImage, 246*, 118745. https://doi.org/10.1016/j.neuroimage.2021.118745

Ghazanfar, A. A., Takahashi, D. Y., Mathur, N., & Fitch, W. T. (2012). Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Current Biology, 22*(13), 1176–1182. https://doi.org/10.1016/j.cub.2012.04.055

Ghitza, O. (2011). Linking speech perception and neurophysiology: Speech decoding guided by cascaded oscillators locked to the input rhythm. *Frontiers in Psychology, 2*. https://www.frontiersin.org/article/10.3389/fpsyg.2011.00130

Ghitza, O., & Greenberg, S. (2009). On the possible role of brain rhythms in speech perception: Intelligibility of time-compressed speech with periodic and aperiodic insertions of silence. *Phonetica, 66*(1–2), 113–126. https://doi.org/10.1159/000208934

Giraud, A.-L. (2020). Oscillations for all ¯_(ツ)_/¯? A commentary on Meyer, Sun & Martin (2020). *Language, Cognition and Neuroscience, 35*(9), 1106–1113. https://doi.org/10.1080/23273798.2020.1764990

Giraud, A.-L., & Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. *Nature Neuroscience, 15*(4), 511–517. https://doi.org/10.1038/nn.3063

Giraud, A.-L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., & Kleinschmidt, A. (2000). Representation of the temporal envelope of sounds in the human brain. *Journal of Neurophysiology, 84*(3), 1588–1598. https://doi.org/10.1152/jn.2000.84.3.1588

Giraud, A.-L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S. J., & Laufs, H. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron, 56*(6), 1127–1134. https://doi.org/10.1016/j.neuron.2007.09.038

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends in Cognitive Sciences, 15*(1), 3–10. https://doi.org/10.1016/j.tics.2010.10.001

Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech—A syllable-centric perspective. *Journal of Phonetics, 31*(3), 465–485. https://doi.org/10.1016/j.wocn.2003.09.005

Guenther, F. H. (2015). *Neural control of speech*. The MIT Press.

Henry, M. J., & Obleser, J. (2012). Frequency modulation entrains slow neural oscillations and optimizes human listening behavior. *Proceedings of the National Academy of Sciences, 109*(49), 20095–20100. https://doi.org/10.1073/pnas.1213390109

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience, 13*(2), 135–145. https://doi.org/10.1038/nrn3158

Hoffman, K., Dragan, M., Leonard, T., Micheli, C., Montefusco-Siegmund, R., & Valiante, T. (2013). Saccades during visual exploration align hippocampal 3–8 Hz rhythms in human and non-human primates. *Frontiers in Systems Neuroscience, 7*. https://www.frontiersin.org/article/10.3389/fnsys.2013.00043

Houtgast, T., & Steeneken, H. J. M. (1985). A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria. *The Journal of the Acoustical Society of America, 77*(3), 1069–1077. https://doi.org/10.1121/1.392224

Howard, M. F., & Poeppel, D. (2010). Discrimination of speech stimuli based on neuronal response phase patterns depends on acoustics but not comprehension. *Journal of Neurophysiology, 104*(5), 2500–2511. https://doi.org/10.1152/jn.00251.2010

International Phonetic Association. (1999). *Handbook of the International Phonetic Association: A guide to the use of the International Phonetic Alphabet*. Cambridge University Press.

Jurkiewicz, M. T., Gaetz, W. C., Bostan, A. C., & Cheyne, D. (2006). Post-movement beta rebound is generated in motor cortex: Evidence from neuromagnetic recordings. *NeuroImage, 32*(3), 1281–1289. https://doi.org/10.1016/j.neuroimage.2006.06.005

Kalashnikova, M., Peter, V., Di Liberto, G. M., Lalor, E. C., & Burnham, D. (2018). Infant-directed speech facilitates seven-month-old infants' cortical tracking of speech. *Scientific Reports, 8*(1), 13745. https://doi.org/10.1038/s41598-018-32150-6

Kayser, C., Ince, R. A. A., & Panzeri, S. (2012). Analysis of slow (theta) oscillations as a potential temporal reference frame for information coding in sensory cortices. *PLoS Computational Biology, 8*(10), e1002717. https://doi.org/10.1371/journal.pcbi.1002717

Keitel, A., & Gross, J. (2016). Individual human brain areas can be identified from their characteristic spectral activation fingerprints. *PLoS Biology, 14*(6), e1002498. https://doi.org/10.1371/journal.pbio.1002498

Keitel, A., Gross, J., & Kayser, C. (2018). Perceptually relevant speech tracking in auditory and motor cortex reflects distinct linguistic features. *PLoS Biology, 16*(3), e2004473. https://doi.org/10.1371/journal.pbio.2004473

Kilavik, B. E., Zaepffel, M., Brovelli, A., MacKay, W. A., & Riehle, A. (2013). The ups and downs of beta oscillations in sensorimotor cortex. *Experimental Neurology, 245*, 15–26. https://doi.org/10.1016/j.expneurol.2012.09.014

Kösem, A., & van Wassenhove, V. (2017). Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. *Language, Cognition and Neuroscience, 32*(5), 536–544. https://doi.org/10.1080/23273798.2016.1238495

Kotz, S. A., & Schwartze, M. (2010). Cortical speech processing unplugged: A timely subcortico-cortical framework. *Trends in Cognitive Sciences, 14*(9), 392–399. https://doi.org/10.1016/j.tics.2010.06.005

Lakatos, P., Gross, J., & Thut, G. (2019). A new unifying account of the roles of neuronal entrainment. *Current Biology, 29*(18), R890–R905. https://doi.org/10.1016/j.cub.2019.07.075

Lee, C. S., Kitamura, C., Burnham, D., & McAngus Todd, N. P. (2014). On the rhythm of infant- versus adult-directed speech in Australian English. *The Journal of the Acoustical Society of America, 136*(1), 357–365. https://doi.org/10.1121/1.4883479

Lee, D. J., Kulubya, E., Goldin, P., Goodarzi, A., & Girgis, F. (2018). Review of the neural oscillations underlying meditation. *Frontiers in Neuroscience, 12*. https://www.frontiersin.org/article/10.3389/fnins.2018.00178

Levelt, W. J. M. (1989). *Speaking: From intention to articulation* (p. xiv, 566). The MIT Press.

Liberman, A. M. (1996). *Speech: A special code*. A Bradford Book.

Long, M. A., Katlowitz, K. A., Svirsky, M. A., Clary, R. C., Byun, T. M., Majaj, N., Oya, H., Howard, M. A., & Greenlee, J. D. W. (2016). Functional segregation of cortical regions underlying speech timing and articulation. *Neuron, 89*(6), 1187–1193. https://doi.org/10.1016/j.neuron.2016.01.032

Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron, 54*(6), 1001–1010.

Meyer, L. (2018). The neural oscillations of speech processing and language comprehension: State of the art and emerging mechanisms. *European Journal of Neuroscience, 48*(7), 2609–2621. https://doi.org/10.1111/ejn.13748

Monod, J. (1971). *Chance and necessity: An essay on the natural philosophy of modern biology* (1st American ed.). Knopf.

Moran, L. V., & Hong, L. E. (2011). High vs low frequency neural oscillations in schizophrenia. *Schizophrenia Bulletin, 37*(4), 659–663. https://doi.org/10.1093/schbul/sbr056

Morillon, B., & Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proceedings of the National Academy of Sciences, 114*(42), E8913–E8921. https://doi.org/10.1073/pnas.1705373114

Morillon, B., & Schroeder, C. E. (2015). Neuronal oscillations as a mechanistic substrate of auditory temporal prediction. *Annals of the New York Academy of Sciences, 1337*(1), 26–31. https://doi.org/10.1111/nyas.12629

Morley, J. W., Goodwin, A. W., & Darian-Smith, I. (1983). Tactile discrimination of gratings. *Experimental Brain Research, 49*(2), 291–299. https://doi.org/10.1007/BF00238588

Nolan, F., & Jeon, H.-S. (2014). Speech rhythm: A metaphor? *Philosophical Transactions of the Royal Society, B: Biological Sciences, 369*(1658), 20130396. https://doi.org/10.1098/rstb.2013.0396

Park, H., Ince, R. A. A., Schyns, P. G., Thut, G., & Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Current Biology, 25*(12), 1649–1653. https://doi.org/10.1016/j.cub.2015.04.049

Pascal Auzou, C. O., Morris, R. J., Jan, M., Eustache, F., & Hannequin, D. (2000). Voice onset time in aphasia, apraxia of speech and dysarthria: A review. *Clinical Linguistics & Phonetics, 14*(2), 131–150. https://doi.org/10.1080/026992000298878

Peelle, J. E., & Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Frontiers in Psychology, 3*, 320.

Peña, M., & Melloni, L. (2012). Brain oscillations during spoken sentence processing. *Journal of Cognitive Neuroscience, 24*(5), 1149–1164. https://doi.org/10.1162/jocn_a_00144

Pereira, A. S., Kavanagh, E., Hobaiter, C., Slocombe, K. E., & Lameira, A. R. (2020). Chimpanzee lip-smacks confirm primate continuity for speech-rhythm evolution. *Biology Letters, 16*(5), 20200232. https://doi.org/10.1098/rsbl.2020.0232

Pikovsky, A., Rosenblum, M., & Kurths, J. (2001). *Synchronization: A universal concept in nonlinear sciences*. Cambridge University Press. https://doi.org/10.1017/CBO9780511755743

Poeppel, D., & Assaneo, M. F. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience, 21*(6), 322–334. https://doi.org/10.1038/s41583-020-0304-4

Riecke, L., Formisano, E., Sorger, B., Başkent, D., & Gaudrain, E. (2018). Neural entrainment to speech modulates speech intelligibility. *Current Biology, 28*(2), 161–169.e5. https://doi.org/10.1016/j.cub.2017.11.033

Rimmele, J. M., Morillon, B., Poeppel, D., & Arnal, L. H. (2018). Proactive sensing of periodic and aperiodic auditory patterns. *Trends in Cognitive Sciences, 22*(10), 870–882. https://doi.org/10.1016/j.tics.2018.08.003

Rimmele, J. M., Poeppel, D., & Ghitza, O. (2021). Acoustically driven cortical δ oscillations underpin prosodic chunking. *eNeuro, 8*(4). https://doi.org/10.1523/ENEURO.0562-20.2021

Risueno-Segovia, C., & Hage, S. R. (2020). Theta synchronization of phonatory and articulatory systems in marmoset monkey vocal production. *Current Biology, 30*(21), 4276–4283.e3. https://doi.org/10.1016/j.cub.2020.08.019

Roux, F., & Uhlhaas, P. J. (2014). Working memory and neural oscillations: Alpha–gamma versus theta–gamma codes for distinct WM information? *Trends in Cognitive Sciences, 18*(1), 16–25. https://doi.org/10.1016/j.tics.2013.10.010

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical*

Transactions of the Royal Society of London. Series B: Biological Sciences, 336(1278), 367–373. https://doi.org/10.1098/rstb.1992.0070

Saarinen, T., Laaksonen, H., Parviainen, T., & Salmelin, R. (2006). Motor cortex dynamics in visuomotor production of speech and non-speech mouth movements. *Cerebral Cortex, 16*(2), 212–222. https://doi.org/10.1093/cercor/bhi099

Salmelin, R., Schnitzler, A., Schmitz, F., & Freund, H.-J. (2000). Single word reading in developmental stutterers and fluent speakers. *Brain, 123*(6), 1184–1202. https://doi.org/10.1093/brain/123.6.1184

Schroeder, C. E., Wilson, D. A., Radman, T., Scharfman, H., & Lakatos, P. (2010). Dynamics of active sensing and perceptual selection. *Current Opinion in Neurobiology, 20*(2), 172–176. https://doi.org/10.1016/j.conb.2010.02.010

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*(5234), 303–304. https://doi.org/10.1126/science.270.5234.303

Singer, W. (2018). Neuronal oscillations: Unavoidable and useful? *European Journal of Neuroscience, 48*(7), 2389–2398. https://doi.org/10.1111/ejn.13796

Spencer, T. L., Clark, A., Fonollosa, J., Virot, E., & Hu, D. L. (2021). Sniffing speeds up chemical detection by controlling air-flows near sensors. *Nature Communications, 12*(1), 1232. https://doi.org/10.1038/s41467-021-21405-y

Stevens, K. N. (1999). *Acoustic phonetics*. The MIT Press.

Strogatz, S. H. (2018). *Nonlinear dynamics and chaos with student solutions manual: With applications to physics, biology, chemistry, and engineering* (2nd ed.). CRC Press.

Symons, A. E., El-Deredy, W., Schwartze, M., & Kotz, S. A. (2016). The functional role of neural oscillations in non-verbal emotional communication. *Frontiers in Human Neuroscience, 10*. https://www.frontiersin.org/article/10.3389/fnhum.2016.00239

ten Oever, S., & Martin, A. E. (2021). An oscillating computational model can track pseudo-rhythmic speech by using linguistic predictions. *eLife, 10*, e68066. https://doi.org/10.7554/eLife.68066

Titze, I. R. (1994). *Principles of voice production*. Prentice Hall.

Travers, J. B., Dinardo, L. A., & Karimnamazi, H. (1997). Motor and premotor mechanisms of licking. *Neuroscience & Biobehavioral Reviews, 21*(5), 631–647. https://doi.org/10.1016/S0149-7634(96)00045-0

Varnet, L., Ortiz-Barajas, M. C., Erra, R. G., Gervain, J., & Lorenzi, C. (2017). A cross-linguistic study of speech modulation spectra. *The Journal of the Acoustical Society of America, 142*(4), 1976–1989. https://doi.org/10.1121/1.5006179

Versfeld, N. J., & Dreschler, W. A. (2002). The relationship between the intelligibility of time-compressed speech and speech in noise in young and elderly listeners. *The Journal of the Acoustical Society of America, 111*(1), 401–408. https://doi.org/10.1121/1.1426376

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *The Journal of the Acoustical Society of America, 66*(5), 1364–1380. https://doi.org/10.1121/1.383531

Vihman, M. M. (2014). *Phonological development: The first two years* (2nd ed.). Wiley-Blackwell.

Wilsch, A., Neuling, T., Obleser, J., & Herrmann, C. S. (2018). Transcranial alternating current stimulation with speech envelopes modulates speech comprehension. *NeuroImage, 172*, 766–774. https://doi.org/10.1016/j.neuroimage.2018.01.038

Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal, 12*(1), 1–24. https://doi.org/10.1016/S0006-3495(72)86068-5

Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience, 7*(7), 701–702. https://doi.org/10.1038/nn1263

Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences, 6*(1), 37–46. https://doi.org/10.1016/S1364-6613(00)01816-7

Zlatin, M. A., & Koenigsknecht, R. A. (1976). Development of the voicing contrast: A comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research, 19*(1), 93–111. https://doi.org/10.1044/jshr.1901.93

Zoefel, B., ten Oever, S., & Sack, A. T. (2018). The involvement of endogenous neural oscillations in the processing of rhythmic input: More than a regular repetition of evoked neural responses. *Frontiers in Neuroscience, 12*. https://doi.org/10.3389/fnins.2018.00095

Timing Patterns in the Extended Basal Ganglia System

Michael Schwartze and Sonja A. Kotz

Abstract

The human brain is a constructive organ. It generates predictions to modulate its functioning and continuously adapts to a dynamic environment. Increasingly, the temporal dimension of motor and non-motor behaviour is recognised as a key component of this predictive bias. Nevertheless, the intricate interplay of the neural mechanisms that encode, decode and evaluate temporal information to give rise to a sense of time and control over sensorimotor timing remains largely elusive. Among several brain systems, the basal ganglia have been consistently linked to interval- and beat-based timing operations. Considering the tight embedding of the basal ganglia into multiple complex neurofunctional networks, it is clear that they have to interact with other proximate and distal brain systems. While the primary target of basal ganglia output is the thalamus, many regions connect to the striatum of the basal ganglia, their main input relay. This establishes widespread connectivity, forming the basis for first- and second-order interactions with other systems implicated in timing such as the cerebellum and supplementary motor areas. However, next to this structural interconnectivity, additional functions need to be considered to better understand their contribution to temporally predictive adaptation. To this end, we develop the concept of *interval-based patterning*, conceived as a temporally explicit hierarchical sequencing operation that underlies motor and non-motor behaviour as a common interpretation of basal ganglia function.

Keywords

Temporal processing · Interval timing · Patterning · Chunking · Temporal prediction

All state changes in an individual and in the environment generate events that are defined by their type ('what') and timing ('when'). To successfully interact with an ever-changing environment, individuals continuously adapt to these what and when dimensions of events (Schwartze & Kotz, 2013). In this fundamental action, humans not only react to but also predict events (Friston, 2009; Friston et al., 2006). This essentially affords optimised allocation of neural and cogni-

---

Michael Schwartze and Sonja A. Kotz contributed equally with all other contributors.

M. Schwartze · S. A. Kotz (✉)
Department of Neuropsychology and Psychopharmacology, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, Netherlands
e-mail: sonja.kotz@maastrichtuniversity.nl

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5_15

tive resources and timely action, e.g. in the case of an athlete commencing a sprint with the last sound of a starting signal, or a musician falling in with an orchestra on a specific note. The underlying principle is simple and thus readily transferable between different contexts: knowing when something happens affords better cognitive and behavioural adaptation to what happens.

The efficiency of this temporally predictive adaptation partially depends on the capacity to exploit 'when', i.e. timing information, to tune into the dynamic environment. Models of 'predictive coding' (Friston, 2009; Friston et al., 2006) suggest that this capacity involves a fundamental division of labour, in which slower neural dynamics inform and interact with faster neural dynamics to guide adaptation (Friston, 2012; Schwartze et al., 2012a). However, unlike 'what' information, which can be manifold (e.g. touch, movement, colour, words, musical tones), 'when' information is strictly one-dimensional and thus solely defined by the interplay of change (event) and persistence (interval; Fig. 1). Consequently, other timing-related characteristics such as regularity, periodicity, but also grouping, are ultimately temporal ordering principles that specify the configuration of change and persistence. However, although these components of timing are well-defined and quantifiable in physics, it has long been recognised that they are much less so for perception and neurocognitive functions, and thus require a distinct taxonomy of temporal experience to capture the mechanisms underlying temporally predictive adaptation in humans (Jones, 1976; Pöppel, 1978).

Neurocognitive processes add further subjective components to temporally predictive adaptation. A well-established example is the phenomenon of 'subjective accentuation', i.e. the emergence of a perceptual dissociation of 'strong' and 'weak' events when listening to a train of physically identical stimuli such as metronome clicks that commonly leads to the 'tick-tock' illusion (Brochard et al., 2003; Abecasis et al., 2005; Criscuolo et al., 2023). Physical and neurocognitive timing are therefore not necessarily equivalent in how temporal ordering principles are realised. However, both, either independent or in combination, can lead to perceived temporal regularity, which, in turn, seems a prerequisite for efficient temporally predictive adaptation. Perceived temporal regularity essentially corresponds to the recognition of a 'pattern *of* time', which also may or may not constitute a 'pattern *in* time' (Handel, 1974). Although the ability to produce, perceive and synchronise movement

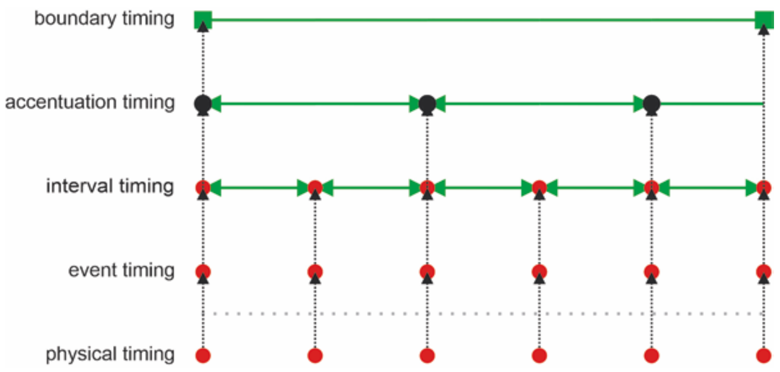


Fig. 1 Patterns of time. The interplay of change and persistence (e.g. the clicks of a metronome) generates successive events that constitute an environmental timing pattern (physical timing: red circles). Neural encoding generates representations of the temporal locus of events (event timing) and inter-event relations (interval timing: green lines). Additional mechanisms and functions can lead to subjective accentuation of some events (accentuation timing: black circles) and to distinct markings of the beginning and the end of a pattern (boundary timing: green squares)

with patterns of time is increasingly recognised as a fundamental of neurocognitive function (Merchant et al., 2015; Penhune & Zatorre, 2019), the dissociation of physical and neurocognitive timing warrants further differentiation of their basic constituents.

Perceived temporal regularity and grouping imply at least two intervals (and correspondingly three events). As indicated, additional mechanisms and functional components can lead to subjective accentuation over the course of a longer sequence but also mark boundaries, i.e. the beginning and the end of a pattern such as that established by an action sequence (Fujii & Graybiel, 2003; Graybiel, 2008). The resultant accentuation timing and boundary timing (Fig. 1) are essentially chunking phenomena that may relate to specific cognitive constructs and constraints, e.g. the dynamic allocation of attention and working memory capacities (Jones, 1976; Large & Jones, 1999; Schwartze et al., 2020). However, these examples illustrate why next to a taxonomy of temporal experience (Pöppel, 1978), additional mechanistic and functional components have to be considered to obtain a better understanding of temporally predictive adaptation as a form of interaction of the organism with a pattern structure (Jones, 1976).

Eventually, these additional subjective components are anchored in the brain and different, partly interdependent, neural mechanisms have been associated with the precise encoding of when an event occurs and the encoding of the respective inter-event relations or intervals that form the basic constituents of any pattern of time (Ivry & Schlerf, 2008; Spencer & Ivry, 2013; Buhusi & Meck, 2005; Bares et al., 2019).

Next to multiple other functions, chunking and interval timing specifically engage the subcortical basal ganglia system, which, in turn, interfaces with other cortical and subcortical systems. Together, these systems form a large-scale network that seems to support interleaved and differential aspects of adaptive timing (Buhusi & Meck, 2005; Ivry & Schlerf, 2008; Merchant et al., 2013; Petter et al., 2016). This interactive network architecture may not only explain phenomena such as subjective accentuation in timing but also how timing factors into other well-established basal ganglia functions, e.g. in relating an event to a subsequent reward (Schultz, 2004) or in triggering sequential behaviour (Graybiel, 2008).

The basal ganglia system and associated thalamocortical circuits link to prefrontal and supplementary motor cortices and the cerebellum. However, much less is known about if and how this core timing network interfaces with further systems and whether such interaction guides temporally predictive adaptation. One example of a direct functional interaction between several systems is the initiation and termination of the basal ganglia interval timing mechanism through dopaminergic bursts emitted by the ventral tegmental area (VTA; Buhusi & Meck, 2005; Petter et al., 2016). However, when considering such interactions between systems, critical questions arise such as how sensory input triggers these bursts in the first place, and how they respond to repeating events that delineate the consecutive intervals of a pattern. To answer these questions, it seems necessary to take a more holistic perspective on timing that considers further second-order interactions and information flow from the earliest stages of sensory processing.

Previous work developed such a holistic perspective into an integrative timing framework (Schwartze & Kotz, 2013). A central tenet of this framework is the adaptive cerebellar filtering of sensory input into an event-based stimulus representation and its rapid transmission to the thalamus as a means to translate physical timing into an intermittent neural ‘clock’ signal (Fig. 2).

Following the cerebellar encoding of event timing, the clock signal is transmitted to cortical targets that include the supplementary motor area (SMA), where successive events trigger continuous oscillatory activity at different frequencies. The thalamus relays and amplifies the signal by employing a bursting firing mode (Sherman, 2001). As originally suggested by the striatal beat frequency model of interval timing, a snapshot of the oscillatory activity at the end of an interval is integrated by the basal ganglia into a distinct code of an inter-event interval (‘timestamp’; Matell et al., 2003; Matell & Meck, 2004; Buhusi & Meck, 2005) and relayed to the cortex. However, one may speculate that the intermittent

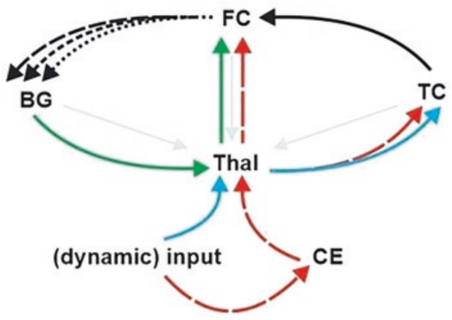


Fig. 2 A timing network. Dynamic input (here exemplified by a sound signal) is differentially encoded along ascending connections that include a direct bottom-up route via the cerebellum (CE). The CE performs temporally precise sparse coding of events (e.g. onsets, offsets, abrupt changes in energy) of the input into an intermittent (dashed red) 'clock signal' transmitted via the thalamus (THAL) to frontal cortices (FC). Thalamic signals trigger cortical oscillatory activity. A snapshot of the concerted cortical oscillatory activity (dashed black) is integrated by the basal ganglia (BG) into a relational (interval) representation of the time between events. Successive events end the current and open the next interval. Interval representations are transmitted to frontal cortices (green) for timing judgments and pattern recognition. Parallel to the intermittent clock signal, the ascending auditory pathway transmits a detailed continuous input representation. This signal reaches sensory cortices (here temporal cortex, TC) to interface with memory representations (blue) that are then conveyed to frontal areas to bridge input elements (e.g. successive words forming utterances, tones forming melodies). The thalamus also receives modulatory input from other structures, potentially allowing top-down modulation of functioning and reactivity (grey) to predictively adapt the organism to the environment

cerebellar clock signal and entirely subcortical pathways are not only instrumental in triggering thalamic bursting but also the dopaminergic bursting that opens and ends the integration of cortical oscillatory activity by the basal ganglia (Petter et al., 2016).

Among the cortical targets of the cerebellar clock signal the medial frontal SMA insofar plays a special role as it tightly connects to the basal ganglia as well as to prefrontal, premotor and primary motor cortices. Already in his original definition of the SMA, Wilder Penfield (Penfield, 1950) provided evidence that electrostimulation of this area just anterior of the sensorimotor cortex produces varied motor responses that include vocalisations. Most of the responses could be categorised as either slow and sustained (i.e. persistent) or rhythmic (i.e. patterned). SMA activity in monkeys has been shown to indicate total elapsed time as well as rhythm intervals (Cadena-Valencia et al., 2018). The SMA is furthermore one of the only brain systems that is consistently recruited in perceptual and motor timing tasks (Wiener et al., 2010). However, this general SMA engagement across different contexts and pattern levels also illustrates that in addition to a holistic perspective, further structural and functional differentiation of the systems that make up the timing network is critical to better understand temporally predictive adaptation.

The general importance of the SMA for timing is supported by its anatomical differentiation into the more anterior pre-SMA and the more posterior SMA-proper (Picard & Strick, 2001). Connectivity between the SMA and the striatum of the basal ganglia maintains this anterior-posterior organisation, while distinct cortical connections link the pre-SMA to prefrontal cortices and the SMA-proper to premotor and primary motor cortices (Picard & Strick, 2001; Lehéricy et al., 2004; Akkal et al., 2007; Kotz et al., 2013). This organisation is likely part and parcel of functional separations, with single and longer interval timing hosted by the pre-SMA and multiple interval timing by the SMA-proper (Schwartze et al., 2012b; Cona et al., 2021). Alternatively, it may reflect a shifting functional gradient that indicates the initial recruitment of single-interval pre-SMA timing in a cyclic manner, and an associated activation of the SMA-proper with an evolving pattern of time. In both cases, it seems likely that information flow through this anterior-posterior system would be modulated by factors such as interval duration or tasks demands (Coull et al., 2004; Macar et al., 2006), reflecting the distinct prefrontal and premotor connectivity patterns of pre-SMA and SMA-proper in interaction with cognitive processes such as attentional resource allocation, working memory, or movement control.

These interactive dynamics and functional gradients furthermore allow considering whether the same or at least similar concepts and mechanisms that have been identified in one field of research can be applied to another field of inquiry to derive refined and testable hypotheses. This rationale can be exemplified by SMA to basal ganglia connectivity and by linking basal ganglia functions that are expressed on a timescale that spans days or even years to the milliseconds-to-seconds range that is typically relevant for sensorimotor timing research. Accordingly, the previously documented role of the SMA in establishing a temporal link between action and effect (Moore et al., 2010) may differentially engage interval timing in a context-dependent manner that reflects interval duration or single interval as opposed to multiple interval patterns and corresponding recruitment of anterior-posterior SMA to basal ganglia connections. Similarly, the formation and expression of 'habits', both associated with the basal ganglia system (Graybiel, 2008; Smith & Graybiel, 2016), may engage interval timing in a cyclic manner (Fig. 3). According to Graybiel (2008), habits are largely defined by five central characteristics: they are learned, occur repeatedly, are performed almost automatically, manifest as a triggered action or thought sequence and can be completed without constant oversight. Structural and functional differentiation of the SMA and connections to the basal ganglia as identified for timing may factor into these characteristics, forming an explicitly temporal component of the general motor and cognitive patterning and chunking mechanisms that have been ascribed to the basal ganglia system (Graybiel, 1997, 1998).

Next to refined conceptions in several domains, this integrative perspective may also help to better explain a range of pathological phenomena. For example, insufficient differentiation along the structural and functional gradients of the timing network may diffuse production and perception of patterns of time, while compensatory strategies may target such insufficient differentiation. This may be the case in stuttering, a disorder that affects the temporal pattern of speech sequences. However, delayed auditory feedback can typically improve speech fluency dramatically in people who stutter. This effect may indicate a lower degree of automaticity or that speakers are prevented from hearing their own errors, reducing or prohibiting erroneous basal ganglia activity (Alm, 2004; Guenther & Hickok, 2016). Alternatively, the higher degree of temporal differentiation introduced by the

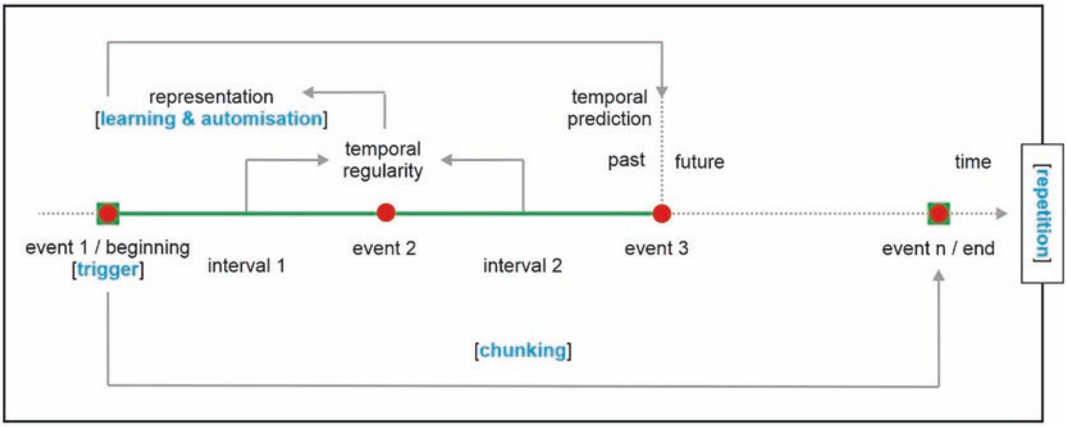


Fig. 3 Interval timing and habitual behaviour. The interval timing capacity of the basal ganglia system is initiated via a dopaminergic burst with the first event (event1/beginning) of sequential behaviour. Subsequent events close and open successive intervals until the final event ends the sequence (event n/end). The temporal regularity extracted from a pattern (through representation and learning) drives temporally adaptive behaviour in prediction of future events, e.g. by allocating attention to times at which events are predicted to occur. In turn, this facilitates automisation of sequential behaviour, e.g. by reducing overall attentional demands for rigid and habitual forms of behaviour. Subjective accentuation through 'boundary timing' facilitates the chunking of the sequence and specifies an interval that corresponds to the duration of the entire sequence

delay may improve a speaker's sense of agency and the differentiation of perceptual and sensorimotor timing as linked to the interplay of the SMA and basal ganglia. This could potentially explain why not only the iterative build-up of a pattern of time during speech production is affected but also that people who stutter can show subtle sensorimotor and sensory as well as nonverbal timing dysfunctions (Etchell et al., 2014; Falk et al., 2015; Schwartze et al., 2020).

Dysfunctional timing is also a hallmark of Parkinson's disease. Although Parkinson's disease is not a unitary pathology and comprises various forms and subtypes, the two most commonly associated structural and functional features are the well-documented cell loss in the substantia nigra pars compacta of the basal ganglia system and cardinal motor symptoms (Weingarten et al., 2015). However, Parkinson's disease is a progressive disease that affects multiple brain systems outside the basal ganglia, and also leads to non-motor symptoms that can precede motor symptoms in early non-medicated patients (Pont-Sunyer et al., 2015). Dysfunctional timing may manifest in disturbances of gait fluency, problems to produce and to maintain a steady movement pace, or an impaired ability to synchronise with pacing stimuli (Allman & Meck, 2012; Dalla Bella et al., 2015). However, it is currently largely unclear at which point in the progression of the neurodegenerative process timing is affected, suggesting that more subtle timing dysfunctions manifest before or after other typical symptoms. Especially if one considers the expression of a pattern of time as the outcome of general operations that define not only the order of successive motor but also of non-motor events and cognitive behaviour such as complex problem solving (Graybiel, 2008), it would be meaningful to explore the timing aspect of existing diagnostic tools more systematically. Dysfunctional timing may already be present and indicative of pre-diagnostic disease mechanisms due to the engagement of the basal ganglia, associated pathways and secondary regions in the selection and sequencing of most forms of motor and non-motor behaviour (Graybiel, 2008).

The notion of aberrant timing as an early marker of PD is further supported by neuroimaging studies that show dynamic patterns of SMA and cerebellar hyper- and hypo-activity that may indicate compensatory mechanisms during early disease stages (Kotz & Schwartze, 2011; Schwartze & Kotz, 2016). Complementing such findings, post-mortem studies of PD patients indicate a selective loss of pyramidal neurons in the SMA and suggest that this loss may precede basal ganglia pathology (McDonald & Halliday, 2002).

Although certainly selective, these examples clearly illustrate that the human capacity for temporally predictive adaptation to the environment is a fundamental but complex phenomenon. A better understanding of this capacity requires combining a taxonomy of temporal experience with a holistic neurofunctional perspective that explains how the human organism encodes and decodes patterns of time. One important open question concerns the dimensionality of facilitatory compared to pathological manifestations of temporal adaptation. Although temporally predictive adaptation necessarily involves an aspect of temporal regularity, overreliance on regularity might mark inflexible, habitual behaviour.

The main argument put forward here is that decomposition of basal ganglia function and their interaction with other brain systems in timing can be used to derive refined explanations and testable assumptions across different contexts. This approach has not least a number of methodological implications, as the parameters that define temporal experience and subjective components of temporally predictive adaptation impose several constraints for the design of respective empirical investigations, e.g. in terms of interval durations, stimulus rates, grouping, regularity, or jitter that specify the interplay of change and persistence.

References

Abecasis, D., Brochard, R., Granot, R., & Drake, C. (2005). Differential brain response to metrical accents in isochronous auditory sequences. *Music Perception, 22*, 549–562.

Akkal, D., Dum, R. P., & Strick, P. L. (2007). Supplementary motor area and presupplementary motor area: Targets of basal ganglia and cerebellar output. *The Journal of Neuroscience, 27*, 10659–10673.

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

Alm, P. A. (2004). Stuttering and the basal ganglia circuits: Acritical review of possible relations. *Journal of Communication Disorders, 37*, 325–336.

Bares, M., Apps, R., Avanzino, L., Breska, A., DÁngelo, E., Filip, P., Gerwig, M., Ivry, R. B., Lawrenson, C. L., Louis, E. D., Lusk, N. A., Manto, M., Meck, W. H., Mitoma, H., & Petter, E. A. (2019). Consensus paper: Decoding the contributions of the cerebellum as a time machine. From neurons to clinical applications. *Cerebellum, 18*, 266–286.

Brochard, R., Abecasis, D., Potter, D., Ragot, R., & Drake, C. (2003). The "ticktock" of our internal clock: Direct brain evidence of subjective accents in isochronous sequences. *Psychological Science, 14*, 362–366.

Buhusi, C. V., & Meck, W. H. (2005). What makes us tick?: Functional and neural mechanisms of interval timing. *Nature Reviews. Neuroscience, 6*, 755–765.

Cadena-Valencia, J., Garcia-Garibay, O., Merchant, H., Jazayeri, M., & de Lafuente, V. (2018). Entrainment and maintenance of an internal metronome in supplementary motor area. *eLife, 7*, e38983.

Cona, G., Wiener, M., & Scarpazza, C. (2021). From ATOM to GradiATOM: Cortical gradients support time and space processing as revealed by a meta-analysis of neuroimaging studies. *NeuroImage, 224*, 117407.

Coull, J. T., Vidal, F., Nazarian, B., & Macar, F. (2004). Functional anatomy of the attentional modulation of time estimation. *Science, 303*, 1506–1508.

Criscuolo, A., Schwartze, M., Henry, M. J., Obermeier, C., & Kotz, S. A. (2023). Individual neurophysiological signatures of spontaneous rhythm processing. *NeuroImage, 273*, 120090.

Dalla Bella, S., Benoit, C.-E., Farrugia, N., Schwartze, M., & Kotz, S. A. (2015). Effects of musically cued gait training in Parkinson's disease: Beyond a motor benefit. *Annals of the New York Academy of Sciences, 1337*, 77–85.

Etchell, A. E., Johnson, B. W., & Sowman, P. F. (2014). Behavioral and multimodal neuroimaging evidence for a deficit in braintiming networks in stuttering: A hypothesis and theory. *Frontiers in Human Neuroscience, 8*, 467.

Falk, S., Müller, T., & Dalla Bella, S. (2015). Non-verbal sensorimotor timing deficits in children and adolescents who stutter. *Frontiers in Psychology, 6*, 847.

Friston, K. (2009). The free-energy principle: A rough guide to the brain? *Trends in Cognitive Sciences, 13*, 293–301.

Friston, K. (2012). Prediction, perception and agency. *International Journal of Psychophysiology, 83*, 248–252.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology, Paris, 100*, 70–87.

Fujii, N., & Graybiel, A. M. (2003). Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science, 301*, 1246–1249.

Graybiel, A. M. (1997). The basal ganglia and cognitive pattern generators. *Schizophrenia Bull, 23*, 459–469.

Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiology of Learning and Memory, 70*, 119–136.

Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience (Palo Alto, CA), 31*, 359–387.

Guenther, F. H., & Hickok, G. (2016). Neural models of motor speech control. In G. Hickok & S. L. Small (Eds.), *Neurobiology of language* (pp. 725–740). Academic Press.

Handel, S. (1974). Perceiving melodic and rhythmic auditory patterns. *Journal of Experimental Psychology, 103*, 922–933.

Ivry, R. B., & Schlerf, J. E. (2008). Dedicated and intrinsic models of time perception. *Trends in Cognitive Sciences, 12*, 273–280.

Jones, M. R. (1976). Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review, 83*, 323–355.

Kotz, S. A., & Schwartze, M. (2011). Differential input of the supplementary motor area to a dedicated temporal processing network: Functional and clinical implications. *Frontiers in Integrative Neuroscience, 5*, 86.

Kotz, S. A., Anwander, A., Axer, H., & Knösche, T. R. (2013). Beyond cytoarchitectonics: The internal and external connectivity structure of the caudate nucleus. *PLoS One, 8*, e70141.

Large, E. W., & Jones, M. R. (1999). The dynamics of attending: How people track time-varying events. *Psychological Review, 106*, 119–159.

Lehéricy, S., et al. (2004). 3-D diffusion tensor axonal tracking shows distinct SMA and pre-SMA projections to the human striatum. *Cerebral Cortex, 14*, 1302–1309.

Macar, F., Coull, J., & Vidal, F. (2006). The supplementary motor area in motor and perceptual time processing: fMRI studies. *Cognitive Processing, 7*, 89–94.

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: Coincidence detection of oscillatory processes. *Cognitive Brain Research, 21*, 139–170.

Matell, M. S., Meck, W. H., & Nicolelis, M. A. (2003). Interval timing and the encoding of signal duration by ensembles of cortical and striatal neurons. *Behavioral Neuroscience, 117*, 760–773.

McDonald, V., & Halliday, G. M. (2002). Selective loss of pyramidal neurons in the presupplementary motor cortex in Parkinson's disease. *Movement Disorders, 17*, 1166–1173.

Merchant, H., Harrington, D. L., & Meck, W. H. (2013). Neural basis of the perception and estimation of time.

Annual Review of Neuroscience (Palo Alto, CA), 36, 313–336.

Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: A neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B, 370*, 20140093.

Moore, J. W., Ruge, D., Wenke, D., Rothwell, J., & Haggard, P. (2010). Disrupting the experience of control in the human brain: Pre-supplementary motor area contributes to the sense of agency. *Proceedings of the Royal Society B, 277*, 2503–2509.

Petter, E. A., Lusk, N. A., Hesslow, G., & Meck, W. H. (2016). Interactive roles of the cerebellum and striatum in sub-second and supra-second timing: Support for an initiation, continuation, adjustment, and termination (ICAT) model of temporal processing. *Neuroscience and Biobehavioral Reviews, 71*, 739–755.

Penfield, W. (1950). The supplementary motor area in the cerebral cortex of man. *Arch Psychiatr Nervenkr Z Gesamte Neurol Psychiatr, 185*, 670–674.

Penhune, V. B., & Zatorre, R. J. (2019). Rhythm and time in the premotor cortex. *PLoS Biology, 17*, e3000293.

Picard, N., & Strick, P. L. (2001). Imaging the premotor areas. *Current Opinion in Neurobiology, 11*, 663–672.

Pont-Sunyer, C., Hotter, A., Gaig, C., et al. (2015). The onset of nonmotor symptoms in Parkinson's disease (the ONSET PD study). *Movement Disorders, 30*, 229–237.

Pöppel, E. (1978). Time perception. In R. Held, W. H. Leibowitz, & H. Teuber (Eds.), *Handbook of sensory physiology* (pp. 713–729). Springer.

Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Current Opinion in Neurobiology, 14*, 139–147.

Schwartze, M., & Kotz, S. A. (2013). A dual-pathway neural architecture for specific temporal prediction. *Neuroscience and Biobehavioral Reviews, 37*, 2587–2596.

Schwartze, M., & Kotz, S. A. (2016). Regional interplay for temporal processing in Parkinson's disease: Possibilities and challenges. *Frontiers in Neurology, 6*, 270.

Schwartze, M., Tavano, A., Schröger, E., & Kotz, S. A. (2012a). Temporal aspect of prediction in audition: Cortical and subcortical neural mechanisms. *International Journal of Psychophysiology, 83*, 200–207.

Schwartze, M., Rothermich, K., & Kotz, S. A. (2012b). Functional dissociation of pre-SMA and SMA-proper in temporal processing. *NeuroImage, 60*, 290–298.

Schwartze, M., Brown, R. M., Biau, E., & Kotz, S. A. (2020). Timing the "magical number seven": Presentation rate and regularity affect verbal working memory performance. *International Journal of Psychology, 55*, 342–346.

Sherman, S. M. (2001). A wake-up call from the thalamus. *Nature Neuroscience, 4*, 344–346.

Smith, K. S., & Graybiel, A. M. (2016). Habit formation. *Dialogues in Clinical Neuroscience, 18*(1), 33–43.

Spencer, R. M. C., & Ivry, R. B. (2013). Cerebellum and timing. In M. Manto, D. L. Gruol, J. D. Schmahmann, N. Koibuchi, & F. Rossi (Eds.), *Handbook of the cerebellum and cerebellar disorders* (pp. 1201–1219). Springer.

Weingarten, C. P., Sundman, M. H., Hickey, P., & Chen, N. (2015). Neuroimaging of Parkinson's disease: Expanding views. *Neuroscience and Biobehavioral Reviews, 59*, 16–52.

Wiener, M., Turkeltaub, P., & Coslett, H. B. (2010). The image of time: A voxel-wise meta-analysis. *NeuroImage, 49*, 1728–1740.

Index

H. Merchant, V. de Lafuente (eds.), *Neurobiology of Interval Timing*, Advances in Experimental Medicine and Biology 1455, https://doi.org/10.1007/978-3-031-60183-5