Auditory Cortical Ensemble-Mechanisms Facilitating Auditory-Driven Behaviors and Perception

by

Harini Suri

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Psychology)
in the University of Michigan
2024

Doctoral Committee:

Assistant Professor Gideon Rothschild, Chair
Associate Professor Omar Ahmed
Assistant Professor Ada Eban-Rothschild
Assistant Professor Michael Roberts

Harini Suri

harinis@umich.edu

ORCID iD: 0000-0003-0580-5249

© Harini Suri 2024

Dedication

To my mother

And

To friends who are family

Acknowledgements

I would like to start by thanking my advisor, Dr. Gideon Rothschild, for his unwavering faith in my work ethics and hard work from the very beginning. He has been there to celebrate the small victories like getting my first ever neuronal images under the two-photon microscope and the big ones like my first peer-reviewed publication. He has trusted me to be independent in the lab, design and run my experiments and allowed me the freedom to find my way through all the numerous runs of data analysis. I am grateful to Gideon for giving me this opportunity and all his encouragement throughout these 5+ years.

My dissertation committee has probably been a dream come true in many ways. Dr. Omar Ahmed and Dr. Ada Eban-Rothschild have taken the time to guide me through techniques and analysis even before I officially formed my dissertation committee. Dr. Michael Roberts has been a truly positive presence and so supportive in all my committee meetings and has been the first person at the end of every meeting to ask me what he can do to help me achieve the goals of my dissertation.

I started my PhD by boldly declaring that I do not want to invest too much time or effort in teaching. This completely changed when I serendipitously got assigned to teach with Dr. Jen Cummings. She gave me the time to find my footing, and then grow as a graduate student instructor. She always listened to my ideas, let me assist with her courses, and believed in me to keep up even when we had to switch from in-person to virtual classes within a span of a week when the pandemic hit. Working with Jen for over

two years certainly bolstered my ability to confidently deliver lectures and learn how to interact and mentor undergraduate students. Besides all of this, she has taken the time to talk me through ups and downs of graduate school, for which I am extremely grateful.

I started in the Rothschild lab along with (almost Dr.) Carlos Vivaldo. Carlos has been there to share all the good and the bad times over the years, and probably remembers as much as I do about all the fancy equipment (read tape) we used to fix and get our rigs to work.

Half of my work in this dissertation and much more beyond would never have been conceived, much less worked, had it not been for Dr. Karla Salgado-Puga. Karla kept me on my toes and made sure I thought about my experimental designs and research goals. She was generous with her time to help me explore projects and was happy to collaborate in on a part of the second study of this dissertation, which was conceived during a casual conversation. I have been lucky to have Karla by my side and her expertise is something I am still learning from.

I have been fortunate to work with several motivated and keen undergraduate research assistants during my PhD. My very first undergraduate students, Viviana Quezada and Haiqi Guo, gave me my very first mentoring experience and helped me learn the process. Nathanial Nass started working with me when we had just resumed working full-time after the pandemic and ensured that I got the much-needed pilot data that determined the course of the second study of this dissertation. I am grateful to Nayomie Allen for being such a reliable individual, working with me over 2 years, learning and contributing significantly to the data collection, and for going the extra mile to ease my responsibilities many a times. Kaitlynn Lane, Kyra Granroth, Bianca Wang and Alberto

Olivei demonstrated exceptional competence and a strong commitment to learning new techniques and got me through the demanding data collection process. Thanks to Lily Allen, Troy Swodzinski, Andrew Robson, Zeinab Mansi, Lindsay Cain, Kevin Steffer, and Grant Griesman for consistently guaranteeing the timely completion of experiments. I am honored to have known and worked with all of you and I am particularly grateful for your valuable contributions to this dissertation.

Thanks to Mekhala Kumar for her remarkable ability to keep up with my impromptu conversation starters and directions, and for her willingness to join me on spontaneous outings, which has continued from our undergraduate years. To quote her wise words “I am extremely grateful to you for having borne my incessant questions and for having been a daily source of entertainment”.

I would also like to thank rest of my 2018 cohort, Dr. Megha Ghosh, soon to be Drs. Eryn Donovan, Chris Turner, Mena Davidson, Averill Cantwell and Melissa Painter, for cheering along and being a substantial part of this journey. A special shoutout to Dr. Caitlin Posillico for guiding me through all my teaching-related missteps, and for all our late-night hallway conversations. Thanks to Dr. Cassie Avila, Ileana Morales, Chelsea Markunas, and Alika Sulaman for your support and generosity. A special thank you to all the people I have met and befriended through F.E.M.M.E.S. and GRIN, especially my book brunch folks, who kept me going through the pandemic and lockdown times.

I probably would have never set foot on this journey to a PhD had it not been for my family and friends. I come from a family with one too many doctoral degree holders, making this probably the most obvious course of action in one way. I am grateful to my parents for their love, encouragement, and belief in me, even when everything I do has

not always made sense to them both. I had many reasons for joining this graduate program and a major one was my best friend, Poortata. She has been the only one to reason with me through my numerous circuitous thought processes related to science and life otherwise and has had my back even when I did not realize. I am thankful to Siddharth for being an encouraging, supportive, and entertaining presence in my life and for all his help with debugging my Python-related bugs. My heartfelt thanks to Harsha and Shubham for making this PhD a fulfilling experience even from afar. I finally made it here, you all!

Table of Contents

List of Figures

Abstract

Sounds in our everyday environment play a crucial role in guiding our perception and behaviors. The ability to effectively process these sounds—perceive them, contextualize their meaning, and subsequently harness this information to adapt our behaviors—is an intricate yet fundamental process. Consider, for instance, the familiar experience of recognizing a friend's voice in a bustling crowd, or the moment when a sudden car horn prompts a swift response to ensure safety. These everyday scenarios underscore the significance of auditory perception, which is known to rely on the auditory cortex within the auditory system. The auditory cortex, a higher-level brain region in the auditory pathway, performs a multifaceted role in processing sounds, beyond the mere analysis of their acoustic features. It is essential for processing spectrotemporally rich sounds like human speech or animal vocalizations, which carry ethological relevance, and contributes to our ability to engage in sound-guided behavior and decision-making. However, how the auditory cortex brings together all the moving parts in our acoustic environment to facilitate a stable auditory perception across contextual variations and time, is still an enigma. This dissertation tackles this challenge by examining neural mechanisms in the auditory cortex at two distinct temporal scales, and its functioning under baseline conditions and in behavioral contexts, providing comprehensive insights into the functioning of the auditory cortex in real-world contexts.

The first study in this dissertation addresses the long-term stability of auditory cortical sound representations, comparing the processing of complex sounds like animal vocalizations, with that of simple sounds like pure tones. By recording the sound-evoked neural responses in the auditory cortex using two-photon calcium imaging, this study provides evidence for the distinction in longitudinal sound representations in the auditory cortex based on the acoustic structure and salience of the auditory inputs. The second study moves to investigate auditory cortical mechanisms in a behavioral context. In this study, I adapted the classical appetitive trace conditioning paradigm to train mice in predicting the time to reward, using a sound cue. By combining electrophysiology, chemogenetic and pharmacological interventions, this study establishes the causal and functional role of auditory cortex and its downstream connection to the posterior striatum in sound-triggered interval timekeeping, at a 1-second temporal resolution. Collectively, these studies offer insights about how neural representations in the auditory cortex can simultaneously encode for auditory and relevant non-auditory information like timing, which are necessary for shaping our consequent actions and behaviors. This work makes an essential contribution to the literature on the various auditory cortical mechanisms aiding in auditory perception but also underscores the importance of recognizing the auditory cortex as a region with broader functions beyond primary auditory processing.

Chapter 1 : Introduction

Processing the array of sounds that envelop our world – from twittering of birds to meaningful conversations and the lively music of concerts – requires more than just hearing. It involves recognizing where these sounds come from, understanding their meaning, and reacting appropriately. For instance, being able to recognize familiar voices in a crowded room makes social interactions smoother or the sound of a car honking warns us of our surroundings and ensures our safety. Without such a stable and reliable perception of sounds in our daily environment, our lives would be quite challenging. Impaired ability to process speech and other complex sounds has been observed in several patients with schizophrenia and autism and have been further linked with irregular cognitive function (F. R. Lin et al., 2013; L. Liu et al., 2016; Peelle & Wingfield, 2016; Uchida et al., 2019). Moreover, noise-induced, and age-related hearing loss are considered global public health issues especially due to their detrimental effect on cognition in over 20% of the global population (Bisogno et al., 2021; F. R. Lin et al., 2013; Natarajan et al., 2023; Uchida et al., 2019). Therefore, this ability, known as auditory perception, is a fundamental part of how we experience the world around us, shaping our interactions and helping us make sense of what we hear.

The neural mechanisms of the auditory system aid in achieving robust auditory perception by extracting enormous amounts of information from all the different types of sounds we encounter in our daily lives. Sound representations in the early stations of the

auditory pathway like the cochlear nucleus, medial geniculate body and the inferior colliculus have been well-described. It is known that these regions are highly specialized in detecting spectro-temporal features of the sound like intensity, pitch, and frequency (Casseday et al., 2002; Davis, 2005; Marsh et al., 2006; Rhode & Smith, 1986; Rouiller et al., 1983), aid in sound localization (Shackleton et al., 2003), and put together topographic maps of physical attributes of the incoming sounds (Davis et al., 2003; Eggermont, 2001; Ehret & Schreiner, 2005; King & Moore, 1991; Ress & Chandrasekaran, 2013; Stiebler & Ehret, 1985). Nonetheless, a disparity exists between these early sound representations, primarily based on the sound waveform's physical characteristics, and perceptual representations, which are intricately tied to real-world entities often referred to as "auditory objects" (Bizley & Cohen, 2013; Griffiths & Warren, 2004; Nelken et al., 2003, 2014; Nelken & Bar-Yosef, 2009; Sutter & Shamma, 2011). Representation of these “auditory objects” is believed to be achieved through the diverse characteristics of the neural activity in the auditory cortex (Nelken, 2004; Sutter & Shamma, 2011; Winkler et al., 2009). However, there are many aspects of how the auditory cortex facilitates auditory perception that are yet to be explored. This dissertation will address neural mechanisms in the auditory cortex that facilitate auditory perception through two distinct studies – one aimed at understanding how real-world complex sounds like speech are represented in the auditory cortex across time and the other, investigating the auditory cortical neural mechanisms underlying a sound-triggered time prediction behavior, essential for our daily activities.

1.1 Auditory cortex

The auditory cortex (AC) is a key region involved in higher-level auditory processing and integration, necessary for formation of “auditory objects” (Bizley & Cohen, 2013; Griffiths & Warren, 2004; Nelken et al., 2014; Nelken & Bar-Yosef, 2009). Most studies in the auditory system have focused on comprehending how simple stimuli like pure tones, clicks and noise bursts, which are the fundamental elements of any sound, are represented neuronally (For example, Read et al., 2002; Nelken et al., 2004; Bizley et al., 2005). Individual neurons in all regions of the auditory pathway are typically most sensitive to one component frequency, known as the characteristic or the best frequency, in the incoming sound signal. In AC, pure tones evoke robust frequency-specific neuronal responses (Hromádka et al., 2008). Many studies in bats, cats and nonhuman primates have shown that individual AC neurons respond to multiple frequencies which are often harmonically related (Kadia & Wang, 2003; Sadagopan & Wang, 2009; Suga et al., 1983; Sutter & Schreiner, 1991; X. Wang, 2013). These findings suggest that AC neurons are a necessary component of circuits involved in processing harmonic patterns in speech or sounds from musical instruments.

Studies using simple sounds have also given us an essential understanding of the local and tonotopic sound representations in AC. For instance, rats trained on an appetitive classical conditioning paradigm with a pure tone as the conditioning stimulus, changed the frequency selectivity of individual AC neurons to that of the conditioning stimulus frequency (Bakin & Weinberger, 1990) and enlarged the representation of the conditioning sound frequency in the global tonotopic map of AC (Polley et al., 2006; Recanzone et al., 1993). Some studies used repeated stimulation of auditory cortical

neurons with pure tones over days in different stages of cortical development to modulate the auditory cortical tonotopic map and understand the extent of its experience-dependent plasticity (Froemke et al., 2007; Recanzone et al., 1993; Zhang et al., 2001; Zhou et al., 2011). To characterize sound-evoked AC neuronal firing patterns, AC neurons in marmoset monkeys were tonically driven with a preferred sound stimulus that had a defined frequency, modulation frequency and intensity, and were compared against AC firing patterns to a non-preferred tonic stimulation (X. Wang et al., 2005). It was seen that preferred stimuli evoke a sustained firing pattern over the stimulus duration in AC, compared to AC responses to the non-preferred stimulus, emphasizing the selective representation of sounds in AC (X. Wang et al., 2005). All these studies elucidate a lot of characteristic properties of AC neuronal activity in response to sounds but are limited in expanding upon the role of AC in processing complex sounds such as speech and animal vocalizations, which carry ethological relevance.

1.2 Complex sound processing in AC

Early studies in AC were shaped by stimulating this region with sound elements ubiquitous in our daily lives like human speech and animal sounds, especially in presence of background noise, famously known as the “cocktail party problem”. Amongst these numerous studies, one particular study provided some of the earliest evidence that neurons in the AC encode for sounds according to their meaning rather than just their structural elements (Wollberg & Newman, 1972). In this study, individual neuronal activity in the AC of squirrel monkeys was recorded in response to modified recordings of species-specific vocalizations and it was shown that, in addition to producing discrete neural responses to each individual vocalization, AC neuronal response to the

subsequent vocalization interacted with that of the previous vocalization, suggesting that AC encodes for the temporal patterns of the vocalizations and analyses complex auditory signals (Wollberg & Newman, 1972).

Following these early studies that provided the foundation for studying AC responses to complex sounds, researchers asked the fundamental question of how AC is involved in this process. Over the decades, both causal (Butler et al., 1957; Goldberg & Neff, 1961; Ohl et al., 1999) and physiological evidence for this has been found in AC. Mongolian gerbils trained to discriminate between two complex sounds showed an impairment in this discrimination ability following bilateral lesioning of AC, compared to another group of gerbils which were able to retain their ability to discriminate between two pure tones after bilateral ablation of AC (Ohl). Once the causal role of AC was established in complex sound processing, extensive research has been done to show that neural representations of complex sounds in AC are unique and are not a linear combination of its responses to the component frequencies that make up the complex sounds (Atencio et al., 2012; Kim et al., 2020; Sadagopan & Wang, 2009; Town et al., 2018). For instance, it was found that 26% of AC neurons in awake marmoset monkeys were not responsive to pure tones (Sadagopan & Wang, 2008) but were highly selective for complex sound features (Sadagopan & Wang, 2009). Using two tone pips as stimuli, it was seen that these non-tone responsive neurons exhibited a nonlinear combination sensitivity to the spectral and temporal characteristics of the sounds rather than the individual tone pips themselves (Sadagopan & Wang, 2009).

A recent study mapped the neural discriminability to communication sounds in guinea pigs along the various regions of the auditory pathway including the cochlear

nucleus, inferior colliculus, auditory thalamus and the primary auditory cortex (Souffi et al., 2020). Compared to the earlier stations of the auditory pathway, AC neural ensembles showed a reduced discrimination ability of vocalizations (Souffi et al., 2020), emphasizing the previous finding that neural responses in AC encode for information beyond the features of the incoming sounds (Chechik & Nelken, 2012). Importantly, this study also showed that representations in AC were invariant to the background noise during communication sounds. Invariance of AC sound representations to sound level, pitch, timbre and spatial location has been demonstrated by studies in songbirds (Bilimoria et al, 2008), rodents (Carruthers et al., 2015; Klein et al., 2006; Sadagopan & Wang, 2008; Schneider & Woolley, 2013) and ferrets (Bizley et al., 2009, 2013). It has also been shown that some AC neurons selectively respond to the background sounds, allowing for sound source segregation (Bar-Yosef & Nelken, 2007). Such invariance of sound representations has been the key to what is known as “perceptual constancy”, where one is able to identify sound features despite variation in sensory input. This perceptual ability was addressed by neural recordings in ferret AC, which were proficient in identifying the sound identity across variations in acoustic features of the sound (Town et al., 2018). AC responses encoded the sound identity, while also representing task-irrelevant sound features, influencing the animals’ behavioral performance accuracy. These invariant sound representations in AC are key to our ability to identify the same sound across variations in our environment.

1.2.1 In behavioral contexts

. A fundamental role of AC is to process sounds to drive behavior. Natural sounds that animals encounter in their environment, are particularly important in developing

experiences over animals' lifetime and consequently shaping their behaviors. Ultrasonic vocalizations emitted by rodents in different contexts have been shown to be crucial for various behaviors. For example, mouse pup calls are selectively recognized and behaviorally relevant to mothers and not to non-naïve virgin female mice (Ehret, 1987; Ehret & Haack, 1982; Ehret & Koch, 1989). Single unit and LFP recordings in the AC of mothers show a pup call-evoked sustained inhibition compared to AC responses seen in virgin females, suggesting that auditory cortical plasticity allows for enhanced detection of pup calls in mothers (Galindo-Leon et al., 2009; F. G. Lin et al., 2013). On the other hand, rats produce vocalizations at different frequencies in appetitive and aversive situations and attract social interactions with juvenile rats only in appetitive conditions (Brudzynski, 2013). There are numerous examples of stereotyped vocalizations produced by songbirds to attract mates and warn others of dangers (Doupe & Kuhl, 1999). Recent studies showed that the ability to identify complex sounds from musical instruments depends on the subject's ability to distinguish the rise and fall in sound intensity and time, which were encoded in AC neural responses in humans (Cutting & Rosner, 1974) and in mice (Deneux et al., 2016). Such a vast array of literature implicates how AC is important for detecting and processing frequency modulations and patterns of structural harmonics that are an essential part of vocalizations in animals and human speech (Kuchibhotla & Bathellier, 2018).

While these findings emphasize how AC is important for shaping behavior using sound cues, how AC is involved in auditory-driven behaviors is perplexing. Multiple studies have addressed this by distinctly suggesting that while AC is required for perceptual learning of auditory-instructed behaviors, it is not required for processing other

auditory aspects of the task. For instance, inactivation of AC in ferrets did not impair their ability to localize sounds but influenced the accuracy of sound localization following a period of monoaural sensory deprivation (Bajo et al., 2019). It has been established by numerous studies that AC is not required for discrimination between two simple sounds like pure tones (Ohl et al., 1999). However, through focal optogenetic perturbation of AC activity, a recent study showed that while AC is dispensable for discrimination of dissimilar tones in mice, it is necessary for context-dependent sound discrimination requiring temporal integration (Ceballo et al., 2019). Further, many studies have characterized the role of AC in sound-guided decision-making tasks in which the sound representation in AC indicates the choice made by the subjects (Aizenberg & Geffen, 2013; Bathellier et al., 2012; Fritz et al., 2003, 2005; Kuchibhotla & Bathellier, 2018). All these findings suggest that plasticity of AC responses is crucial for learning and facilitating adaptation following any hearing-related problems (Bajo et al., 2019) and more importantly, highlight how the functional role of AC is highly dependent on the behavioral context and could potentially extend beyond its role in processing sound-related information.

1.2.2 Study 1: How are complex real-life sounds represented in AC across time?

Evidence from all these studies strongly supports the unique and significant role of AC in processing behaviorally relevant sounds, which are spectrotemporally complex. To support reliable auditory perception that is consistent across time, the ability to recognize complex sounds and their associated meaning is expected to rely on their stable neural representations across time. At the same time, it is known that AC response properties undergo rapid and robust experience-dependent functional plasticity (Atiani et al., 2009; David et al., 2012; Fritz et al., 2003; Polley et al., 2013). Therefore, a key

question that remains to be addressed is whether AC forms a stable representation of complex sounds across days. Hence, I hypothesized that auditory cortical representations of complex sounds are more stable across days than those of sounds that lack spectrotemporal structure such as pure tones. Testing this hypothesis requires recording the responses of the same AC neurons across days, which poses a technical challenge. Previous studies have limited their neural recordings to short periods of time (minutes to hours) or have recorded different forms of population-mean signals across days (Galván et al., 2001; Kisley & Gerstein, 2001), which could be distinct from findings at single neuron-level. One study recorded multiunit activity of AC receptive fields in ferrets and found that 73% of the receptive exhibited stability over the course of minutes to hours (Elhilali et al., 2007) but acquiring reliable recording from the same neurons across days using electrophysiology is highly challenging. In *vivo* two-photon calcium imaging offers precise spatial localization of individual neurons and allows for recording the activity of the same neurons across days. Therefore, to quantify the degree of longitudinal stability of auditory cortical representations of complex sounds and pure tones, in *Study 1,* I carried out two-photon calcium imaging of identified excitatory neuronal ensembles in layers 2/3 of AC in awake head-fixed mice across days.

1.3 Predicting auditory-driven future events is essential for auditory perception

Predictive coding theory postulates that AC's ability to predict and explain sensory experiences extends beyond mere sound processing. It entails the iterative refinement of predictions and their comparison with incoming sounds at varying levels of the auditory pathway to reduce prediction errors and to construct a stable neural representation that underlies auditory perception (Kumar et al., 2011). Most of our initial evidence for

predictive coding in AC is a product of human studies using mismatch negativity (MMN) tasks and animal studies examining stimulus specific adaptation (SSA). SSA is considered the “single-cell analog of MMN” (Khouri & Nelken, 2015), showing enhanced responses to surprise or deviant sounds in a repeated sequence of same sounds (Heilbron & Chait, 2018; Ulanovsky et al., 2003). Along with encoding for the unfamiliar sound in a sequence, the reduced representation of the repeated stimuli during SSA, shows the influence of stimulus history on AC activity (Rubin et al., 2016; Ulanovsky et al., 2004). It has also been shown that an omitted sound at the end of a repetitive sequence of sound stimulation evokes an expectation-based response in AC, further driving home the modulation of AC neural activity by stimulus history. Apart from sounds, actions predictive of upcoming salient sound stimuli are known to weaken AC responses to the sound, especially when the sounds are self-generated (Reznik et al., 2021; Rummell et al., 2016). These retrospective representations of prediction based on previous sensory or motor experiences in AC make a strong case for AC in predicting upcoming future events and guide adaptive behavior.

1.3.1 Study 2: How does the auditory corticostriatal pathway support sound-triggered reward time prediction?

A crucial behavioral sequence that we perform daily is to use sounds to determine our next steps. An example of one such behavioral sequence is to use sound cues to estimate and predict time intervals to the upcoming actions effectively. For example, the distinct sound of a phone notification prompts us whether we need to answer it immediately or later, while the sound of an approaching vehicle signals when it is safe to cross the street. However, the neural mechanisms that underlie this ability to use sounds

to reliably estimate seconds-long time intervals to future salient events, and use these estimates for guiding appropriate action, are not well understood. In the auditory pathway, a key candidate brain region for encoding sound-triggered timing is AC, due to its established role in behavior- and decision-making- dependent sound processing (Bathellier et al., 2012; Francis et al., 2018; Fritz et al., 2003, 2005; King & Schnupp, 2007; Kuchibhotla & Bathellier, 2018; Nelken et al., 2014; Town et al., 2018). Numerous studies have demonstrated retrospective coding of the degree to which a sound deviates from expectation in the auditory cortex (Audette et al., 2022; Huang et al., 2023; J. Li et al., 2017; Singer et al., 2018; Ulanovsky et al., 2003), as outlined above. However, much less is known about the existence of prospective coding of anticipated time from a sound to a subsequent event. Two studies in AC addressed aspects of such prospective coding from a sound stimulus to a following sound stimulus which predicts reward (Jaramillo & Zador, 2011) or from the sound stimulus generated from an action such as a lever press to initiate reward delivery (Cook et al., 2022). But it remains unknown whether and how the auditory cortex is involved in sound-triggered predictive timing of future salient events on the timescale of seconds. Using a combination of behavioral monitoring, electrophysiological recording, pharmacological and chemogenetic inactivation techniques, I investigated the neural mechanisms underlying sound-triggered prediction of time to consequent reward, in *Study 2*.

Together, the findings from both these studies advance our understanding of two distinct neurophysiological mechanisms underlying auditory driven behaviors, and further underscore how the role of auditory cortex is unique and multifaceted, to achieve reliable auditory perception. Note that results of study 1 have already been published (Suri &

Rothschild, 2022) and findings from study 2 is in preparation to be submitted as a manuscript (Suri et al, 2023).

Chapter 2 : Enhanced Stability of Complex Sound Representations Relative to Simple Sounds in the Auditory Cortex

2.1 Abstract

Typical everyday sounds, such as those of speech or running water, are spectrotemporally complex. The ability to recognize complex sounds (CxS) and their associated meaning is presumed to rely on their stable neural representations across time. The auditory cortex is critical for processing of CxS, yet little is known of the degree of stability of auditory cortical representations of CxS across days. Previous studies have shown that the auditory cortex represents CxS identity with a substantial degree of invariance to basic sound attributes such as frequency. We therefore hypothesized that auditory cortical representations of CxS are more stable across days than those of sounds that lack spectrotemporal structure such as pure tones (PTs). To test this hypothesis, we recorded responses of identified L2/3 auditory cortical excitatory neurons to both PTs and CxS across days using two-photon calcium imaging in awake mice. Auditory cortical neurons showed significant daily changes of responses to both types of sounds, yet responses to CxS exhibited significantly lower rates of daily change than those of PTs. Furthermore, daily changes in response profiles to PTs tended to be more stimulus-specific, reflecting changes in sound selectivity, as compared to changes of CxS responses. Lastly, the enhanced stability of responses to CxS was evident across longer time intervals as well. Together, these results suggest that spectrotemporally CxS are

more stably represented in the auditory cortex across time than PTs. These findings support the role of the auditory cortex in representing CxS identity across time.

2.2 Introduction

Everyday sounds such human speech, animal vocalizations, the sound of running water or rustling of leaves, are spectrotemporally complex (Doupe & Kuhl, 1999; Ehret & Haack, 1982; Gygi et al., 2007). A key brain region involved in the perception of spectrotemporally complex sounds is the auditory cortex (AC) (Bizley et al., 2009; Griffiths et al., 2004; King et al., 2018b; Maor et al., 2020; Nelken, 2004, 2008; Nelken & Bar-Yosef, 2009; Rauschecker, 1998). For example, AC lesions result in a more profound impairment in processing CxS in comparison to PTs and other simple sounds in both humans (Griffiths, 2003; Kaga et al., 1997) and animal models (Harrington et al., 2001; Ohl et al., 1999; Rybalko et al., 2006). Responses of AC neurons to CxS can often not be predicted from a linear combination of responses to the PT components of the CxS (Angeloni & Geffen, 2018; Atencio et al., 2008; Barbour & Wang, 2003; Harper et al., 2016; Mizrahi et al., 2014; Nelken et al., 1999; Sadagopan & Wang, 2009; Schreiner et al., 2011; Schwartz et al., 2020; X. Wang et al., 2005). Furthermore, studies using a range of approaches have shown that AC responses to CxS represent sound “identity” with a substantial invariance to its frequency components and other acoustic parameters (Blackwell et al., 2016; Carruthers et al., 2015; Chechik & Nelken, 2012; Harpaz et al., 2021; Nelken et al., 2003, 2014; Town et al., 2018). While these studies suggest an important role of the AC in representing the identity and meaning of CxS, to what degree these representations are stable across time remains unknown.

To support the ability to recognize sensory stimuli and their associated meaning, the neural representations of the stimuli are expected to be stable across time (Lütcke et al., 2013; Schoonover et al., 2021). At the large-scale spatial resolution, the representation of tone frequency across the AC tonotopic map is indeed generally stable in adulthood in the absence of instructive learning or manipulation of the acoustic environment (W. Guo et al., 2012; Merzenich et al., 1976). At the single-cell level, receptive fields of most auditory cortical neurons have been found to be stable across up to 2 hours of recording, though a minority of neurons exhibited significant changes within this timeframe (Elhilali et al., 2007a). However, whether AC sound representations are stable across days and whether the representations of CxS and PTs are similarly stable, remains unknown. Given the suggested involvement of AC in representing CxS identity, we hypothesized that CxS would be more stably represented in the AC across time as compared to PTs. Here, we tested this hypothesis by recording the responses of identified L2/3 AC excitatory neurons to both PTs and CxS across days in awake mice using two-photon calcium imaging.

2.3 Materials and Methods

All animal procedures were performed in accordance with the University of Michigan animal care committee's regulations.

2.3.1 Animals

We used 13 (10 males, 3 females, 8-15 weeks old) Thy1-GCaMP6f mice (C57BL/6J-Tg (Thy1-GCaMP6f) GP5.17Dkim/J, Jax number: 025393) which express the GCaMP6f calcium indicator in excitatory pyramidal neurons (Dana et al., 2014). Mice

were housed under a reverse 12h light/12h dark cycle, with lights on at 8:30pm and off at 8:30am. Experiments were conducted between 11am and 4pm and each animal was imaged around the same time of day across all days of data collection so that the time gap between consecutive imaging days was ~24 hours.

2.3.2 Surgical procedure

All surgeries were performed on mice anesthetized using ketamine (100mg/kg, i.p.) and xylazine (10mg/kg, i.p.). Anesthetized mice were placed in a stereotaxic frame (Kopf 514 Instruments, CA, USA), and an anti-inflammatory drug (Carprofen, 5mg/kg, subcutaneous injection) and a local anesthetic (lidocaine, subcutaneous injection) were administered. A craniotomy was performed over the right primary AC (AP: -3.1 mm; ML: 4.6 mm lateral from midline (Supplementary Figure 2.1) using a 3mm biopsy punch (Integra Inc) and a 3mm diameter round glass cranial window was secured over this craniotomy. A custom-made lightweight (<1 gr) titanium head bar was attached to the left side of the skull using dental cement and cyanoacrylate glue to allow for head-fixed imaging. During the surgery, body temperature was maintained at 38^{o}C and the depth of anesthesia was regularly assessed by checking pinch withdrawal reflex. Mice were treated with Carprofen for 48 hours post surgically and allowed to recover for a week.

2.3.3 Two-photon calcium imaging

Mice were first habituated to the imaging setup and the sound protocols for 3 days. During the 3-day habituation period, the animals were exposed to the same PT and CxS stimuli as during imaging days 1-5 while being head-fixed in the same setup under the

two-photon microscope while being positioned on a circular treadmill (without imaging). Each stimulus was presented 30-35 times in total across the 3-day habituation period.

During imaging, the objective of the microscope was placed perpendicular to the surface of the cranial window to access the AC. Imaging was carried out using an Ultima IV two-photon microscope (Bruker) through water immersion objectives (Nikon, 40X, NA = 0.65 (n= 2 mice) and a 16X, NA = 0.8 (n = 8 mice)) and a pulsed laser was used to provide excitation at 940nm (MaiTai eHP DeepSee by Spectra Physics). Data was collected using galvanometric ("galvo") scanning of 256X256 pixel images at 3 frames/second. We conducted a separate set of recordings from the same neurons using galvo scanning and faster resonant scanning at 60 frames/second (averaging every 4 frames to yield 15 frames/sec) and found that responsiveness, response magnitude and trial-to-trial consistency were not underestimated by the slower galvo imaging sample rate (Supplementary Figure 2.2). During the period of habituation, focal planes with a high yield of neurons were determined in L2/3 (imaged at depths of 150-330μm (Meng et al., 2017)). The overlying blood vessel patterns and position with respect to the cortical surface were noted for these identified focal planes and were used to image the same focal planes across five consecutive days of the experiment.

2.3.4 Auditory stimuli

Stimuli were generated at a sampling rate of 97.6 kHz using MATLAB and presented to the animal using an SA1 speaker amplifier, ED1 speaker driver and a multi-field magnetic speaker (MF1) positioned ~10cm in front of the animal, all by Tucker-Davis Technologies (TDT). Acoustic stimuli consisted of two protocols: PTs consisted of 8 pure

tone stimuli at 2-32 kHz (Figure 2.1D), while CxS consisted of 8 sounds, including 4 animal vocalizations (cricket, macaque, chiffchaff, water shrew) and 4 environmental sounds (glass, thump, scratch and water, Figure 2.1D). The CxS had significantly higher frequency bandwidth, spectral entropy, and spectrotemporal modulation as compared to the PTs (Figure 2.1C). The duration of each sound was 500 ms (padded with silence for some of the CxS) and sound intensity was 65-70 dB SPL. In a given imaging session for each focal plane, each sound within a protocol was repeated 10 times in a pseudorandom order with an inter-stimulus interval of 1.5 ± 0.3 seconds. The order of the sound protocols was shuffled across experiments.

Frequency bandwidth, spectral entropy, and spectrotemporal modulation were quantified for all sounds as attributes of sound complexity. Occupied frequency bandwidth quantifies the range of frequencies a sound is composed of and was calculated as the difference in frequency between the points where the integrated power crosses 0.5% and 99.5% of the total power in the spectrum. Spectral entropy of a sound quantifies how distributed its frequency content is and was calculated as the Shannon entropy of the normalized power distribution of the sound. Spectrogram autocorrelation of each sound measures the similarity of the frequency content of a sound across time bins and was calculated by temporally binning each spectrogram into 20 equally sized time bins (excluding brief periods of silence at the end of some sounds), resulting in column vectors that represent the power distribution of the sound at every time bin. We then calculated the Pearson correlations between all vectors and averaged these correlations values. Thus, the spectrogram autocorrelation of each sound inversely represents the degree of

spectro-temporal modulation. The ‘Spectrotemporal modulation index’ was defined as 1-spectrogram autocorrelation.

2.3.5 Data analysis

2.3.5.1 Preprocessing

Imaging data was run through the open-source Suite2p software package (Pachitariu et al., 2016) to correct for movement and neuropil signal, and to select neuronal regions of interest (ROI). To ensure reliable physiological measurements, we required that in any given imaging session, detected cell bodies show a compactness > 0.8 and that their ΔF/F trace shows a skewness > 1.1 and clear transients (the experimenter was blind to sound responsiveness during the cell inclusion phase). A small minority of responses occurring during locomotion were excluded from all analyses (Supplementary Figure 2.3). All further analysis was done on the data preprocessed and output from Suite2p using custom-written MATLAB scripts (MathWorks, 2019a).

To identify the same neurons across imaging sessions, the average across-frames fluorescence image (with Suite2p’s median-filtering image enhancement) of each focal plane was used. The average fluorescence images of the same focal plane were then manually matched for same neurons across days. We confirmed cell matching using fully automated image registration (MATLAB command: imregcorr) and calculation of structural similarity index (MATLAB command: ssim) of the cell bodies across days and found >95% agreement (Supplementary Figure 2.4).

2.3.5.2 Two-photon imaging data analysis

The relative change in fluorescence (ΔF/F) was defined for each neuron in a given imaging session as $(F(t) - F_0)/F_0$, where F(t) is the raw fluorescence signal of the cell at time t, and F_0 is the median of the raw fluorescence signal across the session. The response magnitude of a given neuron to a sound was defined as the across-trials average ΔF/F within 0-1.5 s from sound onset. Responsiveness of a given neuron to each stimulus was determined using a bootstrap analysis. Specifically, the difference between the sound response magnitude across trials and the mean pre-stim response magnitude (mean ΔF/F during in the pre-stim windows (-1.5 to 0 s) of all sounds in the protocol) was compared to a distribution of similar differences resulting from 1000 random shuffles of the sound responses and pre-stim responses. The neuron was considered responsive to a given stimulus if the difference between the real sound response and mean pre-stim magnitude was larger than 97.5% of the shuffled differences and if the sound response magnitude was at least 10% greater than the pre-stim magnitude. On a given day, a neuron was considered sound responsive if it was responsive to at least one stimulus on that day (with Bonferroni correction for the number of stimuli).

To allow pooling changes in daily responses across neurons with different response magnitudes, the responses of each neuron to all stimuli across the two days of comparison were z-scored prior to further analysis and statistical testing. For each comparison, a neuronal response to a given stimulus was included if the neuron was sound responsive on at least one of the days of comparison.

The significance of a change in response magnitude of a given neuron to a specific sound was quantified using a shuffle test. Specifically, the difference in mean response magnitudes between days was determined to be significant if the difference was larger than 95% of the simulated differences generated from the random shuffling of trials across the days of comparison (nShuffles = 1000) and in addition the magnitude of change was at least 10%. Using this method, we computed the significance of changes across 4 one-day intervals (Day 1 → Day 2; Day 2 → Day 3; Day 3 → Day 4; Day 4 → Day 5), 3 two-day intervals (Day 1 → Day 3; Day 2 → 4; Day 3 → Day 5), 2 three-day intervals (Day 1 → 4; Day 2 → Day 5) and 1 four-day interval (Day 1 → Day5).

The percentage of significant change in daily neuronal responses to a stimulus was calculated by:

$$\frac{(\text{Number of sound responses that showed a significant change}) * 100}{\text{Total number of significant responses}}$$

A neuron was determined to show significant change across days if it showed a significant change in response to at least one stimulus (after Bonferroni correction for the number of stimuli). The fraction of neurons across all pairs of consecutive days showing a significant change was calculated by:

$$\frac{\text{Number of neurons that showed a significant change}}{\text{Total number of sound responsive neurons}}$$

For a given neuron, we computed the average Euclidean Distance between its response profiles (magnitude of responses across stimuli) across pairs of days using the following equation:

$$\sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \ldots + (x_n - y_n)^2}$$

where x_i equals the neuron's response to stimulus i on the first day and y_i equals the neuron's response to stimulus i on the second day.

To test for stimulus-specificity of response change, we tested whether the day of recording (1 or 2) significantly interacted with the stimulus identity in determining response magnitude using a two-way ANOVA with interaction. The ANOVA output was used to compute the effect size (ω^2) of the interaction term.

To test whether there is a significant difference between multi-day or multi-stimuli proportions across CxS and PTs (Figure 2.1F and Figure 2.4), we used a bootstrap analysis. Specifically, for each category across CxS and PTs (e.g., "1-day" in Figure 2.4A), we derived a distribution of 10,000 randomly simulated PTs proportions given the probability of the corresponding CxS category. The p-value was calculated as the fraction of "CxS-simulated" PTs probabilities that were equal to or higher than the real PTs probabilities across categories.

2.3.5.3 Statistical tests

We used statistical tests at a $p<0.05$ significance level and $\alpha = 0.05$ for all comparisons unless otherwise indicated (Table1).

2.4 Results

To quantify the degree of stability of auditory cortical representations of PTs and CxS, we carried out two-photon calcium imaging of identified excitatory neuronal

ensembles in L2/3 of the AC (Supplementary Figure 2.1) in 10 awake head-fixed Thy1-GCaMP6f mice across days. As the degree of sound novelty influences response magnitude in AC (Heilbron & Chait, 2018; Kato et al., 2015; Nelken, 2014; Parras et al., 2017; Ulanovsky et al., 2003, 2004), we familiarized the mice to the experimental sound protocols for three consecutive days while being head-fixed under the two-photon microscope before data acquisition commenced (Figure 2.1A). During this habituation period, in each animal, three optical focal planes were chosen and registered with respect to the overlying blood vessel pattern to allow for repeated imaging of the same neurons across days (Figure 2.1B).

From day 1 to day 5 of the experiment, we imaged the daily responses of the same neuronal ensembles to 8 PTs of varying frequencies and 8 CxS. The CxS consisted of animal vocalizations and environmental sounds that broadly overlapped in frequency content with the PTs, while having significantly higher frequency bandwidth, spectral entropy, and spectrotemporal modulation (Figure 2.1C,D, see Methods). As expected, AC neurons responded to both PTs and CxS with sound-triggered transients in relative change in fluorescence (ΔF/F, Figure 2.1E). We first compared the degree of overall sound-evoked responsiveness to CxS and PTs across the population. We found that response magnitudes to PTs and CxS were not significantly different (Figure 2.1F) and that the rate of responsive neurons to PTs and CxS were also not significantly different (Figure 2.1G). Thus, our chosen set of PTs and CxS evoked similar magnitudes and rates of responses among L2/3 AC excitatory neurons. Responsiveness, response magnitude and trial-to-trial consistency were not underestimated by our imaging sample rate (Supplementary Figure 2.2).

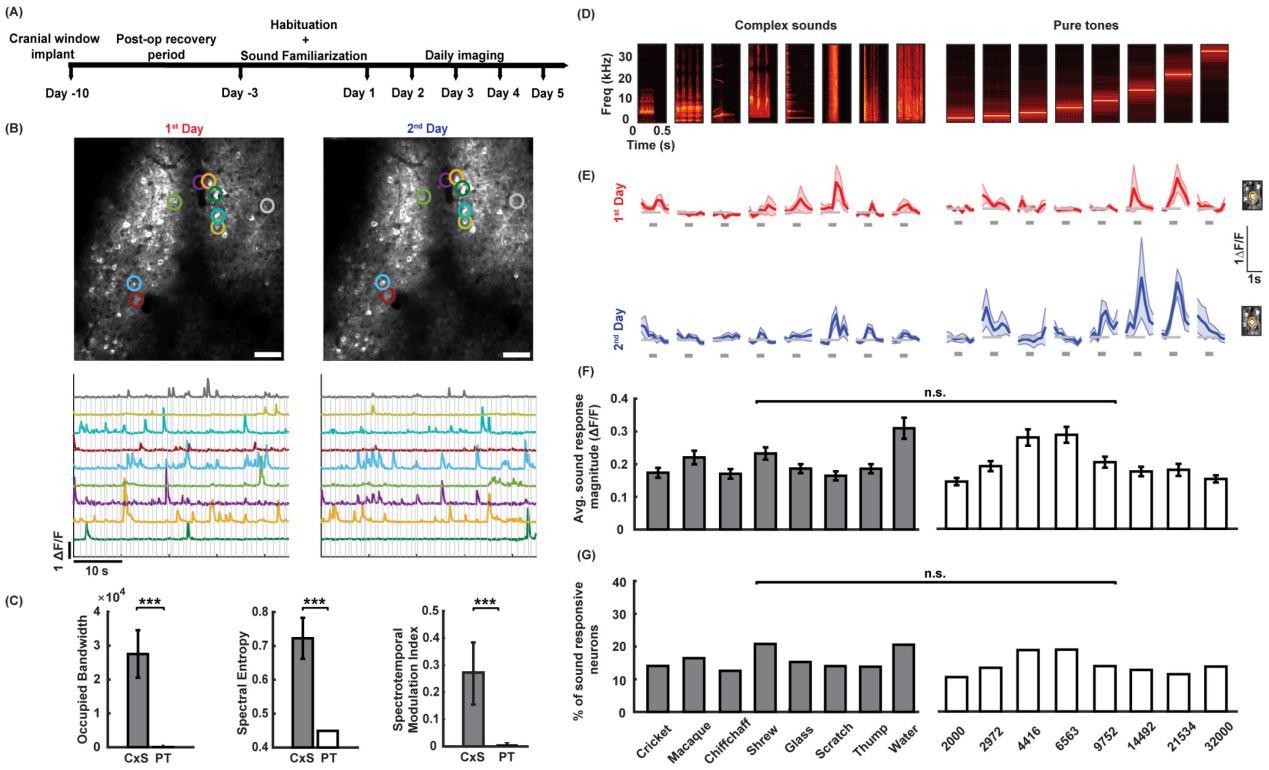


Figure 2.1: Imaging responses of identified L2/3 auditory cortical excitatory neurons to pure tones and complex sounds across days (A) An illustration of the experimental timeline. (B) Top: Example two-photon micrographs from L2/3 of the auditory cortex in an awake mouse from two consecutive days. Scale bar = 60 μm. Colored circles identify example individual neurons matched across two consecutive days. Extended Data Figure 1.1 shows the histological verification of the imaging location. See Extended Data Figure 1.4 for validation of neuron matching across days. Bottom: ΔF/F traces of the neurons marked in the micrographs during a pure tone protocol. Gray lines indicate the stimuli. Scale bar: 10 seconds, 1ΔF/F (C) Comparison of spectrotemporal features of the CxS and PT. From left to right: occupied bandwidth (**p = 0.00015, Mann-Whitney U test), spectral entropy (***p = 0.00015 Mann-Whitney U test), and spectrotemporal modulation index (***p = 0.00015, Mann-Whitney U test). (D)Spectrograms of the sound stimuli presented in the complex sounds (Left) and pure tone (Right) protocols. The identity of the complex sounds and the frequency of the pure tones are labeled at the bottom of panel 'G' (E) Responses of a representative neuron to complex sounds and pure tones (the stimuli correspond to panel 'D' column-wise) across two consecutive days (first day in red and second day in blue). Shaded area marks the mean ± SEM across trials. The gray bar below each response indicates the stimulus time (0.5s). Scale bar: 1s, 1 ΔF/F. The cell body of the neuron as imaged in the two days is highlighted on the right. Supplementary Figure 2.2 validates that responses included in the study were not underestimated due to sampling rate and Supplementary Figure 2.3 shows the distribution of the number of trials included in the data. (F) Average sound response magnitude (mean ΔF/F across all trials over the stimulus window) across all days of imaging in response to complex sounds and pure tones (the stimuli correspond to panels 'D'-'G' column-wise). Error bars indicate mean ± SEM. Number of neurons across all pairs of consecutive days (with repetitions): CXS: 557, PTs: 587, p = 0.98 (Two-way ANOVA). (G) Percentage of sound responsive neurons recorded across all days of imaging in response to complex sounds and pure tones (the stimuli correspond to panels 'D'-'G' column-wise). Number of neurons across all pairs of consecutive days (with repetitions): CXS: 557, PTs: 587, p = 0.875 (bootstrap test, see Methods, Table 1). Sound stimuli for each are indicated below: 8 complex sounds and 8 pure tone frequencies (Hz).

We next quantified the degree of stability of these neuronal responses by comparing responses of identified neurons across pairs of consecutive days. The identical variation in daily experimental and physiological conditions for PTs and CxS allowed us to compare the relative degrees of change in responses between the two sound protocols. We observed that while most responses of individual AC neurons showed stability across days, some displayed significant daily variation (Figure 2.2A). To measure changes in sound responses we first focused on responses of individual neurons to individual stimuli across pairs of consecutive days and restricted our analyses to responses that were significant in at least one of the two days. Across this population, we found that while the majority of responses were stable across days, 22% (114/518) of significant responses to PTs showed a significant change in response magnitude across successive days (Figure 2.2B). These results suggest that underlying a generally stable representation, responses of AC neurons to PTs show a moderate degree of daily dynamics. Interestingly, however, only 12.15% (66/543) of significant responses to CxS showed a significant change in magnitude across the same time interval (Figure 2.2B). This proportion of daily response change to CxS was significantly lower than that of PTs (Figure 2.2B), suggesting that AC responses to CxS are more stable than to PTs across days. The degree of stability of CxS with well-defined spectral centroids at <10 KHz (Cricket, Chiffchaff and Macaque) did not significantly differ from those of more distributed spectra (Glass, Shrew, Thump, Scratch and Water) (12.36% vs. 12.05%, respectively, $p = 0.92$, Chi square test for proportions).

As a complementary approach, we quantified a similar measure at the single-neuron rather than single-stimulus level. To this end, we calculated the fraction of sound-

responsive neurons that exhibited a significant change in response magnitude to at least one of the eight PTs or CxS for each pair of consecutive days. Consistent with our findings at the single-stimulus level, we found that the fraction of neurons showing a significant change in response to CxS was significantly lower than to PTs (Figure 2.2C).

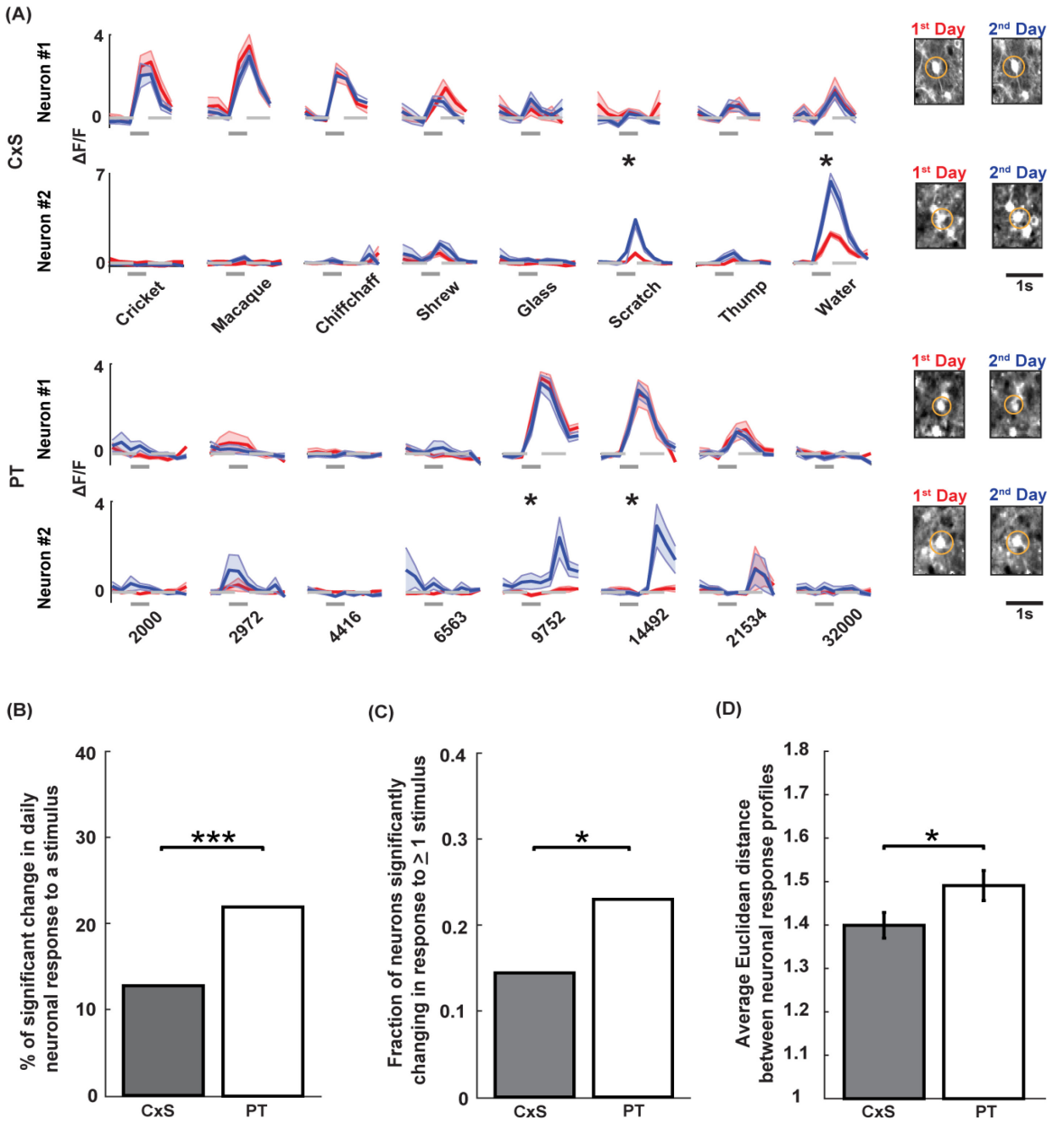


Figure 2.2: Auditory cortical responses to complex sounds are more stable than to pure tones across days (A) Responses of two representative neurons to complex sounds (Rows 1 and 2) and pure tones (Rows 3 and 4) across two consecutive days (first day in red and second day in blue). Shaded area marks the mean ± SEM across trials. Gray bars indicate the stimulus time. Stars indicate significant response changes (see Methods). Scale bar: 1s. The cell bodies of the imaged neurons are shown on the right. Rows 1 and 3 show neuronal responses that were stable from one day to the next and rows 2 and 4 show neuronal responses that changed from one day to the next. (B) Percentage of significant changes in response to complex sounds and pure tones across pairs of consecutive days. CXS: 12.15% (66/543), PTs: 22.01% (114/518),***p = 1.91×10^{-5} (Chi-square test). (C) Fraction of neurons that show a significant change in response to at least one stimulus across pairs of consecutive days. CXS: 0.14 (34/241), PTs: 0.23 (59/256), *p = 0.011 (Chi-square test). (D) Average Euclidean distance between the response profile of a neuron (to either CXS or PTs) from one day to the next day, *p = 0.046 (two-sided t-test). Extended Data Figure 2.1 shows the relationship between a neuron's changes in responsiveness to CxS and PT.

To quantify stability/plasticity of sound responses at the level of response profiles across stimuli, we computed for each neuron the Euclidean distance between its response profile (to either PTs or CxS) on one day and that of the next day. A larger Euclidean distance reflected a higher degree of response change across stimuli. Consistent with the findings above, we found that the Euclidean distance between daily response profiles to PTs was significantly higher than to CxS (Figure 2.2D). There was no significant correlation between the Euclidean Distance of the same neurons to PTs and CxS (Supplementary Figure 2.5A) and changes in responses to CxS were not significantly more strongly correlated with changes in frequency-overlapping PT as compared to frequency-non-overlapping PT (Supplementary Figure 2.5B). Together, these findings across varying quantification methods indicate that AC neuronal responses to CxS are more stable than to PTs across consecutive days.

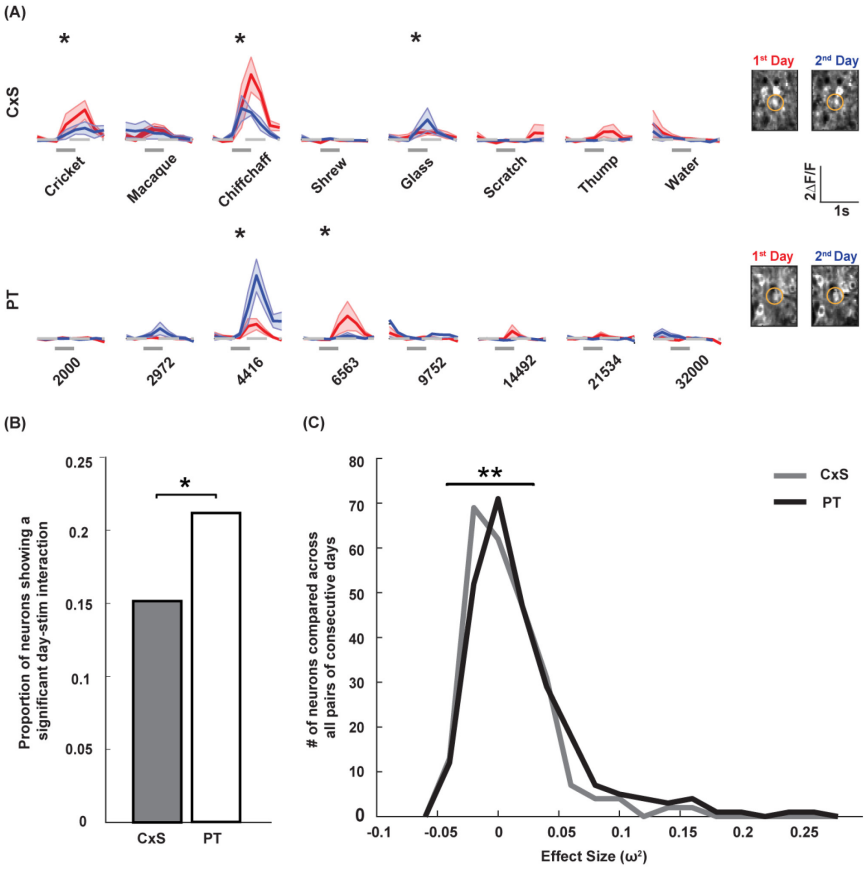


Figure 2.3: Daily plasticity in responses to pure tones is more stimulus-specific than to complex sounds. (A)Responses of two representative neurons to CxS (Row 1) and PTs (Row 2) across two consecutive days, showing stimulus-specific changes. Shaded area marks the mean ± SEM across trials. Gray bars indicate stimulus timing. Stars indicate significant response changes. Scale bar: 1s, 2 ΔF/F. The cell bodies of the imaged neurons are shown on the right. (B)Proportion of neurons across all pairs of consecutive days showing a significant day/stimulus interaction, computed via two-way ANOVA. CXS: 14.9% (36/241) neurons, PTs = 21.1% (54/256), *p = 0.037 (Z-test for proportions) (C) Distributions of the effect size (ω^2) indicating the strength of the interaction between day and stimulus for CXS (gray) and PTs (black) for each neuron across all consecutive days. **p =0.0039 (Mann-Whitney U Test)

A change in the response profile of a neuron across days may include a change in response gain, manifesting as similar changes in response magnitude across stimuli, or it may be stimulus-specific, reflecting a change in the neuronal sound selectivity (Figure 2.3A). To test whether changes in responses to PTs and CxS differed in the nature of change, we compared the degree of stimulus-specificity of response change for each of the stimuli classes. We tested for each neuron's responses, whether there was a significant interaction between the day of recording and the different stimuli. A significant day/stimulus interaction indicates that responses to the different stimuli were differentially modulated across days, reflecting stimulus-specificity in response change. We found that a significantly higher proportion of neurons showed stimulus specificity in daily changes in responsiveness to PTs as compared to CxS (Figure 2.3B). Further, the strength of the day/stimulus interaction was significantly higher for PTs than CxS (Figure 2.3C). These findings indicate that in addition to showing higher overall rates of daily change in responsiveness, the changes in responses to PTs were more stimulus specific, and therefore reflected a higher degree of change in sound selectivity, in comparison to CxS.

Finally, we investigated how the rates of change across pairs of days relate to rates of change across longer durations. To this end, we quantified the changes in responsiveness in a similar manner across intervals of 1-4 days. We found that the degree of response plasticity increased with increasing time interval between days for both CxS

and PTs (Figure 2.4A, B). Moreover, the elevated rates of change in responses to PTs as compared to CxS that were observed across pairs of days also manifested across these intervals (Figure 2.4A). The fraction of neurons showing a significant change to at least one stimulus showed a similar trend, though did not reach significance (Figure 2.4B). Lastly, the Euclidean distance between the PTs response profiles were significantly higher than those of CxS across these intervals (Figure 2.4C). Consistent with our previous results, this suggests that AC representations of CxS are more stable compared to PTs over a range of daily time intervals.

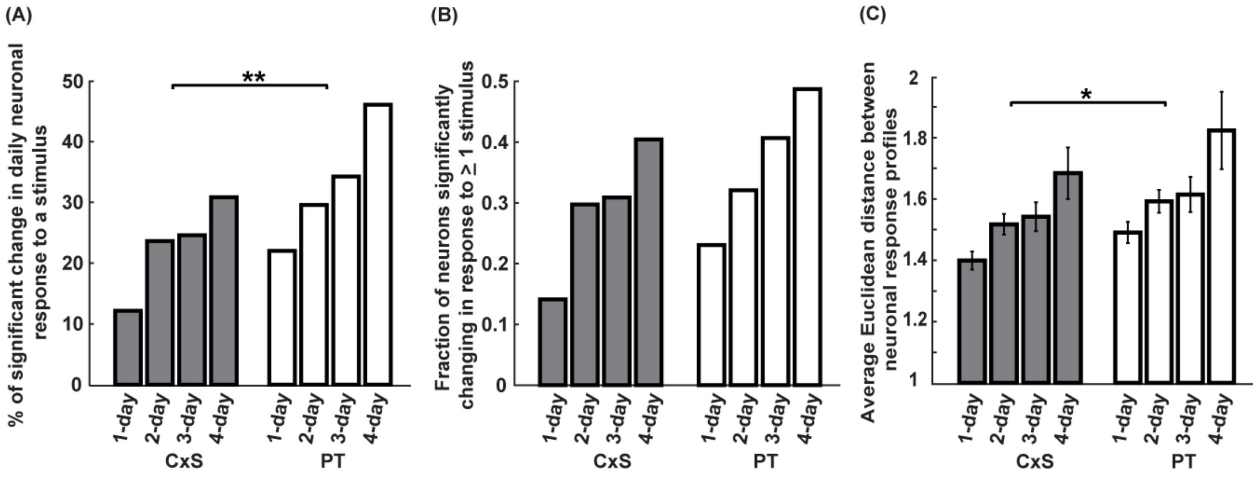


Figure 2.4: Auditory cortical responses to complex sounds are more stable than to pure tones across multiple days. (A) Percentage of significant change in daily neuronal response to a given stimulus in the CxS (gray bars) and PTs (white bars) protocol across varying daily intervals. **p = 0.0015 (bootstrap test, see Methods). (B)Fraction of neurons significantly changing in response to at least one stimulus in the CxS (gray bars) and PTs protocol (white bars) across varying daily intervals. p = 0.1084 (bootstrap test, see Methods). (C) Average Euclidean distance between responses of a neuron for CXS (gray bars) and PTs (white bars) across varying daily intervals. *p = 0.015 (Two-way ANOVA).

2.5 Discussion

In this study, we used two-photon calcium imaging to record the degree of stability and plasticity of sound-evoked responses of L2/3 AC excitatory neurons to PTs and CxS across days. We found that most responses to both PTs and CxS were stable, with a

moderate but significant degree of change across pairs of consecutive days. Importantly, we report that responses to CxS exhibited significantly enhanced stability across days as compared to PTs. Furthermore, the structure of response profiles to PTs exhibited larger degrees of change than to CxS across days, as evidenced by a higher degree of stimulus-specific changes. Finally, we found that the enhanced degree of stability in CxS representations generalizes to longer daily time intervals.

Our findings of a significant degree of ongoing daily changes in auditory cortical representations of both CxS and PTs add to a number of recent studies describing “representational drift” in other sensory modalities (Deitch et al., 2021; Pérez-Ortega et al., 2021; Peron et al., 2015; Ranson, 2017; Rule et al., 2019; Schoonover et al., 2021). Together, these studies point to a potential common principle, by which despite the well-established link between perception and cortical function (Bergman, 1990; Ceballo et al., 2019; Chait et al., 2010; Chapuis & Wilson, 2012; Frégnac & Bathellier, 2015; Kuchibhotla & Bathellier, 2018; Lee & Rothschild, 2021; Leopold, 2012), a stable sensory perception does not rely on fixed cortical sensory representations. Instead, representational dynamics may reflect a general principle of cortical function. Indeed, the locally heterogeneous organization of AC L2/3 ensembles has been suggested to be well suited to support rapid synaptic reorganization in response to changing environmental conditions (Bandyopadhyay et al., 2010; Bathellier et al., 2012; Francis et al., 2018; Kanold et al., 2014; Kato et al., 2015; J. Liu et al., 2019; J. Liu & Kanold, 2021; Maor et al., 2016; Rothschild et al., 2013, 2010; Rothschild & Mizrahi, 2015). Whether sound representations in the thalamorecipient L4 are more stable than those in L2/3 remains for future studies.

While auditory cortical representations of both classes of sounds exhibited significant degrees of daily change, representations of CxS were significantly more stable as compared to PTs across varying quantification methods. These findings likely result from the differences in the acoustic properties of these stimuli. In particular, CxS are decomposed into narrow frequency channels at the cochlea and reconstructing their wideband frequency contents throughout the auditory pathway requires re-integration across frequency channels. In contrast, a pure tone evokes responses in a narrower channel throughout the auditory system. If daily variation in responses is at least partly independent in different frequency channels, integration across frequency bands as needed to represent CxS may “average out” some of this variation as compared to that of a PTs. Thus, spectrotemporal integration may give rise to enhanced longitudinal stability of CxS in the AC. Future studies could directly test this possibility by, for example, measuring the stability of representations of noise with systematically varying bandwidth. An alternative acoustic property that may determine the degree of AC stability is based on temporal rather than spectral integration. In particular, temporal modulations in the complex sounds may "reset" neuronal responses multiple times within a stimulus, such that the enhanced degree of overall stability is due to temporal averaging of per-modulation fluctuations. This possibility could be tested using sequences of amplitude-modulated tones, which have temporal modulation without spectral bandwidth.

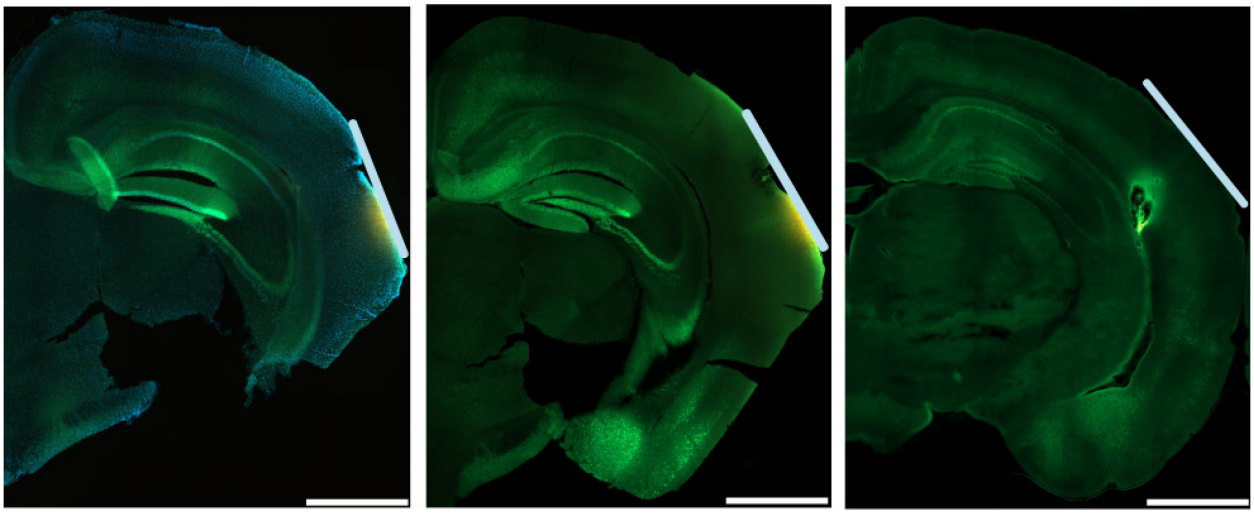
Beyond the higher degrees of change in responses to PTs as compared to CxS, we also found that PT response changes were more stimulus-specific than those of CxS. These findings suggest that changes in responses to CxS tended to be shaped more by global gain factors while changes in responses to PTs tended to reflect stimulus-tuning

changes to a larger degree. If changes to CxS are correlated with changes to the tones that make up the CxS, this finding may be influenced by the frequency overlap between CxS, which is not the case for PTs. Although our finding that responses to CxS do not significantly change as to their frequency-overlapping tones (Supplementary Figure 2.5) argues against this possibility, the experiments described above could directly test it.

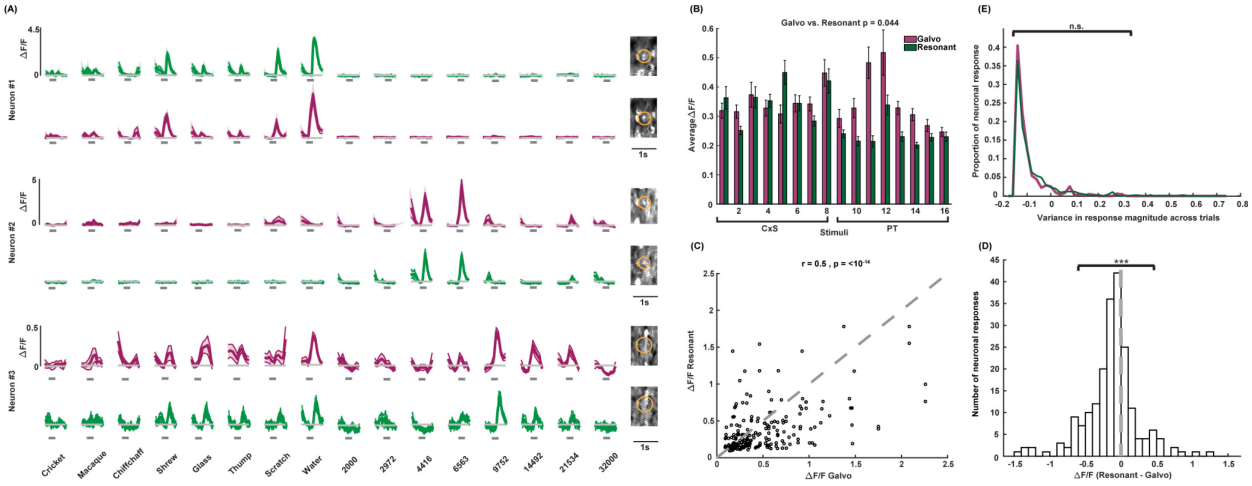
Beyond the acoustic differences between CxS and PTs, a combination of evolution and previous experience may also have contributed to enhanced stability of AC representations of CxS as compared to PTs. Future studies may test this hypothesis by comparing the degree of AC stability to sounds with similar spectrotemporal complexity but varying ethological relevance.

Our findings raise the question of whether enhanced stability of AC representations of CxS are linked with enhanced perceptual stability of these sounds. As the AC is important for sound perception in both humans (Griffiths, 2003; Kaga et al., 1997) and animal models (Ceballo et al., 2019; Frégnac & Bathellier, 2015; Harrington et al., 2001; Kuchibhotla & Bathellier, 2018; Ohl et al., 1999; Rybalko et al., 2006), it is tempting to speculate based on our findings that behavioral measures of perceptual stability, such as sound recognition across days, would be higher for CxS as compared to PTs. Testing this speculation may have important implications as PTs are not just widely used in auditory research but are also the standard in studies using classical conditioning and other learning paradigms.

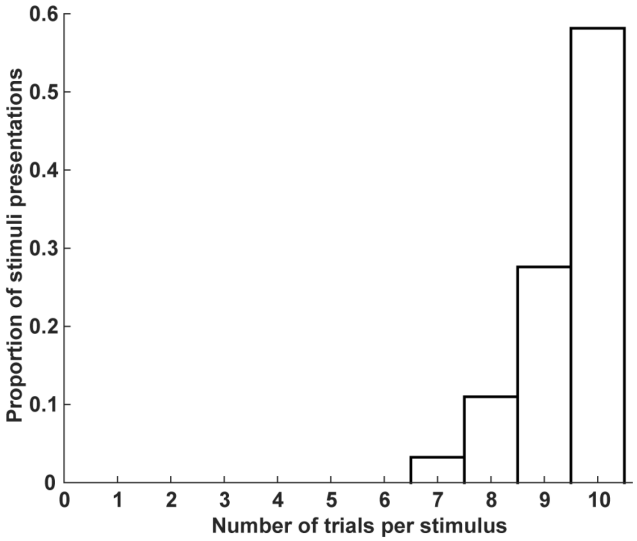
2.6 Supplementary Figures and Table



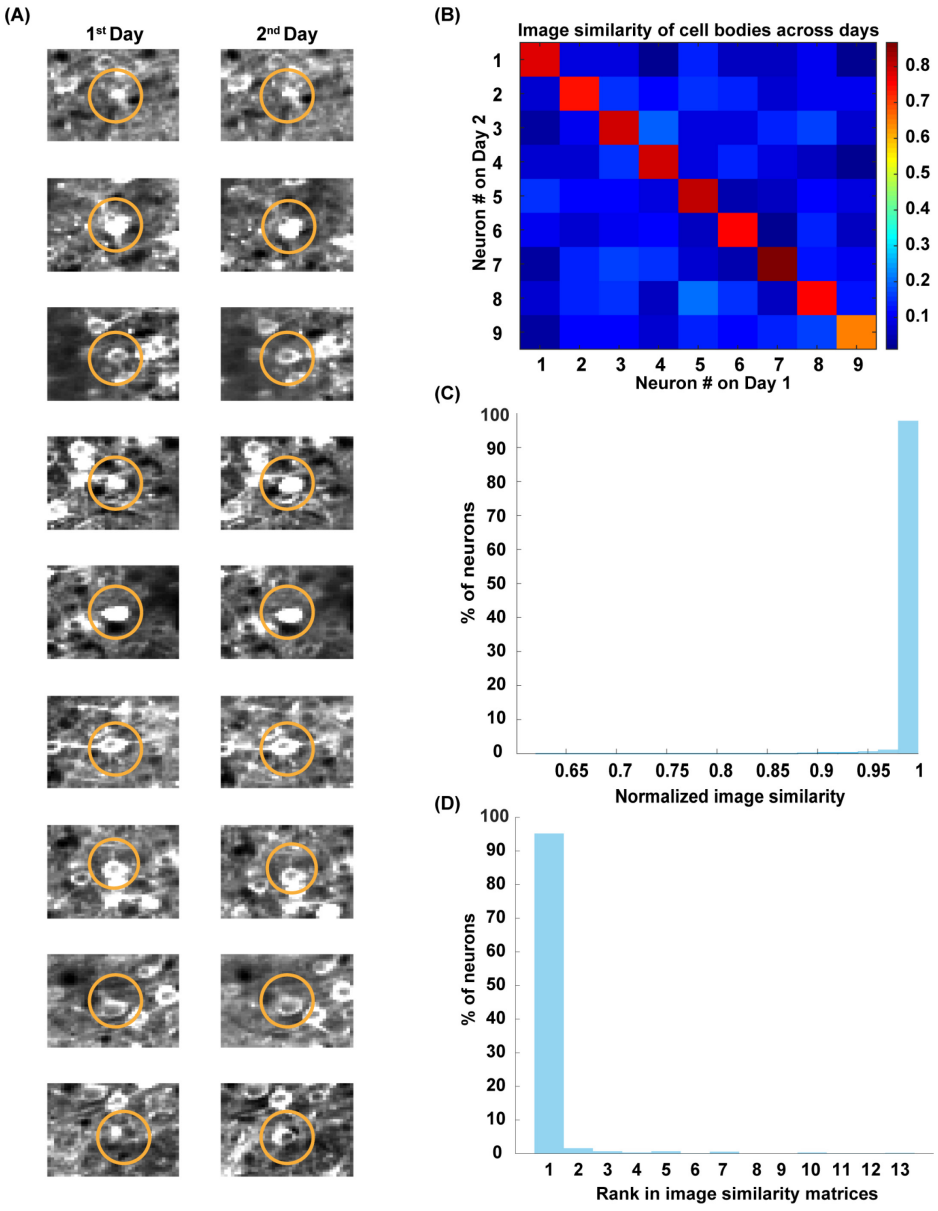
Supplementary Figure 2.1 Histological verification of imaging location. Representative images of coronal brain sections from 3 animals used in this study. Following the completion of the experiments, mice were killed with an overdose of xylazine (10mg/kg, i.p.) and the imaging cranial window was removed. Using a nanoFil needle (Hamilton), we injected 1µl of DiI Tracer (D282 Invitrogen) into the site of imaging identified by blood vessel patterns and covered the brain surface for 5 minutes to maximize labeling and prevent fluorescence loss caused by the perfusion during tissue fixation with 4% PFA. The extracted brains were kept in PFA for 3 days and then transferred to 30% sucrose solution for another 3-4 day before cryosectioning. The brains were sliced in 50µm thickness sections and preserved with Fluoroshield mounting medium with DAPI (Abcam). Recording site confirmation was done by imaging tissue sections positive for GCaMP, DAPI, and DiI fluorescence. DiI fluorescence trace from brain sections were cross-referenced with the Allen Mouse Common Coordinate Framework using the NeuroInfo software (MBF Bioscience). The location of the cranial window is indicated on each brain slice and the specific site of imaging is marked by DiI in yellow. Scale bar: 1000µm.



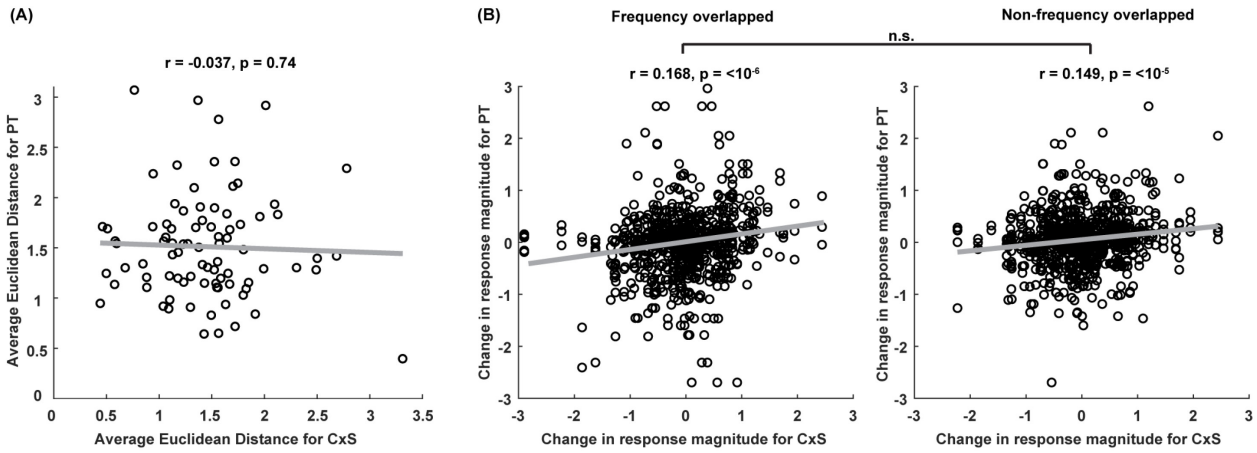
Supplementary Figure 2.2 Comparison of auditory cortical responses to CxS and PT in galvo and resonant scanning modes. (A) Responses of three representative neurons in galvo (in maroon) and in resonant (in green) scanning modes to CxS and PT stimuli (corresponding stimuli are indicated at the bottom of the panel). Shaded area marks the mean ± SEM across trials. The gray bar below each response indicates the stimulus time (0.5s). Scale bar: 1s. The cell body of the neuron as imaged in the two scanning modes is highlighted on the right. (B)Average sound response magnitude (mean ΔF/F across all trials over the stimulus window) across both imaging modes in response to complex sounds (Stimuli #s 1-8) and pure tones (Stimuli #s 9-16). Error bars indicate mean ± SEM. Each focal plane was imaged in galvo and resonant scanning modes alternatingly twice and all pairs of consecutive imaging sessions were included in the comparison between galvo and resonant scanning modes. Number of neurons across all pairs of scanning modes (with repetitions): Galvo: 576, Resonant: 540, p = 0.044 (Two-way ANOVA). In a separate analysis, we found that the likelihood of a neuron to be responsive to a specific CxS and a specific PT using resonant scanning was 7.74% (41/530) and 7.24% (38/525), respectively. Using galvo scanning these values were slightly higher, at 9.23% (53/574) and 10.44% (55/527), respectively, suggesting that significant responses were not underestimated by the use of galvo scanning. (C) Correlation of individual responsive neurons' response magnitude (mean ΔF/F across all trials over the stimulus window) to individual stimuli (both CxS and PT included) in galvo and resonant scanning modes. Neurons were matched across scanning modes using an automated MATLAB algorithm (https://github.com/ransona/ROIMatchPub) and then validated by visual inspection. Dashed grey line represents the diagonal. Correlation coefficient = 0.5, $p = 2.36 \times 10^{-14}$ (Pearson's correlation). (D) Distribution of the difference in response magnitude between galvo and resonant scanning modes across neurons and stimuli (corresponding to the difference between the X and Y values of the points in 'C). $p = 1.6 \times 10^{-6}$ (Wilcoxon Signed Rank test). (E) Distributions of the trial-by-trial variance in response magnitude to given stimuli in galvo (maroon) and in resonant (green) scanning modes. p = 0.09 (Mann Whitney U test).



Supplementary Figure 2.3 Distribution of number of trials included per stimulus across the dataset following trial exclusion due to locomotion.



Supplementary Figure 2.4. Validation of neuron matching across days using image similarity analysis. (A) Cell bodies of neurons matched across a pair of consecutive days. Square dimensions = 39x39 pixels. (B)An example image similarity matrix corresponding to the cell bodies shown in A from a single focal plane, depicting the similarity for each neuron on day 1 compared against all neurons on day 2. Neurons manually matched have the same index assigned on each day. Color bar indicates the image similarity values. Following image registration, image similarity was calculated using the MATLAB structural similarity index (SSIM) for every pair of cell body across consecutive days (see Methods). (C)Distribution of normalized image similarity of manually matched neurons. The image similarity values for each neuron was divided by the maximum value across all its comparisons to yield the normalized image similarity value for each neuron. (D) Distribution of the percentage of neurons that showed the highest similarity rank to its manually matched neuron.



Supplementary Figure 2.5. Relationship between a neuron's changes in responsiveness to CxS and PT. (A) Correlation of the average Euclidean distance of a neuron's response profile from one day to the next for CxS and PT. r =-0.037, p = 0.74 (Pearson's correlation test). (B) Correlation of the change in response magnitude for each responsive neuron to CxS with changes in response magnitude to PTs that had overlapping frequencies with the CxS: left; r = 0.168, p = 10^{-6} (Pearson's correlation test) and with PT stimuli that had minimal overlapping frequencies with CxS: right; r= 0.149, p = 10^{-5} (Pearson's correlation test). The correlations did not significantly differ (p = 0.341, Fisher's z test). PTs with frequency overlap with the CxS were determined as the 3-4 PT frequencies that maximally overlapped with the power spectrum of the CxS.

2.6.1 Table 1: Summary of statistical tests used in the study

Figure	Data Structure	Type of test	Statistical data
1C	Non-normal	Mann-Whitney U test	Fig 1.1 A: Ranksum = 36 , p = 0.00015 Fig 1.1 B: Ranksum = 36, p = 0.00015 Fig 1.1 C: Ranksum = 100, p = 0.00015
1E	Normal distribution with two factors	Two-way ANOVA	F1: F = 0, df = 1, p = 0.98 F2: F = 6.86, df = 7, p <0.0001
1F	Non-normal	Bootstrap test (described above)	10,000 randomly simulated proportions for one group (PT), given the probability of the other group (CxS). p = 0.875
2B & 2C	Non-normal	Chi-square test for proportions and confirmed by Fisher's exact test	Fig. 2B: CXS = 12.15%(66/543), PTs = 22.01%(114/518), Chi^2 statistic = 18.27, p = $1.91e^{-5}$, Fisher's exact test: Odds ratio = 49.03%, CI = [35.22%, 68.26%] Fig. 2C: CXS = 0.14 (34/241), PTs = 0.23 (59/256), Chi^2 statistic = 6.52, p = 0.011,

			Fisher's exact test: Odds ratio =0.55, CI = [0.3445, 0.8730]
2D	Normal distribution	Two-sided t-test	t = -2.0023, p = 0.046, df = 495, sd = 0.51,CI = [-0.1809, -0.0017]
3B	Non-normal	Z-test for proportions	Z value = -1.7811, p = 0.037
3C	Non-normal	Mann-Whitney U test	Z value = -2.8857, p = 0.0039
4A & 4B	Non-normal	Bootstrap test (described above)	10,000 randomly simulated proportions for one group (PT), given the probability of the other group (CxS) computed for each sub category Fig 4A: p = 0.0015 Fig 4B: p = 0.1084
4C	Normal distribution with two factors	Two-way ANOVA with interaction	F1: F = 9.52, df = 3, p < 0.001 F2: F = 5.89, df =1, p = 0.015 F1*F2: F = 0.1, df = 3, p = 0.961

Chapter 3 : A Cortico-Striatal Circuit for Sound-Triggered Prediction of Reward Timing

3.1 Abstract

A crucial aspect of auditory perception is the ability to use sound cues to predict future events and to time actions accordingly. For example, distinct smartphone notification sounds reflect a call that needs to be answered within a few seconds, or a text that can be read later; the sound of an approaching vehicle signals when it is safe to cross the street. Other animals similarly use sounds to plan, time and execute behaviors such as hunting, evading predation and tending to offspring. However, the neural mechanisms that underlie sound-guided prediction of upcoming salient event timing are not well understood. To address this gap, we employed an appetitive sound-triggered reward time prediction behavior in head-fixed mice. We find that mice trained on this task reliably estimate the time from a sound cue to upcoming reward on the scale of a few seconds, as demonstrated by learning-dependent well-timed increases in reward-predictive licking. Moreover, mice showed a dramatic impairment in their ability to use sound to predict delayed reward when the auditory cortex was inactivated, demonstrating its causal involvement. To identify the neurophysiological signatures of auditory cortical reward-timing prediction, we recorded local field potentials during learning and performance of this behavior and found that the magnitude of auditory cortical responses to the sound prospectively encoded the duration of the anticipated sound-reward time interval. Next, we explored how and where these sound-triggered time interval prediction signals

propagate from the auditory cortex to time and initiate consequent action. We targeted the monosynaptic projections from the auditory cortex to the posterior striatum and found that chemogenetic inactivation of these projections impairs animal's ability to predict sound-triggered delayed reward. Simultaneous neural recordings in the auditory cortex and posterior striatum during task performance revealed coordination of neural activity across these regions during the sound cue predicting the time interval to reward. Collectively, our findings identify an auditory cortical-striatal circuit supporting sound-triggered timing-prediction behaviors.

3.2 Introduction

In everyday life, sounds often predict forthcoming events, allowing for planning and execution of appropriate behavioral responses. Consider, for instance, the confidence with which we step out of an elevator's open doors a few seconds after its chime, even without looking up from our phone. Or how the conclusion of a friend's sentence determines the opportune moment for our response (Benichov et al., 2016; Levinson, 2016; Stivers et al., 2009). Similarly, we rely on distinct phone notification sounds to determine whether it is a call that we need to answer within a few seconds, or a text message that we can read a bit later. Likewise, animals rely on sound cues to gauge how swiftly they should vocalize in response to a conspecific call (Benichov & Vallentin, 2020), evade a predator (Z. Li et al., 2021), hone in on prey (Surlykke & Moss, 2000) or approach a needy offspring (Dunlap et al., 2020; Ehret, 2005). These and other examples in everyday life require humans and other animals to utilize sounds to predict and precisely time subsequent salient events and to initiate appropriate behavioral responses within the scale of seconds (Mazzucato, 2022).

The ability to use sounds to predict when future events will occur and consequently when to initiate appropriate action relies on a number of neural processing stages (Bueti, 2011; Buhusi & Meck, 2005; Wiener, Matell, et al., 2011). First, the sound must be detected, processed, and recognized. Second, the predicted amount of time from the sound to the future event is evaluated. And finally, action is initiated when the elapsed time matches the anticipated appropriate time to act. Neural signatures underlying the first stage of this process, namely sound processing and recognition, have been extensively identified in the auditory pathway, and in particular in the auditory cortex (For example, Suri & Rothschild, 2022; Bernal & Ardila, 2016; Geissler & Ehret, 2004; Jasmin et al., 2019; King et al., 2018; King & Schnupp, 2007; Read et al., 2002; Zatorre et al., 2002). Considerably less is known about the second stage, and specifically where and how the brain encodes the predicted amount of time from the sound to a future event. Traditional models of time perception have suggested the existence of a “centralized clock” in the brain (also referred to as the “Internal clock model”) (Hinton & Meck, 1997; Leow & Grahn, 2014; Treisman, 1963; Wearden, 2005). According to this model, the role of sensory regions is to detect the relevant sensory stimulus and communicate this information to higher-order centralized-clock brain regions, where a continuous representation of elapsed time is maintained (Buhusi & Meck, 2005; Hinton & Meck, 1997; Leow & Grahn, 2014; Treisman, 1963). This model proposes that the centralized clock is similarly able to estimate time based on cues from varying modalities, arriving via distinct pathways from the various sensory regions and hence, is “amodal” (Bueti, 2011; Wiener, Matell, et al., 2011). Different studies have implicated a number of brain regions as hosting such centralized clocks, including the nucleus accumbens (Kurti & Matell, 2011),

caudoputamen (Matell et al., 2003), ventral tegmental area (Fiorillo et al., 2008), substantia nigra compacta (Meck, 2006), dorsal striatum (Jones & Jahanshahi, 2011; Meck, 2006; Wiener, Lohoff, et al., 2011) and the medial prefrontal cortex (Jones & Jahanshahi, 2011; Ning et al., 2022; Tunes et al., 2022).

However, the universality of this model has been challenged by studies showing that in addition to a centralized clock, there exist sensory-specific timing mechanisms, and that these are located within early sensory cortical regions (Bueti, 2011; Buhusi & Meck, 2005; Wiener et al., 2011). Early support for this suggestion came from studies demonstrating that the ability to estimate time from a sensory cue depends on the modality of that cue (For example - Hussain Shuler and Bear, 2006; Bueti et al., 2008). Furthermore, recent studies have identified signatures of time estimation within sensory cortices. For example, in vivo neural recordings in the primary visual cortex of rodents show various neural response forms which represent the time interval between the visual stimulus and the anticipated reward (Chubykin et al., 2013; Hussain Shuler & Bear, 2006; Namboodiri et al., 2015). A recent study further showed that this reward timing representation is modulated by an intracortical network of inhibitory interneurons in the visual cortex (Monk et al., 2020).

In the auditory pathway, a key candidate brain region for encoding sound-triggered timing is the auditory cortex (AC), due to its established role in behavior- and decision-making- dependent sound processing (Bathellier et al., 2012; Francis et al., 2018; Fritz et al., 2003, 2005; King & Schnupp, 2007; Lee & Rothschild, 2021; Nelken et al., 2014; Town et al., 2018; Vivaldo et al., 2023). Numerous studies have demonstrated retrospective coding of the degree to which a sound deviates from expectation in AC (Heilbron & Chait,

2018; Khouri & Nelken, 2015; Rubin et al., 2016; Taaseh et al., 2011; Ulanovsky et al., 2003, 2004; Yaron et al., 2012). However, much less is known about the existence of prospective coding of anticipated time from a sound to a subsequent event. In one study, auditory cortical responses to a tone series varied depending on whether a subsequent target sound was expected early (300-450ms) or late (1300-1500ms) within the series (Jaramillo & Zador, 2011). However, this study did not test whether AC encodes expectation of non-auditory cues or whether a continuous representation of predicted time is encoded. In a recent study, mice were trained on a self-paced action timing task, in which lever pressing caused reward delivery after 30 seconds. Optogenetic stimulation and inactivation of the secondary AC implicated its responses to the sound of lever press as being causally involved in timing reward-preparatory action (Cook et al., 2022). However, this study did not test whether and how AC is involved in encoding varying sound-reward time intervals, in particular on the scale of seconds. Thus, it remains unclear whether and how AC is involved in sound-triggered predictive timing of future salient events on the timescale of seconds.

The final step of sound-dependent prediction of imminent events requires initiation of appropriate action once the anticipated time from sound to action has elapsed. To carry this out, the neural information assimilated in the first two stages of this process needs to be sent to downstream brain regions to induce consequent action timing and initiation. Previous studies have implicated a number of brain regions that mediate sound-triggered action initiation, including the striatum (Guo et al., 2018; Matell et al., 2003; Tunes et al., 2022), the medial prefrontal cortex (Tunes et al., 2022; Yumoto et al., 2011) and the supplementary motor area (Mita et al., 2009; Wiener, Matell, et al., 2011). Anatomically,

the most prominent candidate brain region to receive sound-triggered time keeping signals from AC and participate in converting this information to action, is the posterior tail of the dorsal striatum (hereafter referred to as the posterior striatum). The posterior striatum (pStr) receives monosynaptic projections from AC (Bertero et al., 2020; Huang et al., 2023; Znamenskiy & Zador, 2013), and these projections are known to be causally involved in auditory-guided decision making tasks (Znamenskiy & Zador, 2013) and auditory associative learning tasks (Huang et al., 2023). Moreover, pStr itself responds to auditory stimulation (Guo et al., 2018, 2019) and has been shown to play a key role in various stimulus-driven time keeping behaviors (Matell et al., 2003; Tunes et al., 2022). However, whether and how information about sound-triggered time interval estimation arriving from AC engages pStr remains unknown. To collectively address these gaps, we investigated the causal and functional role of the auditory cortical-striatal circuit in sound-triggered prediction of time to consequent reward.

3.3 Materials and Methods

All animal procedures were in accordance with the NIH Guide for the Care and Use of Laboratory Animals and approved by the University of Michigan Institutional Animal Care and Use Committee.

3.3.1 Animals

We used 56 (38 males, 18 females, 8-16 weeks of age) C57BL/6J mice (Jax number: 000664). Mice were individually housed under a reverse 12h light/12h dark cycle, with lights on at 8:30 pm and off at 8:30 am, and had access to ad libitum food, water,

and enrichment. During the behavioral training period, mice were on water restriction and given 2-5ml of water each day.

3.3.2 Surgical procedure

All surgeries were performed on mice anesthetized using isoflurane (1.5-2% vol/vol). Anesthetized mice were placed in a stereotaxic frame (Kopf 514 Instruments, CA, USA), and an anti-inflammatory drug (Carprofen, 5mg/kg, subcutaneous injection) and a local anesthetic (lidocaine, subcutaneous injection) were administered. A custom-made lightweight (<1 gr) titanium head bar was attached to the back of the skull using dental cement and cyanoacrylate glue to allow for head-fixed behavior. During the surgery, body temperature was maintained at 38℃, and the depth of anesthesia was regularly assessed by checking the pinch withdrawal reflex. A small craniotomy was performed over target coordinates relative to the bregma (AC: -2.7mm posterior, ±4.3mm from midline, -0.55mm ventral, 0° angle; pStr: -1.7mm posterior, ±3.35mm from midline, -2.8mm ventral, 0° angle).

To chemogenetically target the neural projections from AC to pStr and for pharmacological inactivation experiments using muscimol, custom-made cannulae (25-gauge tubing) or guide cannulas (Plastics One) were placed at the surface of the brain in these craniotomies at the target regions and secured to the skull using dental cement. Dummy cannulae were inserted into these cannulae to prevent outside debris from entering the cannula.

Mice were treated with Carprofen for 48 hours post-surgically and were allowed to recover for a week.

3.3.3 Electrophysiological recordings

Tungsten wire electrodes (two 50 µm wires bundle) with < 50kΩ of impedance were used to acquire local field potential (LFP) responses in AC and pStr. Electrodes were dipped in DiI dye (Invitrogen, Catalog # 22885) before insertion. The ground screw was positioned over the cerebellum (2.0 mm posterior to lambda, 3.5 mm from midline). Electrodes were unilaterally implanted into AC and pStr on the right hemisphere, (distance between the two electrodes within a region was <50 µm, distance between electrode arrays in AC and pStr was ~2.67 mm). LFP signals were acquired using a Tucker-Davis Technologies (TDT) acquisition system and Synapse Lite Software. The output bioelectrical signal was digitized, sampled at 6 kHz, and bandpass filtered in 0.5-300Hz for LFP recordings. All data acquired was saved for offline data processing.

3.3.4 Virus injections

Viral vectors were acquired from Addgene to inactivate the anterograde projections from AC to pStr chemogenetically. To achieve projection specificity in C57BL/6J WT mice, we used an established dual viral approach as shown in Figure 5A. We bilaterally injected Cre-dependent DREADD viral vector (Roth, 2016): AAV5-hSyn-DIO-hM4D(Gi)-mCherry (2.4E+13 vg/ml, 350nl, Addgene catalog # 44362) or AAV5-hSyn-DIO-mCherry (2.6E+13 vg/ml, 350nl, Addgene catalog # 50459) into the AC and a retrograde Cre viral vector: pENN/AAVrg-hSyn-Cre-WPRE-hGH (1.8E+13 vg/ml, 200nl, Addgene catalog #105553) into the pStr. The infusions were done using a 32-gauge injection needle (customed-length per infusion site) or through thin internal cannulas (Plastics One) inserted into the implanted cannulae in the brain, connected to a 10 µl Hamilton syringe at a rate of 50nl/min.

3.3.5 Drug administration

3.3.5.1 Muscimol infusions

Mildly sedated mice were bilaterally infused with 0.5µg/µl muscimol (BODIPY TMR-X fluorophore-conjugated, ThermoFisher, Catalog Number – M23400) dissolved in phosphate-buffered saline (PBS) and 1.5% DMSO or PBS with 1.5% DMSO as a control (Volume per hemisphere = AC: 750nl (Aizenberg et al., 2015; Sun et al., 2022), pStr: 360nl (L. Guo et al., 2018)) at a rate of 150nl/min, into cannulae implanted in target sites. The infusions were done via custom-made injectors or thin internal cannulas (Plastics One) as previously described.

3.3.5.2 CNO injections

5mg Clozapine-N-Oxide (CNO) (HelloBio) was diluted in 0.9% saline solution. All animals in the chemogenetic inactivation experiments were first injected with saline (5mg/kg, i.p.) as a control and then were injected with the prepared CNO (5mg/kg, i.p.) solution the following day, to chemogenetically inactivate the projections from AC to pStr.

3.3.6 Behavior

All our behavioral setups were custom built and controlled by an Arduino (Arduino Uno board with an Adafruit Music Maker shield) circuit. Behavioral data acquired through the Arduino IDE software was saved in text files for analysis. Videos of animal behavior were acquired using Logitech C920 HD Pro camera on the LogiCapture software and using the Angetube 1080p web camera on the Bandicam software, under red light conditions.

3.3.6.1 Paradigm

Mice were trained on an appetitive sound-triggered reward time prediction task. In this task, mice on water restriction were head-fixed inside a tube to reduce movement-related artifacts, presented a sound cue from a speaker (4Ω, 3W magnetic speakers, placed ~10cm away from the animal's head on the left) and trained to consume a water reward from a reward port placed close to its mouth. A trial constituted a 1.5s long sound cue and a water reward delivery separated by a fixed time interval (0.5-5s), with randomized inter-trial intervals in the range of 2-6s (Figure 3.1A, Trial block). The sound cue was a sequence of three 0.5s long pure tones (8kHz, 12kHz, 16kHz; 5ms rise/fall time) generated at a 25kHz sampling rate using MATLAB (Mathworks 2019a).

3.3.6.2 Training

Water-restricted mice were handled and habituated to the experimental setup for ~7 days. In this period, mice were head fixed and trained to lick the reward port through which a water reward was delivered randomly at 3-10s intervals without any sound cue. The reward port consisted of a metal tube that delivered a fixed amount of water (~3ul) each trial, connected to a capacitance-based lick detector that allowed recording individual lick times. Mice were also familiarized with the sound cue used in behavioral training over the last 3 days of habituation through random sound presentations (~50 times across all 3 days) at 3-10s intervals without any reward delivery.

Habituated mice started training on trials with a fixed time interval of 1.5s from sound termination time to reward, with 150-250 trials per daily training session. After 7-10 days of training, catch trials in which reward was withheld were randomly introduced 15-20% of the trials/session. Once the animal learnt to predict reward time at 1.5s interval,

it was then trained to use the same sound to predict a different interval of time between sound and reward (sound-reward interval). We trained each animal to predict timed reward using the same sound cue at four sound-reward intervals – 0.5s, 1.5s, 2.5s and 5s and always trained them to learn the time intervals in this order – 1.5s → 2.5s → 5s → 0.5s (Figure 3.1A).

To identify whether mice predicted the sound-reward interval duration from sound onset or sound termination, we tested a subset of mice trained to predict reward at 1.5s from sound cue, to use a shorter duration sound cue to predict reward. On this testing day, we randomly interspersed 35% of the trials with 1s long sound cue (a sequence of 8kHz and 12kHz pure tones, each 0.5s long, with 5ms rise/fall time) to deliver reward at the same time interval – short sound trials (Figure 3.3A), along with standard 1.5s long sound cue trials.

3.3.6.3 Behavioral training for the pharmacological inactivation experiments

In these experiments, we used a GABA-A receptor agonist, muscimol, to inactivate AC or pStr to establish their causal role in sound-triggered reward time prediction task. Each of these experiments consisted of two different cohorts of mice. One cohort of animals were trained to predict timed reward at 1.5s from sound termination (Figures 3.2, 3.5, and 3.6; 1.5s Delay task) and another cohort of animals were trained on an alternative version of the task where reward immediately followed the sound cue (Figures 3.2, 3.5, and 3.6; No-Delay task).

3.3.6.4 Behavioral training for acquiring electrophysiological recordings in AC and pStr

Mice implanted with wire electrodes underwent the same habituation protocol as described above. In this habituation phase, these mice were habituated to cables plugged to the electrode connectors on their head implants and trained to lick the reward port to consume the water reward. The reward port for these experiments was a tube fitted with an IR sensor to detect licks through beam breaks. Following habituation, they were trained on the previously described sound-triggered reward time prediction task on the 4 different time intervals between sound and reward in the order 1.5s → 2.5s → 5s → 0.5s, while their LFP responses in AC and pStr were recorded throughout each daily training session. LFP responses were monitored for movement using video recording and periods of movement were eliminated prior to analysis.

3.3.6.5 Behavioral training for chemogenetic inactivation of AC-pStr projections experiments

We used chemogenetic inactivation of anterograde projections from AC to pStr to identify their role in sound-triggered reward time prediction task. Animals in these experiments started by learning to predict timed reward at 1.5s from sound (1.5s Delay task) and then underwent chemogenetic AC-pStr projection inactivation with CNO injection to test for effect on behavioral performance. After a 4-day washout period, a subset of these same animals underwent training to predict sound-guided timed reward at 5s interval (5s Delay task) and chemogenetic inactivation with CNO injection to test for behavioral effect to predict timed reward at 5s from sound. Following another 4-day

washout period, these animals trained and underwent chemogenetic manipulation on the No-Delay task (Figure 3.5A, Task timeline).

In a subset of animals in which DREADDs were expressed in the AC-pStr projections, we also simultaneously recorded LFP responses in pStr using tungsten wire electrodes, while they trained on the 1.5s Delay and No-Delay tasks.

3.3.7 Data analysis

All analyses were done using custom-written MATLAB (Mathworks 2022a) scripts unless otherwise mentioned.

3.3.7.1 Behavioral data analysis

To quantify the animal's ability to predict reward at a fixed time interval from a sound cue, we extracted individual lick times per trial and averaged these licks across trials for each daily training session to get a predictive licking response curve. To measure learning, we computed the slope of this predictive licking response curve (MATLAB command: polyfit, degree 1) in the predictive lick period, which was defined differently for rewarded and catch trials. For rewarded trials, the predictive lick period was defined as the time from sound termination to 100ms prior to reward delivery time for each sound-reward interval. Contrastingly, we used a more conservative definition of the predictive lick period for catch trials, using the period from 200ms prior to sound termination to the time at which reward was expected for the 0.5s interval, for all the sound-reward intervals. We used this slope measure in catch trials to ascertain when the animals had learnt to consistently predict reward within each sound-reward time interval. We compared the slopes of the predictive licking curve for catch trials across training days for each sound-

reward interval and picked those days for which the slope value crossed the slope threshold (mean + 3*standard deviation of slope values across all training days per sound-reward interval). Amongst the training days that satisfied this criterion, the day with maximum slope value was chosen as the “best” behavior day for each sound-reward interval per animal. Additionally, we computed the full width at half maxima of the predictive licking response curve over the predictive lick period defined above for catch trials to estimate the precision of the animal’s ability to predictive lick for each sound-reward interval (Supplementary Figure 3.1).

We determined the effect of muscimol and chemogenetic inactivation on behavioral performance by comparing the predictive lick responses in catch trials on control training day (PBS infusion for pharmacological inactivation experiments or saline injections (i.p.) for chemogenetic inactivation experiments) to the predictive lick responses on the manipulation day (muscimol (MUS) infusion for pharmacological inactivation experiments or CNO injections (i.p.) for chemogenetic inactivation experiments) using this formula –

$$\log Predictive\ licking\ ratio\ (PLR)$$

$$= \log \frac{Avg.across\ trials\ [\#\ of\ licks\ in\ (Predictive\ lick\ period - Baseline\ period)]on\ PBS\ or\ Saline}{Avg.across\ trials\ [\#\ of\ lick\ in\ (Predictive\ lick\ period - Baseline\ period)]\ on\ MUS\ or\ CNO}$$

Where, baseline period = sound onset time – 750ms to sound onset time, and

predictive lick period for delay tasks = reward time – 250ms to reward time + 500ms

for no-delay task = sound onset time + 750ms.

To check whether animals' ability to lick for reward changed based on sound termination time or not, we compared their predictive lick responses in standard and short sound trials using a variation of the PLR described above –

$$log\ PLR\ for\ Standard\ vs. Short\ Sound\ trials$$

$$= \log \frac{\begin{matrix} Avg. across\ trials\ [\#\ of\ licks\ in\ (Predictive\ lick\ period - Baseline\ period)] \\ for\ [STANDARD\ TRIALS] + for\ [SHORT\ SOUND\ TRIALS] \end{matrix}}{\begin{matrix} Avg. across\ trials\ [\#\ of\ licks\ in\ (Predictive\ lick\ period - Baseline\ period)] \\ for\ [STANDARD\ TRIALS] - for\ [SHORT\ SOUND\ TRIALS] \end{matrix}}$$

Where, baseline period = sound onset time – 750ms to sound onset time, and

predictive lick period = sound offset time to reward time-100s

3.3.7.2 Electrophysiological data processing and analysis

Acquired electrophysiological data was extracted using TDTBin2mat script (provided by TDT) and organized to synchronize it with lick response times for each session across sound-reward intervals per animal. Using slopes of the predictive lick response curve of catch trials, the "best" behavior day was determined for each sound-reward interval per animal, as described above. Trials in each training session with no licks in the 200ms period from sound onset or in the 500ms period from sound offset (hereafter referred to as no-lick trials) were extracted and further analysis was carried out only on these no-lick trials on the "best" behavior days for each sound-reward interval across animals. Session-wise LFP signals for AC and pStr were filtered for movement-related artifacts by eliminating any signal above a threshold of mean LFP signal for the session + 3*standard deviation of the session LFP signal and then were z-scored for each session prior to analysis.

To analyze sound-evoked LFP responses, trial-wise LFP activity in AC and pStr were aligned to the sound onset and baseline-corrected by subtracting its average during the 10ms period from sound onset for computing the onset response magnitude (similar to (Taaseh et al., 2011)) or from sound offset for computing the offset response magnitude. Response magnitude was defined as the amplitude of the maximum trough (most negative) in the 40ms period from sound onset/offset for each no-lick trial per session. Response magnitudes were averaged across trials per session and normalized to the average response magnitude of the shortest sound-reward time interval (0.5s) session per animal. Normalized response magnitudes were combined across animals per sound-reward interval and compared using the Kruskal-Wallis test (MATLAB command: kruskalwallis), with a post-hoc Tukey-Kramer test to determine individual group differences (MATLAB command: multcompare applied on the kruskalwallis output).

To compare the pStr onset response magnitudes in saline and CNO conditions, we averaged the response magnitudes across trials and tested for significant differences per animal between the conditions using the Wilcoxon rank-sum test (MATLAB command: ranksum). These average response magnitudes were then normalized to average of the saline condition to determine the population-level trends for the 1.5 Delay and No-Delay tasks.

We examined the coordination in AC and pStr LFP activity during sound by computing the trial-by-trial correlation of sound onset response magnitudes in AC and pStr (MATLAB command: corrcoef). Significant correlation coefficients across animals were combined per sound-reward interval and compared across intervals using the Kruskal-Wallis test (MATLAB command: kruskalwallis), with a post-hoc Tukey-Kramer

test to determine individual group differences (MATLAB command: multcompare applied on the kruskalwallis output).

To determine the temporal relationship of AC and pStr LFP responses during sound-triggered reward time prediction behavior, we ran a cross-correlation analysis of the LFP responses in AC and pStr for each session (MATLAB command: xcorr, using the "normalized" model, maxlag = ±100ms) over a 2s period from 0.5s prior to sound onset to sound termination time. By shuffling the identity of AC and pStr traces per session and computing their cross-correlation coefficients, we simulated chance cross correlation coefficients between AC and pStr sound responses (Number of iterations = 100). We generated difference traces by calculating the difference in average cross correlation coefficients between real and simulated data, and then determined the time which these difference traces peaked giving us the time lag between AC and pStr sound responses per training session. We calculated the median of this peak time of difference traces across all animals and sound-reward intervals and compared this median against 0s using the Wilcoxon signed rank test (MATLAB command: signrank).

3.3.7.3 Statistical tests

We used statistical tests at a $p<0.05$ significance level and $\alpha = 0.05$ for all comparisons unless otherwise indicated.

3.3.8 Histology

Mice were euthanized with an overdose of isoflurane (5%) or carbon dioxide (2%) and perfused transcardially with PBS (0.9%), followed by 10% paraformaldehyde (PFA). The brain tissue was removed and fixed in 10% PFA for 72 hours. For cryoprotection, the

brain tissue was transferred to 30% sucrose solution for 3-4 days before sectioning. Coronal sections (50µm thickness) were obtained in a cryostat (Leica) and kept in PBS at 4°C before mounting. All sections were mounted in glass slides and covered with Fluoroshield mounting medium with DAPI (Abcam, USA). Images of brain sections were acquired using a fluorescent microscope (Zeiss) equipped with an apotome and the ZenPro software.

For all histological validations, brain sections were imaged with a 10x objective, examined for cell nuclei labeled with DAPI (470 nm), and saved as both multichannel and individual fluorophore channels composite tiff images. Histological validation was done by overlaying brain sections over slice images from the Allen Institute's 10 µm voxel 2017 version from the Allen Mouse Brain Common Coordinate Framework (Q. Wang et al., 2020). Histological validation was used as an inclusion criterion for all behavioral, electrophysiological, and chemogenetic experiments.

3.3.8.1 Tungsten electrode tracks verification

Electrode tracks were identified by DiI (565 nm) or DiO (550 nm) labeling. Images with DiI or DiO labeling and electrode tracks were saved and further analyzed with the open-source SHARP-Track toolkit (Shamash et al., 2018). Verified electrode tip coordinates were compared to the coordinate range showing the projections from AC to pStr shown in the Allen Mouse Brain Connectivity Atlas (Allen Mouse Brain Connectivity Atlas, connectivity.brain-map.org/projection/experiment/146858006).

3.3.8.2 Reconstruction of virus expression

The procedure for reconstructing viral expression volumes was similar to electrode track reconstruction. However, the regions defined in each slice are two-dimensional, and the volume is delineated at the 3D projection of all the combined sections showing virus expression. AP, ML, and DV coordinates from all sections with positive mCherry expression were again compared to the range of AC-pStr projections shown in the Allen Mouse Brain Connectivity Atlas (Allen Mouse Brain Connectivity Atlas, connectivity.brain-map.org/projection/experiment/146858006). Animals showing mCherry cell body expression outside AC areas and mCherry axonal expression outside pStr were not included.

3.3.8.3 Muscimol infusion

To verify the muscimol diffusion within the target areas, brain slices were examined for BODYPY-TmX (BDP-T, 573 nm) labeling. Brain sections showing positive BDP-T labeling were saved per animal. Further comparison with the reference Allen Atlas and 3D projection showed the AP and ML coordinates with muscimol diffusion. Clear hit into AC or pStr was considered a diffusion only within the target areas. For pStr experiments, animals with muscimol diffusion up to rostral striatal areas (anterior to -1.2mm from bregma) were not included.

3.4 Results

3.4.1 Mice use sound to predict reward time with 1-second temporal resolution

To investigate the role of auditory cortical-striatal neural mechanisms underlying sound-triggered reward timing prediction, we employed an appetitive sound-guided trace

conditioning task in water-restricted mice. Following habituation to head fixation, 8 mice underwent training sessions of 150-200 trials each, in which each trial was initiated with a 1.5s long sound stimulus (composed of a sequence of three pure tones), followed by a fixed delay period and then delivery of water reward (Figure 1A). Trials were separated by inter-trial intervals which randomly varied between 2-6s in duration. All animals started behavioral training sessions with a fixed delay period of 1.5s between the sound cue termination and reward ("sound-reward time interval"). After 7-10 days of training, we introduced randomly interspersed 10-20% catch trials, in which the reward was withheld. Mice expressed learning of the sound-reward contingency by increasing their lick rate before the anticipated reward time ("Predictive licking", Figure 1B, left) and by licking before and during the time of anticipated reward in the catch trials (Figure 1B, right).

To determine whether predictive licking reflects a reliable estimation of anticipated reward time, the same mice then went on to train on the same paradigm but with different sound-reward time intervals. We posited that if mice can reliably estimate time on the scale of seconds and use these estimates to guide their behavior, the timing of predictive licking would vary with the duration of the sound-reward interval. To test this, once animals showed reliable performance on the task with the 1.5s sound-reward interval (as evidenced by reliable predictive licking) they relearned versions of this task with a 2.5s, 5s and 0.5s sound-reward time intervals, in this order (Figure 1A, Behavior Timeline). For each animal, we compared the slopes of their predictive licking curves on their "best" behavior days of each of these sound-reward time intervals (see Methods). We computed the slopes of the predictive licking curves for rewarded and catch trials separately (see Methods). We found that the predictive licking slopes varied with the duration of the

sound-reward interval (Figure 1C). Across animals, the average slope of the predictive licking curves significantly varied as a function of the sound-reward interval duration, with the shortest interval duration (0.5s) inducing the steepest slope of predictive licking, followed by 1.5s, 2.5s and 5s (Figure 1D, Kruskal-Wallis test for multiple comparisons, p = 0.0011 for rewarded trials, p = 0.0064 for catch trials). We also quantified the precision of predictive licking for each of the sound-reward intervals by computing the full width at half maxima of the predictive licking curve for catch trials and found that mice are able to predict the shorter intervals with higher precision compared to the longer intervals (Supplementary Figure 1, Kruskal-Wallis test for multiple comparisons, p = 3.85x10-6). These results show that mice can estimate time intervals on the scale of 0.5-5s with at least 1-s temporal resolution and use these estimates to predict the time of expected reward following a sound.

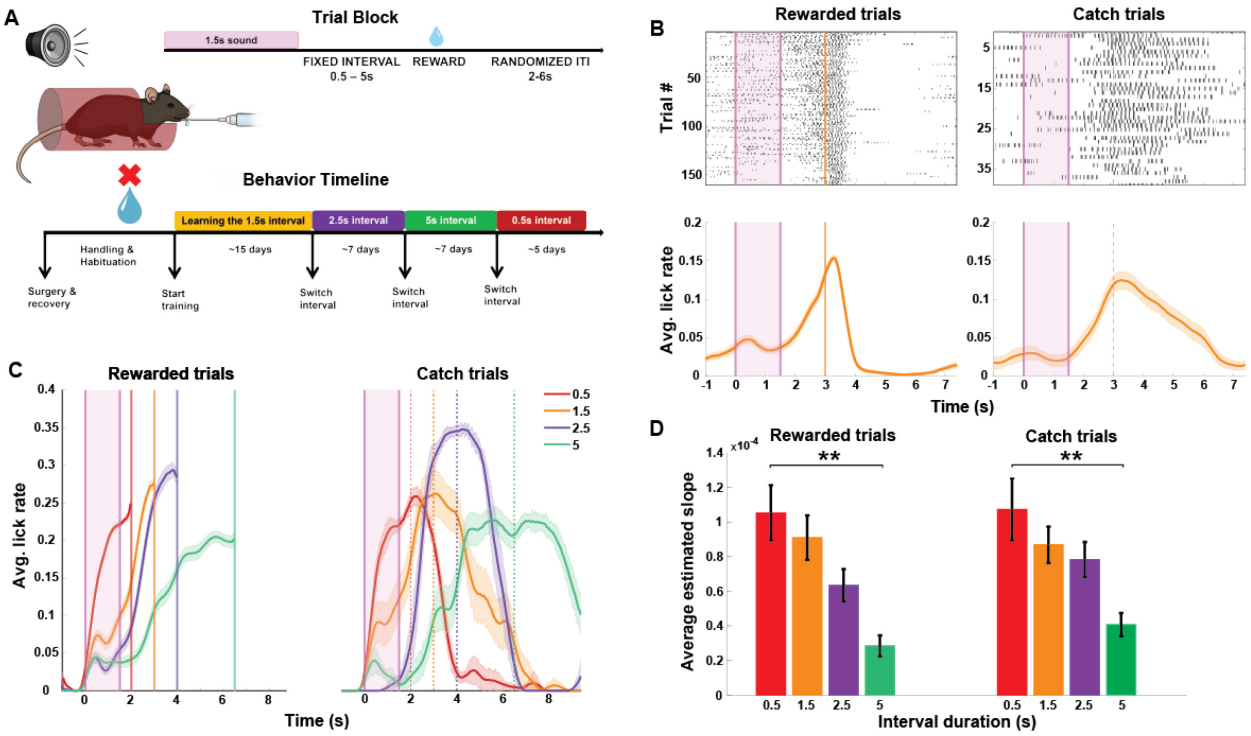


Figure 3.1 Mice predict reward timing using a sound cue. A. An illustration of the behavioral setup for sound-triggered reward time prediction task, components of a trial block and the experimental timeline for behavioral training. B. Top: Peri-sound lick raster of an example behavioral session from a trained animal performing on rewarded (left) and catch (right) trials within the session. Bottom: Average peri-sound lick rate response curve (solid line denotes mean, shaded area represents SEM across trials) for the example behavioral session above for rewarded (left) and catch (right) trials. Shaded pink region represents the 1.5s long sound period. Solid and dotted orange lines represent when reward was given in rewarded trials and expected in catch trials. Black ticks represent licks. C. Average peri-sound lick rate response curves (solid line denotes mean, shaded area represents SEM across trials) of an example animal trained to perform on the four different sound-reward intervals represented by different colors. Left: Rewarded trials; Right: Catch trials. Shaded pink region represents the 1.5s long sound period. Solid and dotted lines represent when reward was given in rewarded trials and expected in catch trials for each of the sound-reward interval. D. Average estimated slope of predictive licking response curves for each of the four sound-reward intervals across all animals (N = 8) for rewarded trials (left,**p=0.0064, Kruskal-Wallis test) and for catch trials (right, **p =0.0011, Kruskal-Wallis test). Error bars represent mean ± SEM across animals.

3.4.2 The auditory cortex is required for sound-triggered delayed reward prediction

The auditory cortex (AC) plays an important role in predictive and behavior-dependent sound processing (Bathellier et al., 2012; Francis et al., 2018; Heilbron & Chait, 2018; Kuchibhotla & Bathellier, 2018; Kuchibhotla et al., 2017; J. Li et al., 2017; Okada et al., 2018). Since our task requires mice to use the sound to time their behavior, we hypothesized that the AC is necessary for successful performance of this task. To test this hypothesis, we measured the influence of bilateral AC inactivation using the GABA-A receptor agonist muscimol on trained animals' ability to predict reward at a 1.5s sound-reward time interval. Muscimol labeling was histologically validated at the end of each experiment and only animals with selective targeting in AC were included (Figure 2A). AC inactivation resulted in an overall reduction in the ability to predictively lick for reward as compared to infusion of inert PBS as a control (Figure 2B). An impaired ability to predictively lick for reward following muscimol infusion was consistently observed in all our animals (N=8, $p<0.05$ for each mouse, Wilcoxon rank-sum test). To rule out the

possibility that reduced predictive licking reflects a simple impairment in sound processing, we trained another cohort of mice on a variation of the task, in which reward was delivered immediately following sound onset (No-Delay Task). Mice trained on the No-Delay task showed evidence for sound-reward association by consistently licking in response to the sound during catch trials, when no reward was delivered (Figure 2C, right). Interestingly, this form of predictive licking following the No-Delay sound-reward association was unaffected by AC inactivation (N=8, $p>0.05$ for each mouse, Wilcoxon rank-sum test; Figure 2C). To directly compare the influence of AC inactivation on prediction of delayed and immediate reward, we calculated the log of the ratio of predictive licking in PBS and muscimol ("log PLR") for each of the task versions. Thus, larger log PLR values indicate a greater influence of AC inactivation on predictive licking. Using this metric, we found that AC inactivation had a significantly more detrimental effect on sound-guided prediction of 1.5 s-delayed reward than on immediate reward ($p=0.000156$, Wilcoxon rank-sum test, Figure 2D). These findings demonstrate that the AC is causally involved in sound-triggered time-delayed reward prediction.

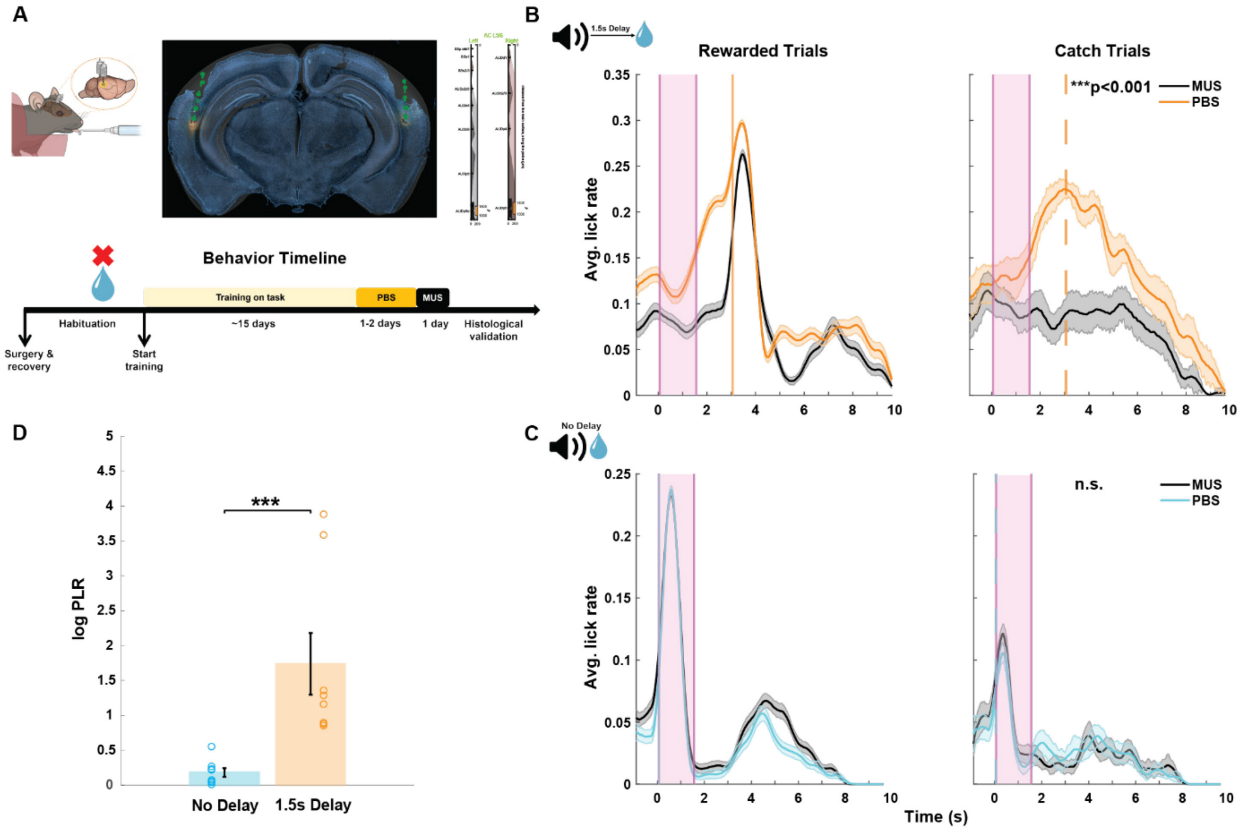


Figure 3.2: AC is required for sound-triggered delayed reward prediction. A. Top: An illustration of the cannula implanted in AC for muscimol infusion. Right: Histological verification of bilateral muscimol infusion into AC. The brain slice acquired from an example animal is overlaid with the corresponding coronal section from the Allen Mouse Brain Atlas (see Methods). The cannula tracks are indicated by the green dotted line on the brain slice and the site of muscimol infusion is seen in orange. Markers on the right indicate the depth at which muscimol was infused in the left and right hemispheres. Scale bar: 1000μm. Bottom: Behavioral timeline for mice that underwent training on the 1.5s Delay task. B. Average peri-sound lick rate response curves (solid line denotes mean, shaded area represents SEM across trials) of an example animal trained to predict reward at 1.5s sound-reward interval when infused with PBS (orange) and muscimol (MUS, black) in AC for rewarded trials (left) and catch trials (right). Shaded pink region represents the sound period. Solid and dotted orange lines represent when reward was given in rewarded trials and expected in catch trials. ***p=0.0007 (Wilcoxon rank-sum test) denotes the significant difference in predictive licking compared to baseline for catch trials on PBS and MUS conditions. C. Average peri-sound lick rate response curves (solid line denotes mean, shaded area represents SEM across trials) of an example animal trained on the No-Delay task when infused with PBS (light blue) and muscimol (MUS, black) in AC for rewarded trials (left) and catch trials (right). Shaded pink region represents the 1.5s long sound period. Solid and dotted blue lines represent when reward was given in rewarded trials and expected in catch trials. n.s. (p=0.939, Wilcoxon rank-sum test) denotes the not significant difference in predictive licking between catch trials on PBS and MUS in the No-Delay task. D. Significant difference in average log predictive licking ratio (log PLR) between No-Delay (N = 8) and 1.5s Delay (N = 8) cohorts (***p =0.000156, Wilcoxon rank-sum test). Error bars represent mean ± SEM across animals. Open circles indicate the PLR for each animal in the No-Delay (blue) and 1.5s Delay cohorts.

3.4.3 Mice predict reward timing from sound onset

Having established that mice can use the sound cue to predict the timing of reward across varying sound-reward time intervals and that this behavior causally involves the auditory cortex, we next asked whether this behavior reflects time estimation from sound onset or from sound termination (Figure 3A). To this end, we compared the predictive lick pattern under the standard paradigm of a 1.5s-long sound and 1.5s-long sound-reward interval, to a similar paradigm in which the sound duration was cropped by 0.5s, to a duration of 1s. We argued that if mice use sound onset to predict reward timing, their predictive lick pattern would be unaffected by a shorter sound duration, whereas if mice use sound termination to predict reward timing, predictive licking will start earlier in the trials with the shorter sound cue. To test this, we trained a cohort of 8 mice to predict reward at a 1.5s sound-reward interval and then introduced them to a training session with 35% short sound trials (Figure 3A). All our animals showed similar behavioral responses on short sound and standard sound trials, with almost identical predictive licking curves (Figure 3B). Across animals, the slopes for the predictive licking curves were not significantly different between the two types of trials (Figure 3C, Wilcoxon rank-sum test, $p = 0.96$), and the average log PLR for standard sound to short sound trials was not significantly different from 0 (See Methods, Figure 3D, Wilcoxon signed rank test compared against 0, $p = 0.12$). These results suggest that mice primarily use the sound onset to estimate the amount of time from sound to reward.

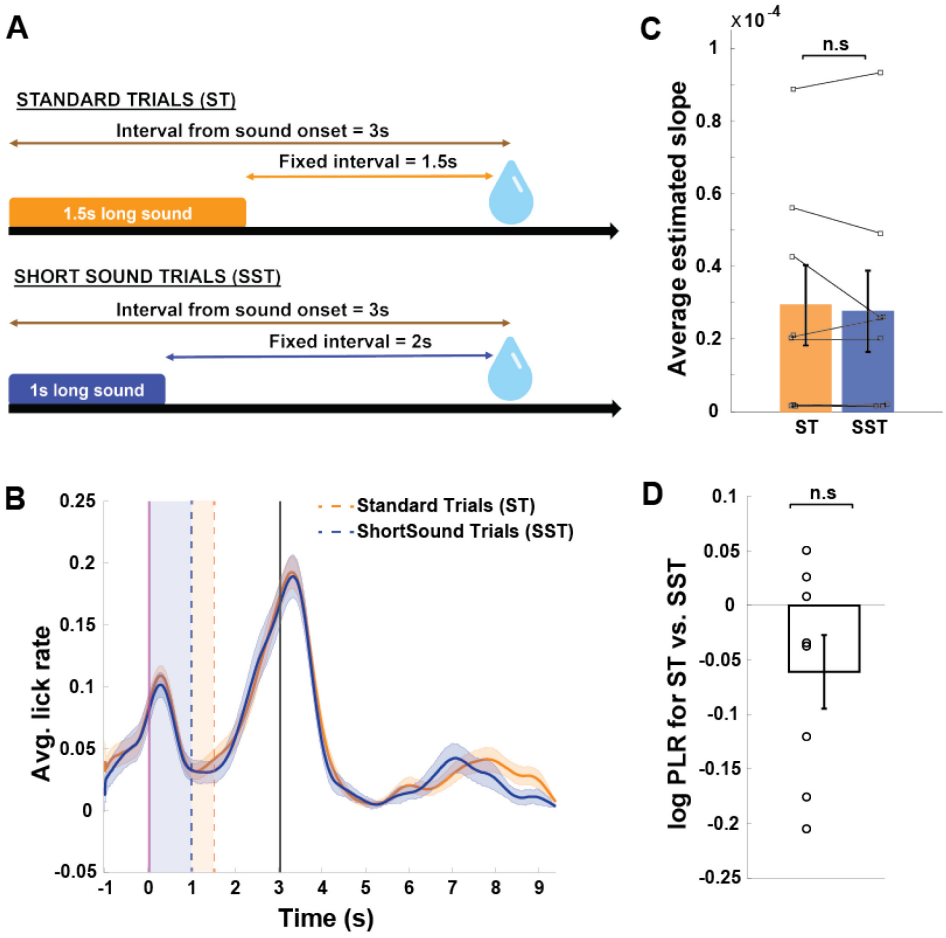


Figure 3.3: Mice predict reward timing from sound onset. A. An illustration of the trial designs. B. Average peri-sound lick rate response curves (solid line denotes mean, shaded area represents SEM across trials) of an example animal trained on the Standard Trials (ST, orange) and Short Sound Trials (SST, navy blue). Dashed lines indicate the sound termination times for ST (orange) and SST (blue) trials. Solid black line represents when reward was given. Comparison of the predictive licking between ST and SST trials compared to baseline yields p = 0.9 (Wilcoxon rank sum test). C. Average estimated slope of the predictive licking curves for ST (orange) and SST (navy blue) across animals (N = 8, p = 96, Wilcoxon rank sum test). Error bars represent mean ± SEM across animals. Lines connecting filled squares represent the estimate slope values for each animal in ST and SST. D. Average log predictive licking ratio (log PLR for SST) across animals is not significantly different from 0 (N = 8, p = 0.12, Wilcoxon signed rank test compared against 0). Error bars represent mean ± SEM across animals. Open circles represent the PLR for each animal.

3.4.4 AC sound responses encode predicted time-to-reward

A prominent model for the neural processing underlying sound-triggered time predictions suggests that following sound coding in the auditory pathway, downstream

brain regions encode the anticipated time interval from the sound to subsequent events (Buhusi & Meck, 2005; Hinton & Meck, 1997; Leow & Grahn, 2014; Treisman, 1963). Alternatively, given the established role of the AC in predictive coding, the predicted time from sound to reward may already be reflected in the AC response to the sound itself. To test this possibility, we recorded local field potential (LFP) activity in the AC of the right hemisphere in 8 mice, as they were trained to predict timed reward using sound at the four different sound-reward intervals in the order described previously. We histologically verified the position of these electrodes at the end of the experiments (Figure 4A). Similar to our findings in the first cohort (Figure 1), animals in this cohort also changed the rate of their predictive licking curves according to the duration of the sound-reward intervals, with the slope increasing with increased interval durations for both rewarded and catch trials (Figure 4B and C, $p = 0.0000544$ for rewarded trials and $p = 0.00012$ for catch trials; Kruskal-Wallis test for group differences). We similarly identified the best behavior performance day of each sound-reward interval for analysis and excluded trials in which the animal moved or in which the animal licked in the 200 ms period from sound onset to avoid a contribution of motor activity to the sound response magnitude (Schneider et al., 2014; Vivaldo et al., 2023; Whitton et al., 2014). We then quantified the magnitude of the AC responses to the sound for each of the interval durations per animal. As our behavioral results showed that mice use the sound onset to estimate the time to reward, we focused on the responses to sound onset. Interestingly, we found that AC responses (to the same sound) increased in magnitude with increasing sound-reward intervals (Figure 4D). Across animals, the average normalized sound response magnitude was significantly different across interval durations ($p=1.9x10-35$, Kruskal-Wallis test), with the response

magnitude increasing as a function of the time interval from sound to reward (Figure 4E). We also quantified the magnitude of responses to the sound offset for each of the interval durations per animal and found that these offset responses did not significantly change in magnitude across sound-reward intervals (p = 0.263, Kruskal-Wallis test, Figure 4F and G). Together, these results indicate that AC neural responses to the sound onset encodes, beyond the sound itself, the predicted time from sound to reward.

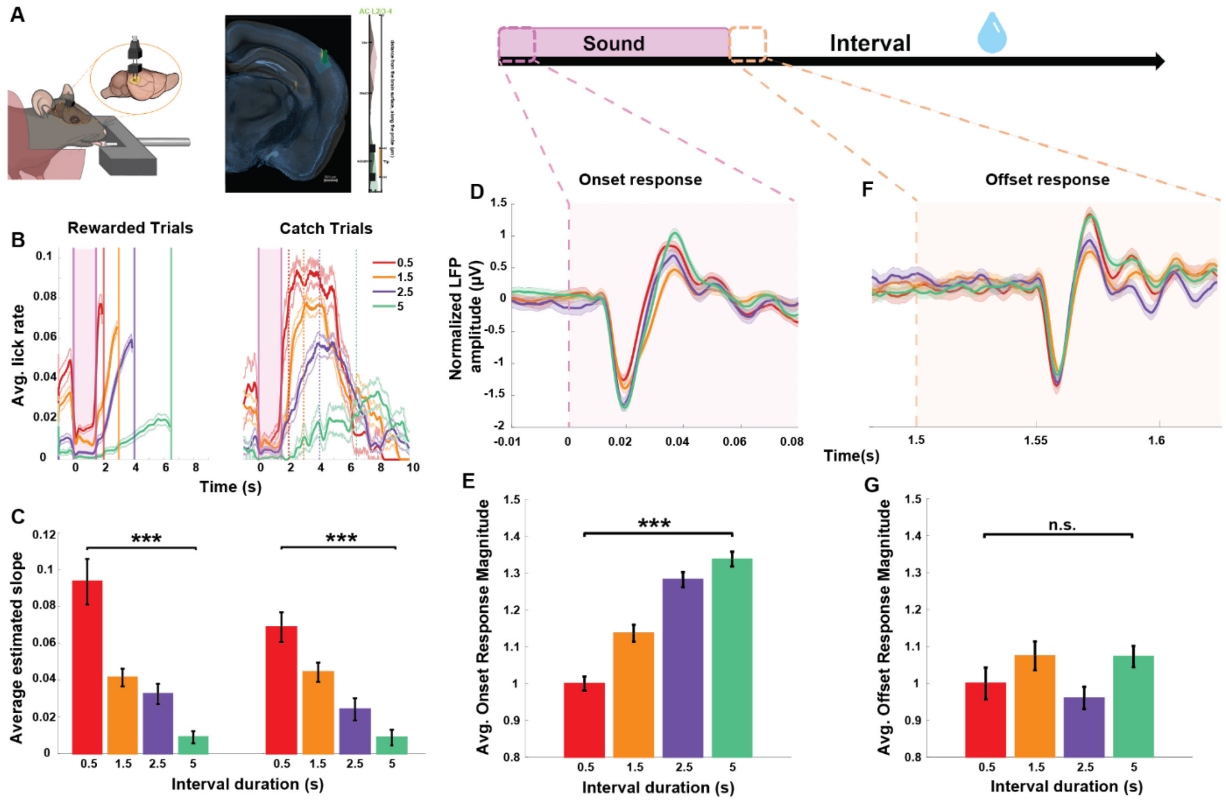


Figure 3.4: Auditory cortical sound responses encode predicted time to reward. A. Left: An illustration of the electrodes implanted in right AC. Right: Histological verification of electrode position in AC of the right hemisphere. The brain slice acquired from an example animal is overlaid with the corresponding coronal section from the Allen Mouse Brain Atlas (see Methods). The electrode track is indicated by the green dotted line on the brain slice and the markers on the right indicate the depth at which the electrode was implanted in the right hemisphere. Scale bar: 500μm. B. Average peri-sound lick rate response curves (solid line denotes mean, shaded area represents SEM across trials) of an example animal trained to perform on the four different sound-reward intervals represented by different colors. Left: Rewarded trials; Right: Catch trials. Shaded pink region represents the 1.5s long sound period. Solid and dotted lines represent when reward was given in rewarded trials and expected in catch trials for each of the sound-reward interval. C. Average estimated slope of predictive licking response curves for each of the four sound-reward intervals across all animals (N = 8) for rewarded trials (left,***p=0.000054, Kruskal-Wallis test) and for catch trials (right, ***p =0.00012, Kruskal-Wallis test). Error bars represent mean ± SEM across animals. D. Normalized average AC LFP (solid line denotes mean, shaded area represents SEM across no-lick trials) recorded in

response to the sound onset from an example animal trained on the four different sound-reward intervals (represented by the different colors). The shaded pink region represents the period from sound onset as shown in the illustration above. E. Average onset response magnitude computed across animals (N = 8) for each of the sound-reward intervals. Error bars represent mean ± SEM across animals. Comparison across sound-reward intervals yields ***$p = 1.9x10^{-35}$ (Kruskal-Wallis test). F. Normalized average AC LFP (solid line denotes mean, shaded area represents SEM across no-lick trials) recorded in response to the sound offset from an example animal trained on the four different sound-reward intervals (represented by the different colors). The shaded yellow region represents the period from sound offset as shown in the illustration above. G. Average offset response magnitude computed across animals (N = 8) for each of the sound-reward intervals. Error bars represent mean ± SEM across animals. Comparison across sound-reward intervals yields p = 0.263 (Kruskal-Wallis test).

3.4.5 Auditory cortical projections to posterior striatum are causally involved in sound-guided prediction of delayed reward

For successful sound-triggered reward timing prediction, animals need to translate their prediction of reward time into motor action, which in the current task is licking. While our data suggests that the AC is involved in sound-triggered reward timing prediction, we hypothesized that this behavior and its translation into motor action would further depend on the posterior striatum (pStr). The pStr is a key candidate brain region to support this function, as it receives strong monosynaptic projections from the AC (Bertero et al., 2020; Huang et al., 2023; Znamenskiy & Zador, 2013) as well as from the thalamic medial geniculate body (Huerta-Ocampo et al., 2014; LeDoux et al., 1991; Smeal et al., 2008) and is involved in sound processing and sound-guided behaviors (Huang et al., 2023; Znamenskiy & Zador, 2013). Moreover, pStr itself is known to be involved in appetitive auditory frequency discrimination tasks (Guo et al., 2018, 2019). Hence, we investigated the role of auditory cortical projections to the pStr in sound-triggered reward time prediction.

To address this, we chemogenetically inactivated the projections from AC to pStr in animals trained to predict timed reward using a sound cue. Using a dual virus approach, we bilaterally injected and expressed a Cre-dependent DREADD viral vector, AAV-hSyn-

DIO-hM4D(Gi)-mCherry, in the AC of mice in the experimental group, (or AAV-hSyn-DIO-mCherry in the control group), and a retrograde-CRE viral vector (pENN/AAVrg-hSyn-Cre-WPRE-hGH) in the pStr. This approach allowed us to specifically target the projections from the AC to pStr for chemogenetic inactivation. The expression of these viruses was histologically validated at the end of the experiments and only those with targeted virus expression in AC cell bodies and pStr axonal projections were included (Figure 5B).

Mice were trained to reliably predict sound-guided reward at a 1.5s sound-reward interval. They were then injected with saline (i.p.) as a control and their behavioral performance was recorded 30 minutes after the injection. 24 hrs later, we injected them with CNO (5mg/kg, i.p.) and after 30 min again recorded their behavioral performance (Figure 5A). Chemogenetic silencing of the AC-pStr projections in the mice of the experimental group significantly reduced their ability to predict reward at 1.5s (Figure 5C). We observed this significant effect in all mice in the experimental group ($p<0.05$ for each mouse, Wilcoxon rank-sum test). In contrast, none of the mice in the control group showed a significant change in predictive licking following CNO injection compared to their performance with saline injection (Figure 5F, $p>0.05$ for each mouse, Wilcoxon rank-sum test). These findings demonstrate that the AC-pStr projection is necessary for sound-triggered prediction of delayed reward.

Following a 4-day CNO-washout period, we trained a subset of these mice in the experimental and control groups to use the same sound cue to predict reward at a 5s interval from sound (Figure 5A). We chemogenetically inactivated the AC-pStr projections again and found that all mice in the experimental group showed poor reward prediction

at 5s interval following CNO injection compared to their performance with saline injection on the previous day (Supplementary Figure 2A, $p<0.05$ for each mouse, Wilcoxon rank-sum test). In contrast, mice in the control group did not show an impairment in their ability to predict timed reward at 5s from saline to CNO behavioral sessions (Supplementary Figure 2B, $p>0.05$ for each mouse, Wilcoxon rank-sum test).

To verify that chemogenetic inactivation of AC-pStr projections does not impair animals' ability to process sounds and form sound-reward associations, we trained all mice in the experimental and control groups on the No-Delay task following another 4-day CNO washout period and followed the same manipulation protocol above (Figure 5A). Mice in both these groups did not show an impairment in their ability to lick for reward immediately following sound in the catch trials (Figure 5D and G).

At a population-level, we quantified the effect of the AC-pStr projection inactivation on behavior using the log PLR for Saline to CNO days (see Methods), such that higher values indicate a reduction in animals' ability to reliably predict the time-to-reward from sound. We found that the average log PLR was significantly higher for the 1.5 s sound-reward delay compared to the No-Delay (Figure 5E, $p=0.0019$, Wilcoxon rank-sum test). These ratios did not significantly differ between the 1.5s and 5s delays (Supplementary Figure 2C, $p =0.808$, Wilcoxon rank-sum test). In contrast, mice in the control group showed no significant difference in the average log PLR across No-Delay, 1.5s and 5s delays (Figure 5H, Supplementary Figure 2D, $p=0.1774$, Kruskal-Wallis test). These findings further establish the causal involvement of the AC-pStr projections in animals' ability to predict delayed reward using a sound cue.

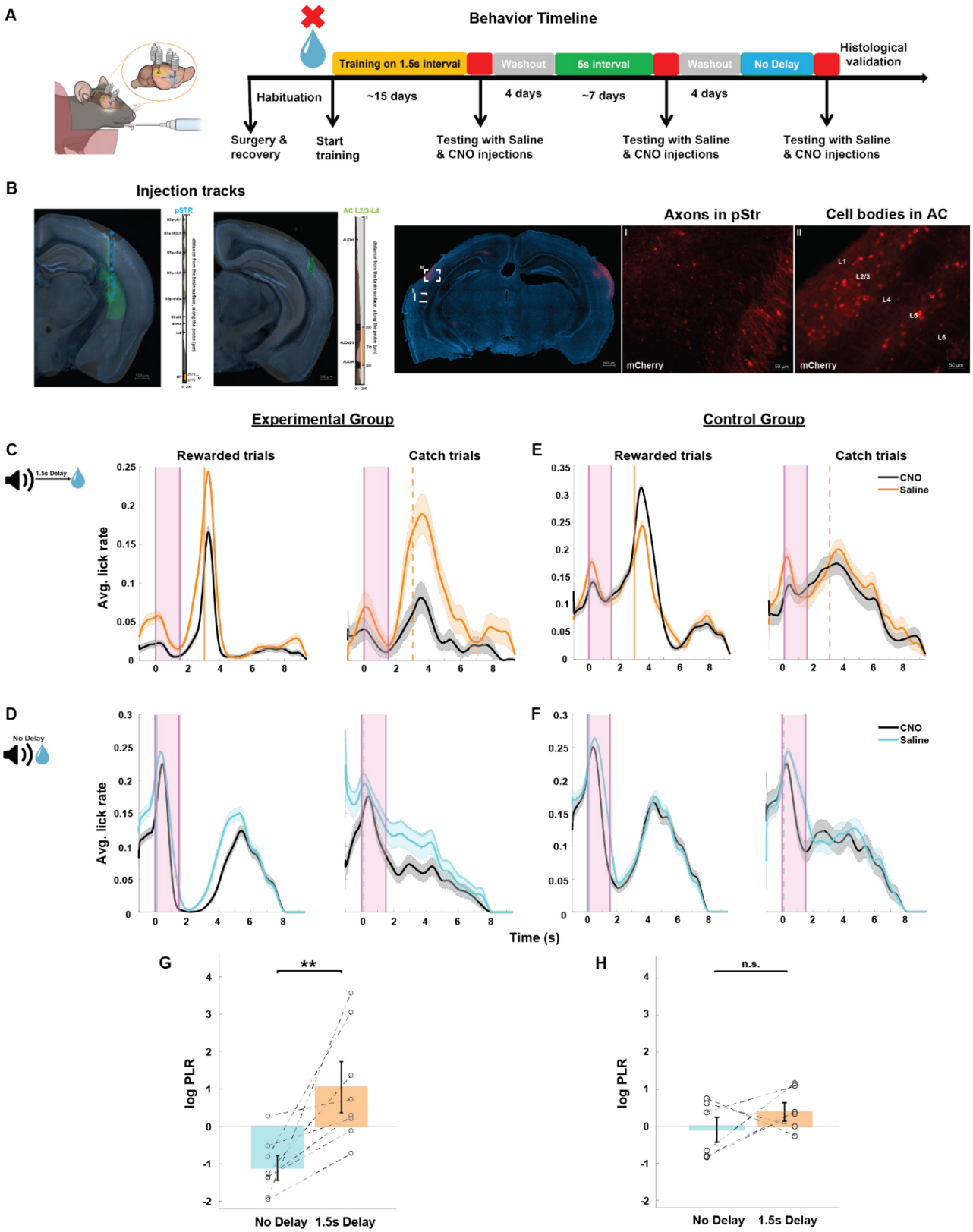

A
Behavior Timeline
Training on 1.5s interval
Washout
5s interval
Washout
No Delay
Histological validation
Habituation
~15 days
4 days
~7 days
4 days
Surgery & recovery
Start training
Testing with Saline & CNO injections
Testing with Saline & CNO injections
Testing with Saline & CNO injections
B
Injection tracks
pSTR
AC L2/3-L4
Axons in pStr
Cell bodies in AC
mCherry
mCherry
L1
L2/3
L4
L5
L6
Experimental Group
Control Group
C
1.5s Delay
Rewarded trials
Catch trials
Avg. lick rate
E
Rewarded trials
Catch trials
CNO
Saline
D
No Delay
F
CNO
Saline
Time (s)
G
**
log PLR
No Delay
1.5s Delay
H
n.s.
log PLR
No Delay
1.5s Delay

Figure 3.5 AC to pStr projections are causally involved in sound-triggered delayed reward prediction. A. Left: An illustration of cannulas implanted in bilateral AC and pStr for virus injections. Right: Behavioral timeline for chemogenetic inactivation experiments. B. Histological verification of selective virus expression in the projections from AC to pStr. The brain slice acquired from an example animal is overlaid with the corresponding coronal section from the Allen Mouse Brain Atlas for both and AC and pStr (see Methods). Left: Virus injection tracks in pStr and AC denoted by green dotted line on the brain slice and the markers on the right indicate the depth at which the electrode was implanted in the right hemisphere. Scale bar: 500μm Right: Magnified images of DREADD virus expression in pStr axons (I) and AC cell bodies (II). Scale bar: 50μm. C. to F. Average peri-sound lick rate response curves (solid line denotes mean, shaded area represents SEM across trials) of an example animal trained to predict reward at either 1.5s sound-reward interval (C and E) or at no delay (D and F) when injected with saline (orange or light blue) and CNO (black) for rewarded trials (left) and catch trials (right). Shaded pink region represents the 1.5s long sound period. Solid and dotted orange or light blue lines represent when reward was given in rewarded trials and expected in catch trials. Left column represents animals from the experimental group and right column represents animals from the control group. G. and H. Average log Predictive Licking Ratio (log PLR) across animals trained on the 1.5s Delay and No-Delay tasks in the experimental group (left, N = 8) and in the control group (right, N=6). Lines connecting the circles represent the log PLR for each animal when trained on the 1.5s Delay and No-Delay tasks. The average log PLR for experimental group animals was significantly higher for 1.5s Delay task than No-Delay task (**p = 0.0019, Wilcoxon rank-sum test) and was not significantly different for the control group animals (p = 0.366, Wilcoxon rank-sum test).

3.4.6 The posterior striatum is involved in sound-guided reward time prediction

We next asked whether the target region of these AC-pStr projections, pStr itself, is a necessary component of the neural circuitry supporting sound-guided reward prediction behavior. Mice were bilaterally implanted with cannulae into their pStr (Figure 3.6A). We compared their ability to predict reward at 1.5s with muscimol infusion to that with PBS infusion on a day prior and found that inactivation of pStr reduced animals' ability to predict the time-delayed reward (Figure 3.6B, $p<0.05$, Wilcoxon rank-sum test). Similar to the previous inactivation experiments, we trained another cohort of mice on the No-Delay task and compared their behavioral performances following PBS and muscimol infusion into pStr. Inactivation of pStr using muscimol did not change animals' behavior on the No-Delay task (Figure 3.6C, $p>0.05$, Wilcoxon rank-sum test), confirming that there was no impairment in animals' ability to process sounds or lick for reward. Across animals, the average log PLR for PBS to muscimol days for the 1.5s Delay task was significantly higher than that of the No-Delay task (Figure 3.6D, $p = 0.0079$, Wilcoxon rank-sum test).

Overall, these findings show that pStr is also required for successful sound-guided reward prediction behavior.

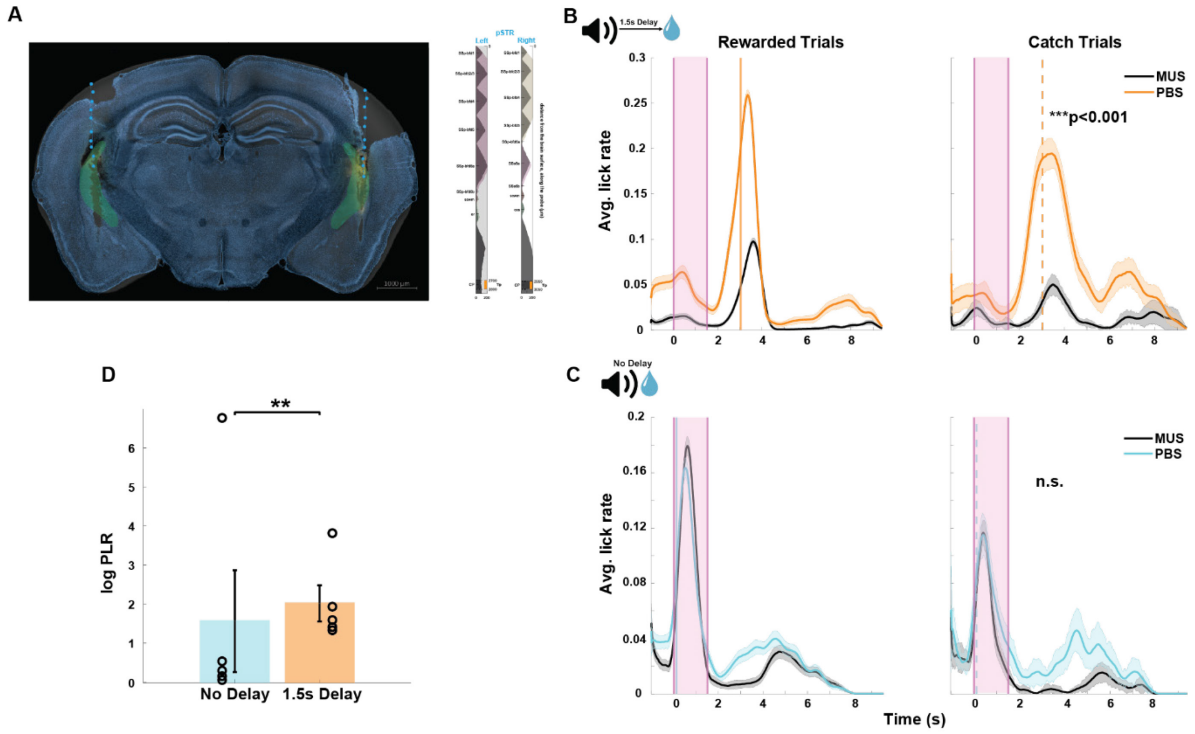


Figure 3.6 pStr is causally involved in sound-triggered delayed reward prediction. A. Histological verification of muscimol labeling in bilateral pStr in a representative animal's brain slice. The brain slice acquired from an example animal is overlaid with the corresponding coronal section from the Allen Mouse Brain Atlas (see Methods). The cannula tracks are indicated by the green dotted line on the brain slice and the site of muscimol infusion is seen in orange. Markers on the right indicate the depth at which muscimol was infused in the left and right hemispheres. Scale bar: 1000µm. B. Average peri-sound lick rate response curves (solid line denotes mean, shaded area represents SEM across trials) of an example animal trained to predict reward at 1.5s sound-reward interval when infused with PBS (orange) and muscimol (MUS, black) in pStr for rewarded trials (left) and catch trials (right). Shaded pink region represents the 1.5s long sound period. Solid and dotted orange lines represent when reward was given in rewarded trials and expected in catch trials. ***p=0.000007(Wilcoxon rank-sum test) denotes the significant difference in predictive licking between catch trials on PBS and MUS. C. Average peri-sound lick rate response curves (solid line denotes mean, shaded area represents SEM across trials) of an example animal trained on the No-Delay task when infused with PBS (light blue) and muscimol (MUS, black) in pStr for rewarded trials (left) and catch trials (right). Shaded pink region represents the 1.5s long sound period. Solid and dotted blue lines represent when reward was given in rewarded trials and expected in catch trials. n.s. (p=0.3, Wilcoxon rank-sum test) denotes the not significant difference in predictive licking between catch trials on PBS and MUS in the No-Delay task. D. Significant difference in average log predictive licking ratio (log PLR) between No-Delay (N = 5) and 1.5s Delay (N = 5) cohorts (**p =0.0079, Wilcoxon rank-sum test). Error bars represent mean ± SEM across animals. Open circles indicate the PLR for each animal in the No-Delay (blue) and 1.5s Delay cohorts.

3.4.7 Coordination of sound-evoked responses in AC and pStr during sound-triggered reward time prediction

To determine the activity patterns in pStr and to test whether the AC and pStr activity is coordinated during sound-triggered reward time prediction behavior, we simultaneously recorded LFP activity in AC and pStr as animals learnt to predict sound-triggered reward timing at the four sound-reward time intervals (Figure 7A). Like we did for analyzing AC responses, we eliminated trials with movement and included only trials with no licks in the first 200ms from sound onset for the best behavior days on each sound-reward interval. As expected from previous studies (Guo et al., 2018, 2019), we found robust sound responses in pStr, albeit of lower magnitude than in the AC. When we compared the average pStr sound responses across all four sound-reward intervals, we found that pStr responses also tended to increase with sound-reward interval duration (Figure 7B). However, at the population level, we noticed that while the pStr response magnitude was significantly different across sound-reward intervals (Figure 7C, $p = 8.96\text{x}10\text{-}6$, Kruskal-Wallis test), the pair-wise comparison of the response magnitude across all pairs of sound-reward intervals did not yield significant differences, unlike responses in the AC.

Next, we asked whether neural activity across the two brain regions was coordinated during sound-guided reward time prediction behavior. We found that AC and pStr LFP response magnitudes showed a significant positive trial-by-trial correlation for each behavioral session (Figure 7D) and the average correlation coefficients across animals did not significantly differ across sound-reward intervals (Figure 7E, $p = 0.99$,

Kruskal-Wallis test). These results suggest the existence of strong coordination across the AC and pStr during sound processing within our task.

pStr receives direct projections from the AC, but both the pStr and AC receive direct projections from the medial geniculate body (Chen et al., 2019; Huerta-Ocampo et al., 2014; Hunnicutt et al., 2016; Smeal et al., 2008). We next sought to test to what extent this anatomical projection pattern is reflected in the temporal relationship of AC and pStr sound-evoked responses. To this end, we calculated trial-wise cross-correlations of LFP responses in AC and pStr over a 2 s sound period, from 0.5s prior to sound onset till sound termination (Figure 7F and G). The trial-averaged cross-correlation showed a clear peak at 0 lag, indicating strong synchrony in activity across the brain regions, as expected from the shared common input from the MGB (Figure 7G, black trace). However, we noticed that the shape of these average cross-correlation curves was not symmetric around 0. To test this, we compared them to cross-correlations generated after randomly shuffling the AC/pStr identity of the traces (see Methods). This form of shuffling generated cross correlations with no temporal directionality across the brain regions by design (Figure 7G, red trace). In comparison to these shuffled cross-correlations, the real cross correlations showed higher values at positive lags, indicating a temporal lead of AC relative to pStr (Figure 7G). To quantify this, we calculated the difference traces of the real and shuffled cross correlations (Figure 7H). Across the data, the median peak time of these difference traces was 38 ms and was significantly different from 0 s (Figure 7I, Wilcoxon signed rank test, $p = 0.0084$). Further, these peak times of difference traces were not significantly different across sound-reward intervals (Figure 7J, $p = 0.94$, Kruskal-Wallis test). This positive median peak time indicates that in addition to the no-

lag synchrony across these brain regions, AC leads pStr during the sound predicting the time to reward in our task.

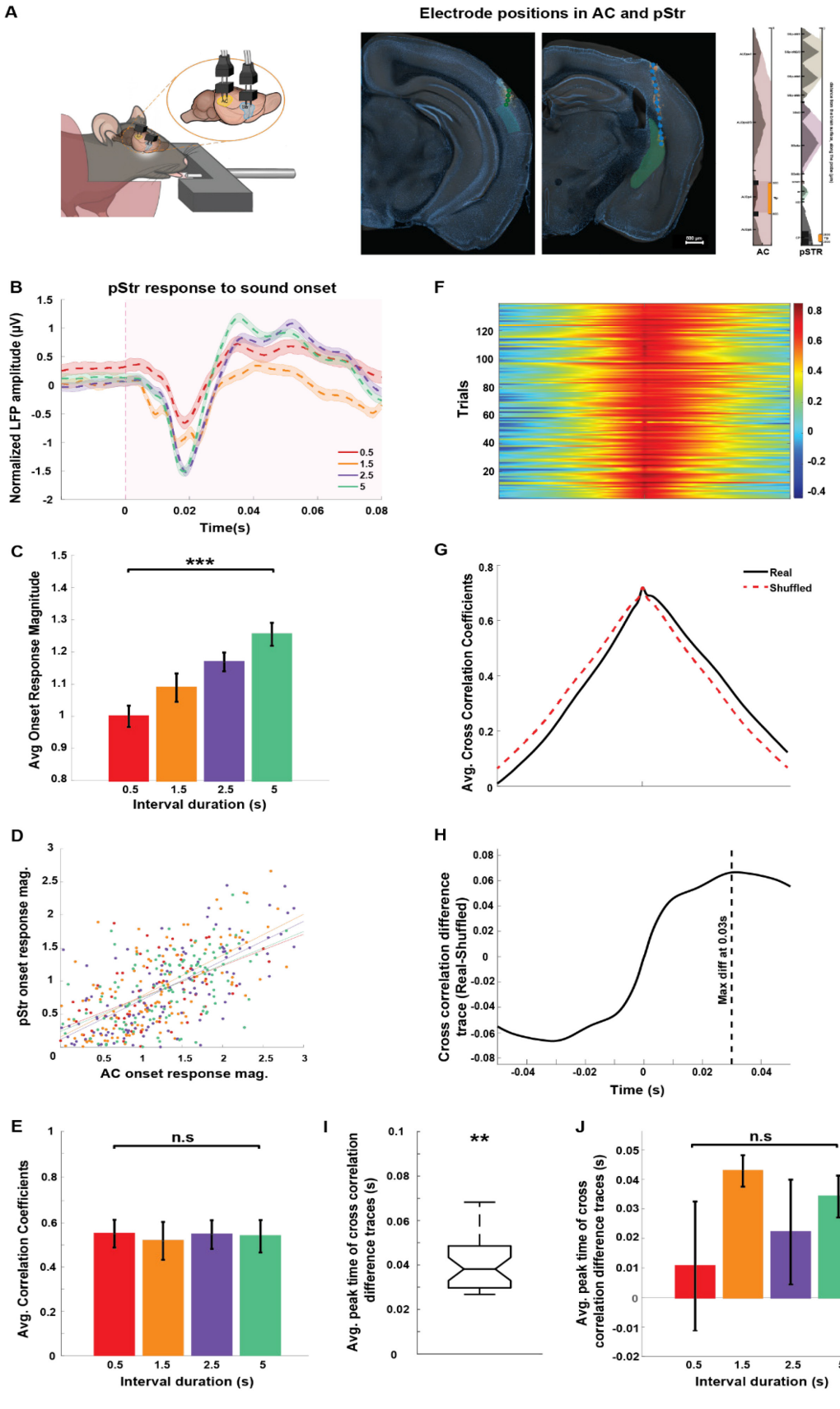

A
Electrode positions in AC and pStr
AC
pSTR
B
pStr response to sound onset
Normalized LFP amplitude (μV)
Time(s)
0.5
1.5
2.5
5
F
Trials
C
***
Avg Onset Response Magnitude
Interval duration (s)
G
Real
Shuffled
Avg. Cross Correlation Coefficients
D
pStr onset response mag.
AC onset response mag.
H
Cross correlation difference trace (Real-Shuffled)
Max diff at 0.03s
Time (s)
E
n.s
Avg. Correlation Coefficients
Interval duration (s)
I
**
Avg. peak time of cross correlation difference traces (s)
J
n.s
Avg. peak time of cross correlation difference traces (s)
Interval duration (s)

Figure 3.7: Coordination of LFP activity across AC and pStr during sound-triggered reward time prediction. A. Left: Illustration of electrode implanted in right AC and pStr. Right: Histological verification of electrode positions in AC (left) and pStr (right). The brain slice acquired from an example animal is overlaid with the corresponding coronal section from the Allen Mouse Brain Atlas (see Methods). The electrode tracks are indicated by the green dotted line on the brain slices and the markers in the middle indicate the depth at which the electrode was implanted in the right hemisphere. Scale bar: 500μm. B. Normalized average pStr LFP (solid line denotes mean, shaded area represents SEM across no-lick trials) recorded in response to the sound onset from an example animal trained on the four different sound-reward intervals (represented by the different colors). The shaded pink region represents the period from sound onset. C. Average pStr onset response magnitude computed across animals (N = 8) for each of the sound-reward intervals. Error bars represent mean ± SEM across animals. Comparison across sound-reward intervals yields ***$p = 8.96x10^{-6}$ (Kruskal-Wallis test). D. Scatter plot of the trial-wise correlation between AC and pStr onset response magnitudes for each of the sound-reward intervals for an example animal. Filled circles of different colors represent the trial-wise onset response magnitudes and the solid lines represent the linear fits for each of the sound-reward intervals. Pearson correlation coefficients for each of the sound-reward intervals was positive and significant: 0.5s interval (red): $r = 0.582$, $p<0.001$; 1.5s interval (orange): $r = 0.673$, $p<0.001$; 2.5s interval (purple): $r = 0.673$, $p<0.001$; 5s interval (green): $r = 0.561$, $p<0.001$. E. The average correlation coefficients across animals across sound-reward intervals are not significantly different ($p = 0.99$, Kruskal-Wallis test). F. Heat map representing the trial-wise cross-correlation of AC and pStr sound-evoked LFP responses for an example animal trained on an 0.5s sound-reward interval behavioral session. Time axis is set such that it denotes AC leading on the positive side and pStr leading on the negative side. Color bar indicates the cross-correlation coefficients. G. Average cross correlation of AC and pStr sound evoked LFP responses across trials for the behavioral session shown above is indicated in black. The dotted red line shows the average of the shuffled cross correlation computed from randomized AC and pStr sound evoked LFP responses from the session. H. The solid black line shows the difference trace between the real and shuffled average cross correlation in panel G. This cross-correlation difference trace has a peak at 0.03s as indicated by the dotted black line. I. The average of the peak times of the cross-correlation difference traces computed across all animals and across all sound-reward intervals is significantly greater than 0s (**$p= 0.0084$, Wilcoxon signed rank test). J. Average peak times of the cross-correlation difference traces across all animals was not significantly different across sound-reward intervals ($p = 0.94$, Kruskal-Wallis test).

3.5 Discussion

In this study, we established a novel behavioral paradigm based on an extension of the classical appetitive trace-conditioning task to assess sound-triggered reward time prediction in mice. Using this behavioral paradigm, we found that mice can use a sound cue to reliably predict time intervals to reward at a 1-second temporal resolution and that the ability to predict delayed reward is dependent on AC. We further found that mice use the sound onset to estimate the time to reward and that the AC LFP responses to sound onset predicted the amount of time to reward. As a downstream pathway to non-auditory brain areas, we found that the neural projections from AC to pStr, as well as the pStr itself, are necessary for sound-triggered delayed reward prediction. Sound responses in pStr also varied based on the time to reward. Finally, using simultaneous recordings in AC and pStr during performance of this task, we found strong coordination of neural activity across these brain regions, with synchronous, as well as AC-leading components of cross-correlation. Together, our findings identify AC-pStr mechanisms for sound-triggered prediction of reward timing.

Decades of research has yielded several competing models for how the brain represents time. The most explored theory underlying the timing mechanisms in animals and humans assumes the existence of an internal clock based on neural counting (Fung et al., 2021; Hinton & Meck, 1997; Leow & Grahn, 2014; Treisman, 1963; Wearden, 2005). According to this theory, higher order “centralized” clock brain regions receive inputs from various modalities and maintain a representation of time. On the other hand, parallel work in this domain supports the existence of multiple timing mechanisms distributed across different brain regions and circuits that are engaged based on the task

design, sensory modality used, and the temporal resolution of the task (Bueti, 2011; Jazayeri & Shadlen, 2015; Tallot & Doyère, 2020; Matell, et al., 2011). Our findings that sound-triggered reward time prediction ability in mice at a 1-second temporal resolution is dependent on AC, strongly supports the distributed modality-specific timing model. By pharmacologically inactivating AC, we show an impairment in this ability to predict delayed reward time based on a sound cue. Further, we found that the magnitude of AC responses to the sound cue in this task encode and maintain a neural representation of the time from the sound onset to reward. These results provide evidence for an auditory-modality specific timing mechanism for sound-triggered time estimation to an imminent salient event.

Previous studies have identified different forms of neurophysiological signatures of interval timing across various brain regions (Tallot & Doyère, 2020). A group of studies in rodent visual cortex (Hussain Shuler & Bear, 2006; Namboodiri et al., 2015), basal amygdala (Pendyam et al., 2013) and basal ganglia (Hikosaka et al., 1989) have found coding of interval timing via sustained increase or decrease in spiking activity from cue onset till upcoming salient event (reward or foot shock). Other studies, primarily in the prefrontal cortex (PFC) and dorsal striatum, have found gradual firing rate ramping up as the expectation of the animal for an upcoming salient stimulus increases (Armony et al., 1998; Matell et al., 2003; Narayanan & Laubach, 2009). Another set of studies have found encoding of time intervals by phasic increases in neural activity at the time of anticipated reward in neurons in the PFC (Yumoto et al., 2011), dopaminergic neurons in the ventral tegmental area (Fiorillo et al., 2008) and in the visual cortex (Hussain Shuler & Bear, 2006). A common factor across most of these studies is that the amount of time to the

salient event is encoded separately from the sensory stimulus that triggers its anticipation. In contrast, we find that the magnitude of the AC responses to the sound itself also encode the anticipated time to reward. Thus, the cue and the reward-timing prediction associated with it are jointly coded in the AC. This form of integration between sound and the temporal expectation associated with it is consistent with a previous study, in which the AC responses to a tone were found to be different when rats expected a target sound to be presented either immediately after the tone or about 1 second later (Jaramillo & Zador, 2011). Our findings extend and generalize these results to demonstrate the existence of timing prediction of non-auditory cues within the AC, and that interval timing duration is represented in a continuous manner within the scale of 0.5-5s.

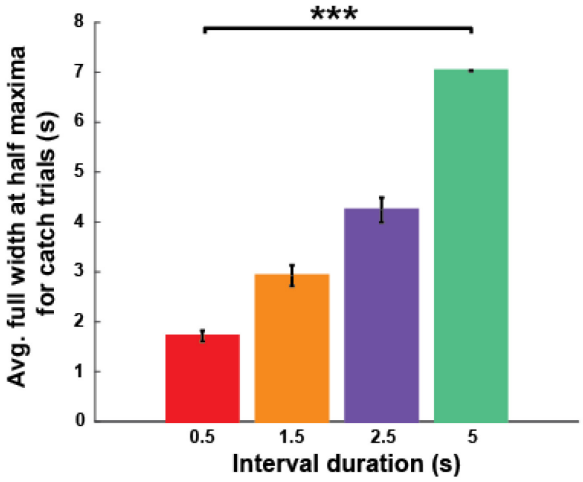
To investigate how the auditory corticostriatal projections are involved in sound-triggered time estimation behavior, we used chemogenetic inactivation to establish the necessity of the anterograde projections from AC to pStr in this task. We acquired simultaneous LFP recordings from AC and pStr to show that neural activity in these regions is temporally synchronized during the sound cue predicting the interval timing to reward. The activity in AC and pStr was highly synchronized at 0s lag, which is consistent with the common input received by both these regions from the medial geniculate body (MGB) (Chen et al., 2019; Huerta-Ocampo et al., 2014; Hunnicutt et al., 2016; Smeal et al., 2008). In the context of our task, this would suggest that the sound-triggered reward timing encoded in AC and in pStr could be simultaneously and differentially modulated by the neural inputs from MGB. In addition to the direct projections from MGB, pStr receives strong monosynaptic projections from AC (Znamenskiy and Zador, 2013; Xiong et al., 2015; Huang et al., 2023). This is consistent with our finding that in addition to the

synchronous activity across these brain regions, a second component of the cross-correlograms points at AC leading the pStr by 38ms on average. Our LFP recordings in AC and pStr provided a quantification of the overall functional and temporal communication patterns across the AC and pStr. Future studies using single unit recordings with cell-type specificity could elucidate how these phenomena are represented at the local ensemble level.

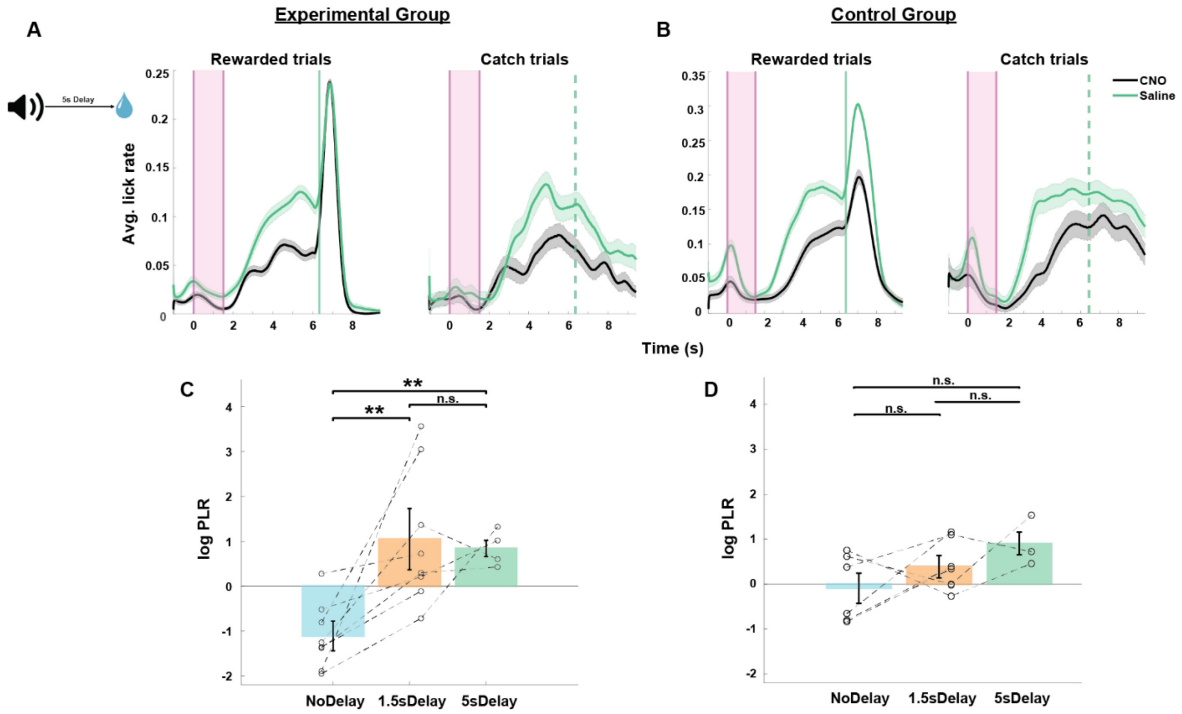
While sound-triggered action timing on the scale of seconds is important for many everyday behaviors, it is a particularly critical ability in humans for verbal communication. In verbal communication, humans use incoming speech sounds to predict when and what sounds are expected to follow (Heinks-Maldonado et al., 2005, 2006; Nixon & Tomaschek, 2021; Tremblay et al., 2013). Furthermore, the sequential nature of speech sounds helps arrange incoming information in a temporal structure and this ability is impaired when there are hearing deficits (Füllgrabe, 2013; Grose & Mamo, 2010; Helfer & Jesse, 2021; Ozmeral et al., 2016; Tallal et al., 1995). Deaf or hard of hearing children face challenges in sequential time perception, which impairs their storytelling ability, a key cognitive development factor (Eden & Leibovitz-Ganon, 2022). In comparison, children with postlingual cochlear implants exhibit greater improvements in time perception and consequently, their storytelling ability, emphasizing the role of hearing acquisition in sound-guided temporal processing. Other research involving children with mild hearing loss have showcased the benefits of interventions on time sequencing and storytelling (Ingber & Eden, 2011). These studies collectively underscore the critical role of sound-guided timekeeping in developing and maintaining language-related skills and highlight

the need to further study the neural mechanisms underlying these processes to develop tailored support for individuals with hearing impairments.

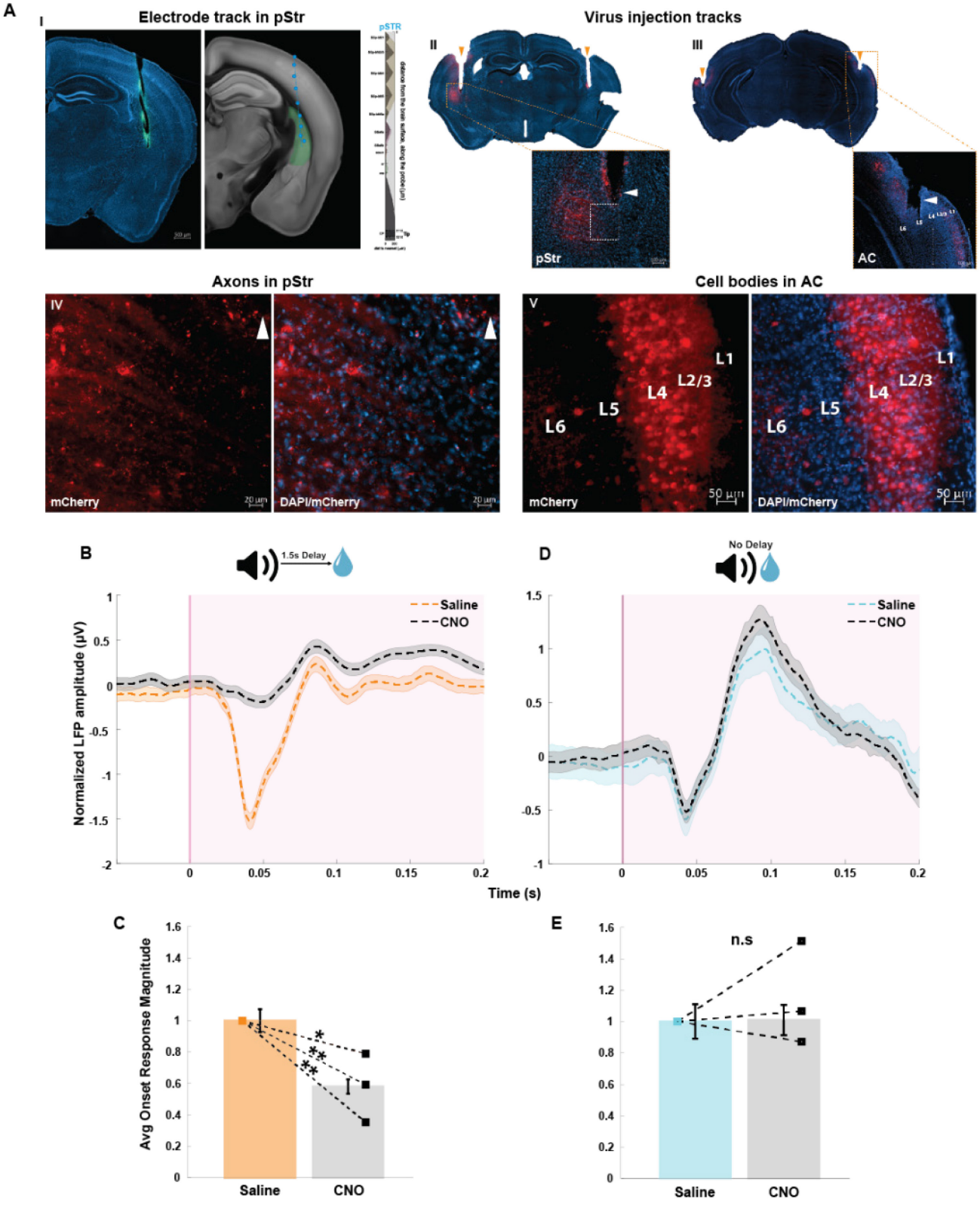
3.6 Supplementary Figures



Supplementary Figure 3.1: Precision of predictive licking ability across sound-reward intervals. The average full-width at half maxima of the predictive licking curves across animals was significantly different across sound-reward intervals (***$p = 3.85 \times 10^{-6}$, Kruskal-Wallis test).



Supplementary Figure 3.2. Effect of chemogenetic inactivation of AC to pStr projections on the 5s Delay task. A. and B. Average peri-sound lick rate response curves (solid line denotes mean, shaded area represents SEM across trials) of an example animal trained to predict reward at 5s sound-reward interval when injected with saline (green) and CNO (black) for rewarded trials (left) and catch trials (right). Shaded pink region represents the 1.5s long sound period. Solid and dotted green lines represent when reward was given in rewarded trials and expected in catch trials. Left column represents animal from the experimental group and right column represents animals from the control group. C. and D. Average log Predictive Licking Ratio (log PLR) across animals trained on the 1.5s Delay, 5s Delay and No-Delay tasks in the experimental group and in the control group reiterated from Fig. 5E and H. Lines connecting the circles represent the PLR for each animal when trained on the 1.5s Delay, 5s Delay and No-Delay tasks. The average PLR for experimental group animals was significantly higher for 5s Delay task than No-Delay task (**p = 0.004, Wilcoxon rank-sum test) and was not significantly different for the control group animals (p = 0.167, Wilcoxon rank-sum test).


A
I
Electrode track in pStr
pSTR
II
Virus injection tracks
III
pStr
AC
Axons in pStr
Cell bodies in AC
IV
V
mCherry
DAPI/mCherry
L1
L2/3
L4
L5
L6
50 µm
B
1.5s Delay
Saline
CNO
Normalized LFP amplitude (µV)
Time (s)
D
No Delay
C
Avg Onset Response Magnitude
E
n.s

Supplementary Figure 3.3 Histological verification of targeting AC to pStr projections and change in pStr LFP responses following chemogenetic inactivation of these projections. A. Histological validation of electrode position in pStr and chemogenetic virus expression in AC to pStr projections. I. Electrode position in pStr is denoted by the green dotted lines. II. and III. Representation of the virus injection tracks in AC and pStr. IV. and V. are magnified images from II and III showing axons in pStr and cell bodies in AC. B.and D. Normalized average pStr LFP (solid line denotes mean, shaded area represents SEM across no-lick trials) recorded in response to the sound onset from an example animal trained on the 1.5s Delay task (left) and on the No-Delay task (right) with saline (orange/light blue) and CNO (black) injections. The shaded pink region represents the period from sound onset. C. and E. Comparison of the normalized average pStr onset response magnitude computed across animals (N = 3) for the 1.5s Delay task (Fig C) and for the No-Delay task (Fig E) on saline and CNO conditions. Error bars represent mean ± SEM across animals. Comparison between saline and CNO conditions yields *$p<0.05$ (Wilcoxon rank-sum test) for each animal when trained on the 1.5s Delay task and was not significantly different when trained on the No-Delay task.

Chapter 4 : Discussion

4.1 Summary

Understanding day-to-day sounds in our environment like speech or music, requires recognition and integration of acoustic features of the sound over a range of time periods, along with the context in which they occur. Addressing how our auditory system facilitates such robust auditory perception has been a fundamental question of research. As the major higher level sound processing region in the auditory pathway, the auditory cortex (AC) has been extensively studied to investigate its role in auditory perception and it has been shown that AC, unlike other primary sensory cortices like the visual cortex, is not a simple detector of acoustic features of the incoming sound signals but also integrates information about the context and behavior (Kuchibhotla & Bathellier, 2018; Nelken et al., 2003, 2014; Nelken & Bar-Yosef, 2009; Sutter & Shamma, 2011; Town et al., 2018). Various neurophysiological signatures have been identified in AC to support its multidimensional role in auditory perception.

This dissertation outlines two distinct auditory cortical neural mechanisms underlying auditory driven behaviors -

In study 1, I found that AC neuronal representations of behaviorally relevant complex sounds are more stable than that of simple stimuli like pure tones. These results demonstrate that sounds with a distinct identity and potential meaning are more stably represented in AC across time.

In study 2, I used an appetitive sound-triggered reward time prediction task to establish that mice can use a sound cue to predict time intervals at 1-second temporal resolution, and that this behavior is dependent on AC. I also found that AC LFP responses encode for the interval duration from sound onset to reward in this task and coordinate this information with the posterior striatum, a downstream non-auditory region. These findings provide the first evidence of AC encoding for information about the sensory cue and timing, a non-auditory component, simultaneously and make an essential contribution about the role of AC in processing aspects beyond the auditory stimuli.

Together, both these studies provide novel evidence that sound representations in AC carry information about sound-guided predictive coding, behavioral context, and subsequent actions, and further our understanding of the role of AC in achieving stable and highly robust auditory perception.

4.2 Multifaceted role of AC

The findings outlined in this dissertation challenges the notion that AC is only involved in simple computations about the incoming sounds while suggesting that AC could be involved in encoding for sounds and all sound-related non-auditory information necessary for achieving a stable auditory perception. As it has been theorized in past works, this dissertation allows us to further ask questions about the distinctive function of the primary AC in the auditory pathway and about how the primary AC is not equivalent to other primary sensory cortices like the primary visual cortex. Moreover, given the diversity and the range of functions that AC has been implicated in, it would not be

farfetched to speculate whether AC performs the role of a higher-order associative region like the prefrontal cortex or the hippocampus.

4.2.1 Exploring the link between AC representational stability and perception

Stability of perception and behavior necessitates some form of neuronal representational stability. The highly dynamic cortical activity in response to sensory inputs, which is also known to change over time, influences how such a representational stability can exist. Such changes in neural representations, also referred to as the "representational drift", has been observed in olfactory bulb (Schoonover et al., 2021), visual (Deitch et al., 2021; Pérez-Ortega et al., 2021; Ranson, 2017), somatosensory (Pancholi et al., 2023; Peron et al., 2015), auditory (Aschauer et al., 2022), motor (Rokni et al., 2007) and hippocampal (Ziv et al., 2013) regions. Many of these studies show a significant proportion of neurons responding similarly across time, while others contribute to the necessary plasticity required for adapting to changes in the environment. Recently, a study in the mouse vibrissal somatosensory cortex showed that greater neural stability was associated with stimulus response during a perceptual learning task, indicating that representational stability of a stimulus evolves not only with time but is strongly influenced by the behavioral relevance of the stimulus to the animal (Pancholi et al., 2023). The findings in my study 1 provide evidence for existence of representational stability in the auditory modality but also suggest that such a representational stability could be correlated with behavioral salience as well, since AC responses are more stable across days to complex sounds compared to pure tones.

Previous studies have shown that there is more to AC than processing a sound's acoustic structure (Kuchibhotla & Bathellier, 2018; Nelken et al., 2003, 2014; Nelken &

Bar-Yosef, 2009; Sutter & Shamma, 2011; Town et al., 2018). Consistent with these studies, results from both my studies showed that AC representation of spectrotemporally complex sounds and in a behavioral context could be modulated by their ethological relevance. This ethological relevance may either be innate through evolution or gained significance through behavioral conditioning. These findings underscore the need to investigate AC neural mechanisms using complex sounds and contexts that extend beyond the conventional use of simple auditory stimuli such as pure tones. This expansion is vital because the long-term stability of sound representations in AC may be directly linked to its capacity to maintain perceptual stability (Lütcke et al., 2013; Sutter & Shamma, 2011). For example, electrophysiological recordings from the mouse AC have revealed that the ability to discriminate between complex sounds improves over time, while the discriminability between pure tones remains unchanged (Maor et al., 2020). Similarly, studies involving non-human primates and human subjects have shown that AC representation of meaningful vocalizations is retained longer than that of non-vocalization complex sounds (Ng et al., 2009). Building on these studies and the findings from study 1, it can be inferred that enhanced stability of AC representations of complex sounds compared to pure tones would predict better behavioral performance on learning and memory tasks using complex sounds rather than pure tones. For example, performance on an auditory recognition task with long time intervals (e.g. hours or days) between sound presentations would be expected to be better for complex sounds compared to pure tones. Whether this is indeed the case, and whether neuronal representational stability and memory-related behavioral performance are indeed linked remains to be addressed.

Another way in which future studies may attempt to identify whether longitudinal stability of AC sound representations is associated with the ethological relevance of the sound is by comparing AC sound responses to pure tones, ethologically relevant complex sounds and to wideband, ethological irrelevant sounds like tone clouds (Bigelow et al., 2019; Chen et al., 2019; Francis et al., 2018; Gilday & Mizrahi, 2023; M. Wang et al., 2020; Znamenskiy & Zador, 2013). If a similar degree of daily plasticity is observed between tone clouds and pure tones, it would suggest that the ethological relevance of the sound is the modulating factor for maintaining representational stability in AC, thereby implicating it in the preservation of perceptual stability.

4.2.2 AC representations encoding predictive non-auditory information

Cortical representations of prediction have largely focused on retrospective coding of sensory information. The most intensively studied phenomenon of stimulus-specific adaptation (SSA), in which neuronal responses to familiar repeated stimuli are reduced while responses to rare, deviant ones are enhanced (Ulanovsky et al., 2003, 2004), is a good example of such predictive coding. This phenomenon of expectation-induced reduction in neural responses have also been observed in the visual cortex (Summerfield & De Lange, 2014) and somatosensory cortex (Musall et al., 2017). Most importantly for this dissertation, temporal expectation of a salient sound cue is known to evoke an enhanced response in AC neurons. This representation of temporal expectation has been suggested to cause improvement in associated behavioral responses in rodents and humans (Jaramillo & Zador, 2011; Stefanics et al., 2010). Other studies have shown the salient sensory consequences of an action like a lever press have also been encoded in AC (Cook et al., 2022). While these studies have given us insights about AC

representation of predictive coding based on sensory experiences or actions, findings in study 2 of this dissertation, provides evidence for prospective predictive coding of future salient event based on a sound cue in AC. I show that AC responses are able to simultaneously encode for sound and reward timing, with the AC sound onset-response magnitude increasing with sound-reward interval duration. A potential explanation for this trend could be based on an already known mechanism in AC, which tells us that AC responses are smaller in magnitude to more predictable and reliable sounds compared to its responses to surprise, deviant sounds (Ulanovsky et al., 2003, 2004). Based on my results that animals' show higher precision in predicting shorter interval durations (Supplementary Figure 3.1), this would suggest that the sound cue associated with the shorter interval durations would be more predictable and reliable and hence, evoke a smaller response in AC compared to the responses to the sound associated with the longer intervals, which are known to have less precision in behavioral prediction.

As past work has shown that AC responses encode for non-auditory aspects like sound-associated movement (Henschke et al., 2021; Schneider et al., 2014; Vivaldo et al., 2023), arousal (P. A. Lin et al., 2019; Schwartz et al., 2020), and other sensory stimuli (Bizley et al., 2007; Cohen et al., 2011; Ghazanfar et al., 2005; Gilday & Mizrahi, 2023; Kayser et al., 2005), the simultaneous encoding of cue and predicted reward timing in AC LFP responses in this study adds another dimensionality to AC's ability to represent information beyond auditory stimuli. However, we are yet to determine whether this ability of sound-triggered prediction of reward timing in AC extends beyond the temporal scale of few seconds and how it interacts with AC encoding of other non-auditory components. For instance, in the context of my task, attention/alertness of the animal to the sound and

other cues in its environment could very well influence how AC LFP responses encode the interval duration from sound to upcoming reward and consequently, the behavioral performance on the task. Moreover, as described above, the influence of ethological relevance of the sounds guiding such behaviors could also modulate the precision of timing incorporated in these predictive representations in AC. One way to test this would be to estimate the attention of the animal by recording its pupil diameter (P. A. Lin et al., 2019; Schwartz et al., 2020) during performance on this task across varying sound-reward intervals and checking if the change in pupil size varies within the different components of the trial, including the sound, the interval period, and the reward and whether this variation is correlated with different interval durations.

4.3 Technical considerations

Both studies conducted within this dissertation represent significant steps in comprehending the neural mechanisms within AC through distinctive approaches. Two-photon calcium imaging allows for cell-type specific recording of the same neurons over long periods of time with spatial precision (Grienberger & Konnerth, 2012; Stosiek et al., 2003; Svoboda & Yasuda, 2006). Importantly, this technique has been used to address the long-term stability of neural representations in sensory cortices like visual cortex (Deitch et al., 2021; Ranson, 2017), olfactory bulb (Schoonover et al., 2021), and somatosensory cortex (Peron et al., 2015) and in the hippocampus (Ziv et al., 2013). In Study 1, I harnessed two-photon calcium imaging to monitor the activity of individual excitatory neurons in the AC, focusing on their responses to complex sounds and pure tones across days. This allowed me to compare representation of sounds in AC based on their ethological relevance and acoustic structure under baseline conditions across

days. However, this study only reflects the beginning of how we can exploit this tool to record cell-type specific activity under baseline and experimental conditions to understand local network dynamics. One question to ask would be how the excitatory and inhibitory neurons across cortical layers of AC interact to achieve the intracortical network dynamics required to facilitate auditory behaviors. A recent study targeted excitatory and PV-expressing interneurons in the dorsal auditory field using two-photon calcium imaging and showed how activity of these neurons are differentially modulated by locomotion (Henschke et al., 2021). Further, with the advent of cortical layer-specific neuronal targeting using two-photon microscopy (Clayton et al., 2021; Romero et al., 2020; Tischbirek et al., 2019), it would be much easier to understand how representations are modulated by the neurons in different cortical layers and the neural inputs they receive from other regions.

In contrast to this approach, in study 2, I acquired LFP recordings to examine AC and pStr responses to behaviorally relevant sounds over much shorter intervals, spanning seconds. In the context of this study, it was imperative that I recorded neural activity using a technique with high temporal resolution and that which allowed me to compare activity in AC and pStr simultaneously within the same hemisphere. In vivo electrophysiology provided these advantages over two-photon calcium imaging and hence, was more suitable for the purpose of this study. Moreover, previous studies particularly implicated LFP neural signatures like oscillations and coherence when multiple brain regions communicated and coordinated to facilitate interval timekeeping (Buhusi & Meck, 2005; Tallot & Doyère, 2020; Treisman, 1963). While LFP recordings in AC and pStr certainly gave us essential insights about their role in sound-triggered prediction of reward timing,

acquiring single-unit activity from both these regions, as the next step, would enable us to test how neural responses in these regions are modulated by different aspects of the sound to timed action sequence of behavior.

In addition to using two key neurophysiological recording techniques, Study 2 also chemogenetically targeted the anterograde projections from AC to pStr while simultaneously acquiring LFP recordings from both these brain regions. The results from the specific experiment employing these techniques, established the causal and functional consequences of inactivating AC to pStr projections on sound-triggered reward timing task. But this was only explored in a subset of animals using chemogenetics, a global inactivation approach, with effects lasting for longer periods of time. To determine how neural information is integrated and communicated from AC to pStr during this task would require using a combination of optogenetic inactivation of these projections and single unit recordings in AC and pStr, such as optrodes (For example, Weible et al., 2014; Guo et al., 2015; O'sullivan et al., 2019). This would allow for inactivating the communication from AC to pStr during specific parts of the trials and examining when during the trial is the information about timing in AC, translated to inform the initiation of the consequent action in pStr. Such studies have been conducted in mice to understand the encoding of short-term auditory memory in AC (Yu et al., 2021).

4.4 Conclusion

The diverse approaches utilized in this dissertation, along with their distinct outcomes, not only underscore the vast expanse of research opportunities within the realm of auditory perception but also highlight the profound enigma that AC continues to

represent in this landscape. Applying similar disparate methods to address AC neurophysiological mechanisms that integrate and interpret complex sounds in various behavioral contexts, could provide a more comprehensive understanding of how they can be modulated to help patients with hearing impairments and patients with motor-related disorders with consequential speech and communication difficulties like dysarthria and apraxia (Kent, 2000). These insights may prove invaluable in developing invasive and non-invasive therapeutic interventions for patients who struggle with planning and sequencing speech sounds (Kent, 2000; Maas et al., 2008; Strand, 1995), and more importantly, for patients who have impaired cognitive abilities due to age-related hearing deficits (Jafari et al., 2021; Uchida et al., 2019). Therefore, the culmination of these studies not only enriches our understanding of AC functioning but also paves the way for diverse applications, particularly in areas aimed at enhancing human communication and overcoming auditory challenges.

References

Aizenberg, M., & Geffen, M. N. (2013). Bidirectional effects of aversive learning on perceptual acuity are mediated by the sensory cortex. Nature Neuroscience, 16(8), 994–996. https://doi.org/10.1038/nn.3443

Aizenberg, M., Mwilambwe-Tshilobo, L., Briguglio, J. J., Natan, R. G., & Geffen, M. N. (2015). Bidirectional Regulation of Innate and Learned Behaviors That Rely on Frequency Discrimination by Cortical Inhibitory Neurons. PLoS Biology, 13(12), 1–32. https://doi.org/10.1371/journal.pbio.1002308

Angeloni, C., & Geffen, M. N. (2018). Contextual modulation of sound processing in the auditory cortex. Current Opinion in Neurobiology, 49, 8–15. https://doi.org/10.1016/j.conb.2017.10.012

Armony, J. L., Quirk, G. J., & Ledoux, J. E. (1998). Differential effects of amygdala lesions on early and late plastic components of auditory cortex spike trains during fear conditioning. Journal of Neuroscience, 18(7), 2592–2601. https://doi.org/10.1523/jneurosci.18-07-02592.1998

Aschauer, D. F., Eppler, J. B., Ewig, L., Chambers, A. R., Pokorny, C., Kaschube, M., & Rumpel, S. (2022). Learning-induced biases in the ongoing dynamics of sensory representations predict stimulus generalization. Cell Reports, 38(6), 110340. https://doi.org/10.1016/j.celrep.2022.110340

Atencio, C. A., Sharpee, T. O., & Schreiner, C. E. (2008). Cooperative Nonlinearities in Auditory Cortical Neurons. Neuron, 58(6), 956–966. https://doi.org/10.1016/j.neuron.2008.04.026

Atencio, C. A., Sharpee, T. O., & Schreiner, C. E. (2012). Receptive field dimensionality increases from the auditory midbrain to cortex. Journal of Neurophysiology, 107(10), 2594–2603. https://doi.org/10.1152/jn.01025.2011

Atiani, S., Elhilali, M., David, S. V., Fritz, J. B., & Shamma, S. A. (2009). Task Difficulty and Performance Induce Diverse Adaptive Patterns in Gain and Shape of Primary Auditory Cortical Receptive Fields. Neuron, 61(3), 467–480. https://doi.org/10.1016/j.neuron.2008.12.027

Audette, N. J., Zhou, W., La Chioma, A., & Schneider, D. M. (2022). Precise movement-based predictions in the mouse auditory cortex. Current Biology, 32(22), 4925-4940.e6. https://doi.org/10.1016/j.cub.2022.09.064

Bajo, V. M., Nodal, F. R., Korn, C., Constantinescu, A. O., Mann, E. O., Boyden, E. S., & King, A. J. (2019). Silencing cortical activity during sound-localization training impairs auditory perceptual learning. Nature Communications, 10(1). https://doi.org/10.1038/s41467-019-10770-4

Bakin, J. S., & Weinberger, N. M. (1990). Classical conditioning induces CS-specific receptive field plasticity in the auditory cortex of the guinea pig. Brain Research, 536(1–2), 271–286. https://doi.org/10.1016/0006-8993(90)90035-A

Bandyopadhyay, S., Shamma, S. A., & Kanold, P. O. (2010). Dichotomy of functional organization in the mouse auditory cortex. Nature Neuroscience, 13(3), 361–368. https://doi.org/10.1038/nn.2490

Bar-Yosef, O., & Nelken, I. (2007). The effects of background noise on the neural responses to natural sounds in cat primary auditory cortex. Frontiers in Computational Neuroscience, 1(November), 1–11. https://doi.org/10.3389/neuro.10.003.2007

Barbour, D. L., & Wang, X. (2003). Contrast Tuning in Auditory Cortex. Science, 299(5609), 1073–1075. https://doi.org/10.1126/science.1080425

Bathellier, B., Ushakova, L., & Rumpel, S. (2012). Discrete Neocortical Dynamics Predict Behavioral Categorization of Sounds. Neuron, 76(2), 435–449. https://doi.org/10.1016/j.neuron.2012.07.008

Benichov, J. I., Globerson, E., & Tchernichovski, O. (2016). Finding the beat: From socially coordinated vocalizations in songbirds to rhythmic entrainment in humans. Frontiers in Human Neuroscience, 10(June), 1–7. https://doi.org/10.3389/fnhum.2016.00255

Benichov, J. I., & Vallentin, D. (2020). Inhibition within a premotor circuit controls the timing of vocal turn-taking in zebra finches. Nature Communications, 11(1), 1–10. https://doi.org/10.1038/s41467-019-13938-0

Bergman, A. S. (1990). Auditory scene analysis : the perceptual organization of sound. Cambridge: The MIT Press. https://doi.org/10.7551/mitpress/1486.001.0001

Bernal, B., & Ardila, A. (2016). From hearing sounds to recognizing phonemes: Primary auditory cortex is a truly perceptual language area. AIMS Neuroscience, 3(4), 454–473. https://doi.org/10.3934/Neuroscience.2016.4.454

Bernal, B., Ardila, A., King, A. J., Schnupp, J. W. H., Jasmin, K., Lima, C. F., Scott, S. K., King, A. J., Teki, S., Willmore, B. D. B., Read, H. L., Winer, J. A., Schreiner, C. E., Zatorre, R. J., Belin, P., & Penhune, V. B. (2002). Understanding rostral–caudal auditory cortex contributions to auditory perception. Current Biology, 6(7), 454–473. https://doi.org/10.1038/s41583-019-0160-2

Bertero, A., Zurita, H., Normandin, M., & Apicella, A. J. (2020). Auditory Long-Range Parvalbumin Cortico-Striatal Neurons. Frontiers in Neural Circuits, 14(July), 1–15. https://doi.org/10.3389/fncir.2020.00045

Bigelow, J., Morrill, R. J., Dekloe, J., & Hasenstaub, A. R. (2019). Movement and VIP interneuron activation differentially modulate encoding in mouse auditory cortex. ENeuro, 6(5). https://doi.org/10.1523/ENEURO.0164-19.2019

Bisogno, A., Scarpa, A., Di Girolamo, S., De Luca, P., Cassandro, C., Viola, P., Ricciardiello, F., Greco, A., De Vincentiis, M., Ralli, M., & Di Stadio, A. (2021). Hearing loss and cognitive impairment: Epidemiology, common pathophysiological findings, and treatment considerations. Life, 11(10). https://doi.org/10.3390/life11101102

Bizley, J. K., & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. Nature Reviews Neuroscience, 14(10), 693–707. https://doi.org/10.1038/nrn3565

Bizley, J. K., Nodal, F. R., Bajo, V. M., Nelken, I., & King, A. J. (2007). Physiological and Anatomical Evidence for Multisensory Interactions in Auditory Cortex. Cerebral Cortex, 17(9), 2172–2189. https://doi.org/10.1093/cercor/bhl128

Bizley, J. K., Nodal, F. R., Nelken, I., & King, A. J. (2005). Functional Organization of Ferret Auditory Cortex. Cerebral Cortex, 15(10), 1637–1653. https://doi.org/10.1093/cercor/bhi042

Bizley, J. K., Walker, K. M. M., King, A. J., & Schnupp, J. W. H. (2013). Spectral timbre perception in ferrets: Discrimination of artificial vowels under different listening conditions. The Journal of the Acoustical Society of America, 133(1), 365–376. https://doi.org/10.1121/1.4768798

Bizley, J. K., Walker, K. M. M., Silverman, B. W., King, A. J., & Schnupp, J. W. H. (2009). Interdependent Encoding of Pitch, Timbre, and Spatial Location in Auditory Cortex. The Journal of Neuroscience, 29(7), 2064–2075. https://doi.org/10.1523/JNEUROSCI.4755-08.2009

Blackwell, J. M., Taillefumier, T. O., Natan, R. G., Carruthers, I. M., Magnasco, M. O., & Geffen, M. N. (2016). Stable encoding of sounds over a broad range of statistical parameters in the auditory cortex. European Journal of Neuroscience, 43(6), 751–764. https://doi.org/10.1111/ejn.13144

Brudzynski, S. M. (2013). Ethotransmission: Communication of emotional states through ultrasonic vocalization in rats. Current Opinion in Neurobiology, 23(3), 310–317. https://doi.org/10.1016/j.conb.2013.01.014

Bueti, D. (2011). The sensory representation of time. Frontiers in Integrative Neuroscience, 5(August), 1–3. https://doi.org/10.3389/fnint.2011.00034

Bueti, D., Bahrami, B., & Walsh, V. (2008). Sensory and association cortex in time perception. Journal of Cognitive Neuroscience, 20(6), 1054–1062. https://doi.org/10.1162/jocn.2008.20060

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

Butler, R. A., Diamond, I. T., & Neff, W. D. (1957). Role of auditory cortex in discrimination of changes in frequency. Journal of Neurophysiology, 20(1), 108–120. https://doi.org/10.1152/jn.1957.20.1.108

Carruthers, I. M., Laplagne, D. A., Jaegle, A., Briguglio, J. J., Mwilambwe-Tshilobo, L., Natan, R. G., & Geffen, M. N. (2015). Emergence of invariant representation of vocalizations in the auditory cortex. Journal of Neurophysiology, 114(5), 2726–2740. https://doi.org/10.1152/jn.00095.2015

Casseday, J. H., Fremouw, T., & Covey, E. (2002). The Inferior Colliculus: A Hub for the Central Auditory System. In D. Oertel, R. R. Fay, & A. N. Popper (Eds.), Integrative Functions in the Mammalian Auditory Pathway (pp. 238–318). Springer New York. https://doi.org/10.1007/978-1-4757-3654-0_7

Ceballo, S., Piwkowska, Z., Bourg, J., Daret, A., & Bathellier, B. (2019). Targeted Cortical Manipulation of Auditory Perception. Neuron, 104(6), 1168-1179.e5. https://doi.org/10.1016/j.neuron.2019.09.043

Chait, M., de Cheveigné, A., Poeppel, D., & Simon, J. Z. (2010). Neural dynamics of attending and ignoring in human auditory cortex. Neuropsychologia, 48(11), 3262–3271. https://doi.org/10.1016/j.neuropsychologia.2010.07.007

Chapuis, J., & Wilson, D. A. (2012). Bidirectional plasticity of cortical pattern recognition and behavioral sensory acuity. Nature Neuroscience, 15(1), 155–161. https://doi.org/10.1038/nn.2966

Chechik, G., & Nelken, I. (2012). Auditory abstraction from spectro-temporal features to coding auditory entities. Proceedings of the National Academy of Sciences, 109(46), 18968–18973. https://doi.org/10.1073/pnas.1111242109

Chen, L., Wang, X., Ge, S., & Xiong, Q. (2019). Medial geniculate body and primary auditory cortex differentially contribute to striatal sound representations. Nature Communications, 10(1), 1–10. https://doi.org/10.1038/s41467-019-08350-7

Chubykin, A. A., Roach, E. B., Bear, M. F., & Shuler, M. G. H. (2013). A Cholinergic Mechanism for Reward Timing within Primary Visual Cortex. Neuron, 77(4), 723–735. https://doi.org/10.1016/j.neuron.2012.12.039

Clayton, K. K., Williamson, R. S., Hancock, K. E., Tasaka, G., Mizrahi, A., Hackett, T. A., & Polley, D. B. (2021). Auditory Corticothalamic Neurons Are Recruited by Motor

Preparatory Inputs. Current Biology, 31(2), 310-321.e5. https://doi.org/10.1016/j.cub.2020.10.027

Cohen, L., Rothschild, G., & Mizrahi, A. (2011). Multisensory Integration of Natural Odors and Sounds in the Auditory Cortex. Neuron, 72(2), 357–369. https://doi.org/10.1016/j.neuron.2011.08.019

Cook, J. R., Li, H., Nguyen, B., Huang, H. 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(3), 330–344. https://doi.org/10.1038/s41593-022-01025-5

Cutting, J. E., & Rosner, B. S. (1974). Categories and boundaries in speech and music. Perception & Psychophysics, 16(3), 564–570. https://doi.org/10.3758/BF03198588

Dana, H., Chen, T. W., Hu, A., Shields, B. C., Guo, C., Looger, L. L., Kim, D. S., & Svoboda, K. (2014). Thy1-GCaMP6 transgenic mice for neuronal population imaging in vivo. PLoS ONE, 9(9). https://doi.org/10.1371/journal.pone.0108697

David, S. V., Fritz, J. B., & Shamma, S. A. (2012). Task reward structure shapes rapid receptive field plasticity in auditory cortex. Proceedings of the National Academy of Sciences of the United States of America, 109(6), 2144–2149. https://doi.org/10.1073/pnas.1117717109

Davis, K. A. (2005). Spectral Processing in the Inferior Colliculus. International Review of Neurobiology, 70(05), 169–205. https://doi.org/10.1016/S0074-7742(05)70006-4

Davis, K. A., Ramachandran, R., & May, B. J. (2003). Auditory processing of spectral cues for sound localization in the inferior colliculus. JARO - Journal of the Association for Research in Otolaryngology, 4(2), 148–163. https://doi.org/10.1007/s10162-002-2002-5

Deitch, D., Rubin, A., & Ziv, Y. (2021). Representational drift in the mouse visual cortex. Current Biology, 31(19), 4327-4339.e6. https://doi.org/10.1016/j.cub.2021.07.062

Deneux, T., Kempf, A., Daret, A., Ponsot, E., & Bathellier, B. (2016). Temporal asymmetries in auditory coding and perception reflect multi-layered nonlinearities. Nature Communications, 7. https://doi.org/10.1038/ncomms12682

Doupe, A. J., & Kuhl, P. K. (1999). Birdsong and Human Speech: Common Themes and Mechanisms. Annual Review of Neuroscience, 22(1), 567–631. https://doi.org/10.1146/annurev.neuro.22.1.567

Dunlap, A. G., Besosa, C., Pascual, L. M., Chong, K. K., Walum, H., Kacsoh, D. B., Tankeu, B. B., Lu, K., & Liu, R. C. (2020). Becoming a better parent: Mice learn sounds that improve a stereotyped maternal behavior. Hormones and Behavior, 124(July), 104779. https://doi.org/10.1016/j.yhbeh.2020.104779

Eden, S., & Leibovitz-Ganon, K. (2022). The effects of cochlear implants on sequential time perception. Deafness and Education International, 24(2), 160–178. https://doi.org/10.1080/14643154.2021.1902644

Eggermont, J. J. (2001). Between sound and perception: Reviewing the search for a neural code. Hearing Research, 157(1–2), 1–42. https://doi.org/10.1016/S0378-5955(01)00259-3

Ehret, G. (1987). Left hemisphere advantage in the mouse brain for recognizing ultrasonic communication calls. Nature, 325(6101), 249–251. https://doi.org/10.1038/325249a0

Ehret, G. (2005). Infant rodent ultrasounds - A gate to the understanding of sound communication. Behavior Genetics, 35(1), 19–29. https://doi.org/10.1007/s10519-004-0853-8

Ehret, G., & Haack, B. (1982). Ultrasound recognition in house mice: Key-Stimulus configuration and recognition mechanism. Journal of Comparative Physiology ? A, 148(2), 245–251. https://doi.org/10.1007/BF00619131

Ehret, G., & Koch, M. (1989). Ultrasound-induced Parental Behaviour in House Mice is Controlled by Female Sex Hormones and Parental Experience. Ethology, 80(1–4), 81–93. https://doi.org/10.1111/j.1439-0310.1989.tb00731.x

Ehret, G., & Schreiner, C. E. (2005). Spectral and Intensity Coding in the Auditory Midbrain. In J. A. Winer & C. E. Schreiner (Eds.), The Inferior Colliculus (pp. 312–345). Springer New York. https://doi.org/10.1007/0-387-27083-3_11

Elhilali, M., Fritz, J. B., Chi, T. S., & Shamma, S. A. (2007). Auditory cortical receptive fields: Stable entities with plastic abilities. Journal of Neuroscience, 27(39), 10372–10382. https://doi.org/10.1523/JNEUROSCI.1462-07.2007

Fiorillo, C. D., Newsome, W. T., & Schultz, W. (2008). The temporal precision of reward prediction in dopamine neurons. Nature Neuroscience, 11(8), 966–973. https://doi.org/10.1038/nn.2159

Francis, N. A., Winkowski, D. E., Sheikhattar, A., Armengol, K., Babadi, B., & Kanold, P. O. (2018). Small Networks Encode Decision-Making in Primary Auditory Cortex. Neuron, 97(4), 885-897.e6. https://doi.org/10.1016/j.neuron.2018.01.019

Frégnac, Y., & Bathellier, B. (2015). Cortical Correlates of Low-Level Perception: From Neural Circuits to Percepts. Neuron, 88(1), 110–126. https://doi.org/10.1016/j.neuron.2015.09.041

Fritz, J., Elhilali, M., & Shamma, S. (2005). Active listening: Task-dependent plasticity of spectrotemporal receptive fields in primary auditory cortex. Hearing Research, 206(1–2), 159–176. https://doi.org/10.1016/j.heares.2005.01.015

Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. Nature Neuroscience, 6(11), 1216–1223. https://doi.org/10.1038/nn1141

Froemke, R. C., Merzenich, M. M., & Schreiner, C. E. (2007). A synaptic memory trace for cortical receptive field plasticity. Nature, 450(7168), 425–429. https://doi.org/10.1038/nature06289

Füllgrabe, C. (2013). Age-dependent changes in temporal-fine-structure processing in the absence of peripheral hearing loss. American Journal of Audiology, 22(2), 313–315. https://doi.org/10.1044/1059-0889(2013/12-0070)

Fung, B. J., Sutlief, E., & Hussain Shuler, M. G. (2021). Dopamine and the interdependency of time perception and reward. Neuroscience and Biobehavioral Reviews, 125, 380–391. https://doi.org/10.1016/j.neubiorev.2021.02.030

Galindo-Leon, E. E., Lin, F. G., & Liu, R. C. (2009). Inhibitory Plasticity in a Lateral Band Improves Cortical Detection of Natural Vocalizations. Neuron, 62(5), 705–716. https://doi.org/10.1016/j.neuron.2009.05.001

Galván, V. V., Chen, J., & Weinberger, N. M. (2001). Long-term frequency tuning of local field potentials in the auditory cortex of the waking guinea pig. JARO - Journal of the Association for Research in Otolaryngology, 2(3), 199–215. https://doi.org/10.1007/s101620010062

Geissler, D. B., & Ehret, G. (2004). Auditory perception vs. recognition: Representation of complex communication sounds in the mouse auditory cortical fields. European Journal of Neuroscience, 19(4), 1027–1040. https://doi.org/10.1111/j.1460-9568.2004.03205.x

Ghazanfar, A. A., Maier, J. X., Hoffman, K. L., & Logothetis, N. K. (2005). Multisensory Integration of Dynamic Faces and Voices in Rhesus Monkey Auditory Cortex. The Journal of Neuroscience, 25(20), 5004–5012. https://doi.org/10.1523/JNEUROSCI.0799-05.2005

Gilday, O. D., & Mizrahi, A. (2023). Learning-Induced Odor Modulation of Neuronal Activity in Auditory Cortex. The Journal of Neuroscience, 43(8), 1375–1386. https://doi.org/10.1523/JNEUROSCI.1398-22.2022

Goldberg, J. M., & Neff, W. D. (1961). Frequency Discrimination After Bilateral Ablation Of Cortical Auditory Areas. Journal of Neurophysiology, 24(2), 119–128. https://doi.org/10.1152/jn.1961.24.2.119

Grienberger, C., & Konnerth, A. (2012). Imaging Calcium in Neurons. Neuron, 73(5), 862–885. https://doi.org/10.1016/j.neuron.2012.02.011

Griffiths, T. D. (2003). The Neural Processing of Complex Sounds. In The Cognitive Neuroscience of Music (pp. 168–177). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198525202.003.0011

Griffiths, T. D., & Warren, J. D. (2004). What is an auditory object? Nature Reviews Neuroscience, 5(11), 887–892. https://doi.org/10.1038/nrn1538

Griffiths, T. D., Warren, J. D., Scott, S. K., Nelken, I., & King, A. J. (2004). Cortical processing of complex sound: A way forward? Trends in Neurosciences, 27(4), 181–185. https://doi.org/10.1016/j.tins.2004.02.005

Grose, J. H., & Mamo, S. K. (2010). Processing of temporal fine structure as a function of age. Ear and Hearing, 31(6), 755–760. https://doi.org/10.1097/AUD.0b013e3181e627e7

Guo, L., Walker, W. I., Ponvert, N. D., Penix, P. L., & Jaramillo, S. (2018). Stable representation of sounds in the posterior striatum during flexible auditory decisions. Nature Communications, 9(1). https://doi.org/10.1038/s41467-018-03994-3

Guo, L., Weems, J. T., Walker, W. I., Levichev, A., & Jaramillo, S. (2019). Choice-selective neurons in the auditory cortex and in its striatal target encode reward expectation. Journal of Neuroscience, 39(19), 3687–3697. https://doi.org/10.1523/JNEUROSCI.2585-18.2019

Guo, W., Chambers, A. R., Darrow, K. N., Hancock, K. E., Shinn-cunningham, B. G., & Polley, D. B. (2012). Robustness of Cortical Topography across Fields , Laminae , Anesthetic States , and Neurophysiological Signal Types. 32(27), 9159–9172. https://doi.org/10.1523/JNEUROSCI.0065-12.2012

Guo, W., Hight, A. E., Chen, J. X., Klapoetke, N. C., Hancock, K. E., Shinn-Cunningham, B. G., Boyden, E. S., Lee, D. J., & Polley, D. B. (2015). Hearing the light: neural and perceptual encoding of optogenetic stimulation in the central auditory pathway. Scientific Reports, 5(1), 10319. https://doi.org/10.1038/srep10319

Gygi, B., Kidd, G. R., & Watson, C. S. (2007). Similarity and categorization of environmental sounds. Perception and Psychophysics, 69(6), 839–855. https://doi.org/10.3758/BF03193921

Harpaz, M., Jankowski, M. M., Khouri, L., & Nelken, I. (2021). Emergence of abstract sound representations in the ascending auditory system. Progress in Neurobiology, 202. https://doi.org/10.1016/j.pneurobio.2021.102049

Harper, N. S., Schoppe, O., Willmore, B. D. B., Cui, Z., Schnupp, J. W. H., & King, A. J. (2016). Network Receptive Field Modeling Reveals Extensive Integration and Multi-feature Selectivity in Auditory Cortical Neurons. PLoS Computational Biology, 12(11). https://doi.org/10.1371/journal.pcbi.1005113

Harrington, I. A., Heffner, R. S., & Heffner, H. E. (2001). An investigation of sensory deficits underlying the aphasia-like behavior of macaques with auditory cortex lesions. NeuroReport, 12(6), 1217–1221. https://doi.org/10.1097/00001756-200105080-00032

Heilbron, M., & Chait, M. (2018). Great Expectations: Is there Evidence for Predictive Coding in Auditory Cortex? Neuroscience, 389, 54–73. https://doi.org/10.1016/j.neuroscience.2017.07.061

Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. Psychophysiology, 42(2), 180–190. https://doi.org/10.1111/j.1469-8986.2005.00272.x

Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. NeuroReport, 17(13), 1375–1379. https://doi.org/10.1097/01.wnr.0000233102.43526.e9

Helfer, K. S., & Jesse, A. (2021). Hearing and speech processing in midlife. Hearing Research, 402, 108097. https://doi.org/10.1016/j.heares.2020.108097

Henschke, J. U., Price, A. T., & Pakan, J. M. P. (2021). Enhanced modulation of cell-type specific neuronal responses in mouse dorsal auditory field during locomotion. Cell Calcium, 96, 102390. https://doi.org/10.1016/j.ceca.2021.102390

Hikosaka, O., Sakamoto, M., & Usui, S. (1989). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. Journal of Neurophysiology, 61(4), 814–832. https://doi.org/10.1152/jn.1989.61.4.814

Hinton, S. C., & Meck, W. H. (1997). Erratum: The “internal clocks” of circadian and interval timing. Endeavour, 21(2), 82–87. https://doi.org/10.1016/S0160-9327(97)01043-0

Hromádka, T., DeWeese, M. R., & Zador, A. M. (2008). Sparse representation of sounds in the unanesthetized auditory cortex. PLoS Biology, 6(1), 0124–0137. https://doi.org/10.1371/journal.pbio.0060016

Huang, W., Wang, Y., Qin, J., He, C., Li, Y., Wang, Y., Li, M., Lyu, J., Zhou, Z., Jia, H., Pakan, J., Xie, P., & Zhang, J. (2023). A corticostriatal projection for sound-evoked and anticipatory motor behavior following temporal expectation. NeuroReport, 34(1), 1–8. https://doi.org/10.1097/WNR.0000000000001851

Huerta-Ocampo, I., Mena-Segovia, J., & Bolam, J. P. (2014). Convergence of cortical and thalamic input to direct and indirect pathway medium spiny neurons in the striatum. Brain Structure & Function, 219(5), 1787–1800. https://doi.org/10.1007/s00429-013-0601-z

Hunnicutt, B. J., Jongbloets, B. C., Birdsong, W. T., Gertz, K. J., Zhong, H., & Mao, T. (2016). A comprehensive excitatory input map of the striatum reveals novel functional organization. ELife, 5(November2016), 1–32. https://doi.org/10.7554/eLife.19103

Hussain Shuler, M. G., & Bear, M. F. (2006). Reward Timing in the Primary Visual Cortex. Scientific Reports, March, 231–237. https://doi.org/10.4324/9780203181232-18

Ingber, S., & Eden, S. (2011). Enhancing Sequential Time Perception and Storytelling Ability of Deaf and Hard of Hearing Children. American Annals of the Deaf, 156, 391–401.

Jafari, Z., Kolb, B. E., & Mohajerani, M. H. (2021). Age-related hearing loss and cognitive decline: MRI and cellular evidence. Annals of the New York Academy of Sciences, 1500(1), 17–33. https://doi.org/10.1111/nyas.14617

Jaramillo, S., & Zador, A. M. (2011). The auditory cortex mediates the perceptual effects of acoustic temporal expectation. Nature Neuroscience, 14(2), 246–253. https://doi.org/10.1038/nn.2688

Jazayeri, M., & Shadlen, M. N. (2015). A Neural Mechanism for Sensing and Reproducing a Time Interval. Current Biology, 25(20), 2599–2609. https://doi.org/10.1016/j.cub.2015.08.038

Jones, C. R. G., & Jahanshahi, M. (2011). Dopamine modulates striato-frontal functioning during temporal processing. Frontiers in Integrative Neuroscience, 5(70).

Kadia, S. C., & Wang, X. (2003). Spectral integration in A1 of awake primates: Neurons with single- and multipeaked tuning characteristics. Journal of Neurophysiology, 89(3), 1603–1622. https://doi.org/10.1152/jn.00271.2001

Kaga, K., Shindo, M., & Tanaka, Y. (1997). Central auditory information processing in patients with bilateral auditory cortex lesions. Acta Oto-Laryngologica, Supplement, 532, 77–82. https://doi.org/10.3109/00016489709126148

Kanold, P. O., Nelken, I., & Polley, D. B. (2014). Local versus global scales of organization in auditory cortex. Trends in Neurosciences, 37(9), 502–510. https://doi.org/doi:10.1016/j.tins.2014.06.003

Kato, H. K., Gillet, S. N., & Isaacson, J. S. (2015). Flexible Sensory Representations in Auditory Cortex Driven by Behavioral Relevance. Neuron, 88(5), 1027–1039. https://doi.org/10.1016/j.neuron.2015.10.024

Kayser, C., Petkov, C. I., Augath, M., & Logothetis, N. K. (2005). Integration of Touch and Sound in Auditory Cortex. Neuron, 48(2), 373–384. https://doi.org/10.1016/j.neuron.2005.09.018

Kent, R. D. (2000). Research on speech motor control and its disorders. Journal of Communication Disorders, 33(5), 391–428. https://doi.org/10.1016/S0021-9924(00)00023-X

Khouri, L., & Nelken, I. (2015). Detecting the unexpected. Current Opinion in Neurobiology, 35, 142–147. https://doi.org/10.1016/j.conb.2015.08.003

Kim, K. X., Atencio, C. A., & Schreiner, C. E. (2020). Stimulus dependent transformations between synaptic and spiking receptive fields in auditory cortex. Nature Communications, 11(1), 1–15. https://doi.org/10.1038/s41467-020-14835-7

King, A. J., & Moore, D. R. (1991). Plasticity of auditory maps in the brain. Trends in Neurosciences, 14(1), 31–37. https://doi.org/10.1016/0166-2236(91)90181-S

King, A. J., & Schnupp, J. W. H. (2007). The auditory cortex. Current Biology, 17(7), 236–239. https://doi.org/10.1016/j.cub.2007.01.046

King, A. J., Teki, S., & Willmore, B. D. B. (2018). Recent advances in understanding the auditory cortex. F1000Research, 7(1555). https://doi.org/10.12688/F1000RESEARCH.15580.1

Kisley, M. A., & Gerstein, G. L. (2001). Daily variation and appetitive conditioning-induced plasticity of auditory cortex receptive fields. European Journal of Neuroscience, 13(10), 1993–2003. https://doi.org/10.1046/j.0953-816X.2001.01568.x

Klein, D. J., Simon, J. Z., Depireux, D. A., & Shamma, S. A. (2006). Stimulus-invariant processing and spectrotemporal reverse correlation in primary auditory cortex. Journal of Computational Neuroscience, 20(2), 111–136. https://doi.org/10.1007/s10827-005-3589-4

Kuchibhotla, K., & Bathellier, B. (2018). Neural encoding of sensory and behavioral complexity in the auditory cortex. Current Opinion in Neurobiology, 52, 65–71. https://doi.org/10.1016/j.conb.2018.04.002

Kuchibhotla, K. V., Gill, J. V., Lindsay, G. W., Papadoyannis, E. S., Field, R. E., Sten, T. A. H., Miller, K. D., & Froemke, R. C. (2017). Parallel processing by cortical inhibition enables context-dependent behavior. Nature Neuroscience, 20(1), 62–71. https://doi.org/10.1038/nn.4436

Kumar, S., Sedley, W., Nourski, K. V., Kawasaki, H., Oya, H., Patterson, R. D., Howard, M. A., Friston, K. J., & Griffiths, T. D. (2011). Predictive Coding and Pitch Processing in the Auditory Cortex. Journal of Cognitive Neuroscience, 23(10), 3084–3094. https://doi.org/10.1162/jocn_a_00021

Kurti, A. N., & Matell, M. S. (2011). Nucleus Accumbens Dopamine Modulates Response Rate but Not Response Timing in an Interval Timing Task. Behavioral Neuroscience, 125(2), 215–225. https://doi.org/10.1037/a0022892

LeDoux, J. E., Farb, C. R., & Romanski, L. M. (1991). Overlapping projections to the amygdala and striatum from auditory processing areas of the thalamus and cortex. Neuroscience Letters, 134(1), 139–144. https://doi.org/10.1016/0304-3940(91)90526-Y

Lee, J., & Rothschild, G. (2021). Encoding of acquired sound-sequence salience by auditory cortical offset responses. Cell Reports, 37(5). https://doi.org/10.1016/j.celrep.2021.109927

Leopold, D. A. (2012). Primary visual cortex: Awareness and blindsight. Annual Review of Neuroscience, 35, 91–109. https://doi.org/10.1146/annurev-neuro-062111-150356

Leow, L.-A., & Grahn, J. A. (2014). Neurobiology of Interval Timing. In Advances in experimental medicine and biology (Vol. 829). http://www.ncbi.nlm.nih.gov/pubmed/25358718

Levinson, S. C. (2016). Turn-taking in Human Communication - Origins and Implications for Language Processing. Trends in Cognitive Sciences, 20(1), 6–14. https://doi.org/10.1016/j.tics.2015.10.010

Li, J., Liao, X., Zhang, J., Wang, M., Yang, N., Zhang, J., Lv, G., Li, H., Lu, J., Ding, R., Li, X., Guang, Y., Yang, Z., Qin, H., Jin, W., Zhang, K., He, C., Jia, H., Zeng, S., … Chen, X. (2017). Primary Auditory Cortex is Required for Anticipatory Motor Response. Cerebral Cortex, 27(6), 3254–3271. https://doi.org/10.1093/cercor/bhx079

Li, Z., Wei, J. X., Zhang, G. W., Huang, J. J., Zingg, B., Wang, X., Tao, H. W., & Zhang, L. I. (2021). Corticostriatal control of defense behavior in mice induced by auditory looming cues. Nature Communications, 12(1), 1–13. https://doi.org/10.1038/s41467-021-21248-7

Lin, F. G., Galindo-Leon, E. E., Ivanova, T. N., Mappus, R. C., & Liu, R. C. (2013). A role for maternal physiological state in preserving auditory cortical plasticity for salient infant calls. Neuroscience, 247, 102–116. https://doi.org/10.1016/j.neuroscience.2013.05.020

Lin, F. R., Yaffe, K., Xia, J., Xue, Q. L., Harris, T. B., Purchase-Helzner, E., Satterfield, S., Ayonayon, H. N., Ferrucci, L., & Simonsick, E. M. (2013). Hearing loss and cognitive decline in older adults. JAMA Internal Medicine, 173(4), 293–299. https://doi.org/10.1001/jamainternmed.2013.1868

Lin, P. A., Asinof, S. K., Edwards, N. J., & Isaacson, J. S. (2019). Arousal regulates frequency tuning in primary auditory cortex. Proceedings of the National Academy of Sciences of the United States of America, 116(50), 25304–25310. https://doi.org/10.1073/pnas.1911383116

Liu, J., & Kanold, P. O. (2021). Diversity of receptive fields and sideband inhibition with complex thalamocortical and intracortical origin in L2/3 of mouse primary auditory cortex. Journal of Neuroscience, 41(14), 3142–3162. https://doi.org/10.1523/JNEUROSCI.1732-20.2021

Liu, J., Whiteway, M. R., Sheikhattar, A., Butts, D. A., Babadi, B., & Kanold, P. O. (2019). Parallel Processing of Sound Dynamics across Mouse Auditory Cortex via Spatially Patterned Thalamic Inputs and Distinct Areal Intracortical Circuits. Cell Reports, 27(3), 872-885.e7. https://doi.org/10.1016/j.celrep.2019.03.069

Liu, L., Shen, P., He, T., Chang, Y., Shi, L., Tao, S., Li, X., Xun, Q., Guo, X., Yu, Z., & Wang, J. (2016). Noise induced hearing loss impairs spatial learning/memory and

hippocampal neurogenesis in mice. Scientific Reports, 6, 1–9. https://doi.org/10.1038/srep20374

Lütcke, H., Margolis, D. J., & Helmchen, F. (2013). Steady or changing? Long-term monitoring of neuronal population activity. Trends in Neurosciences, 36(7), 375–384. https://doi.org/10.1016/j.tins.2013.03.008

Maas, E., Robin, D. A., Hula, S. N. A., Freedman, S. E., Wulf, G., Ballard, K. J., & Schmidt, R. A. (2008). Principles of motor learning in treatment of motor speech disorders. American Journal of Speech-Language Pathology, 17(3), 277–298. https://doi.org/10.1044/1058-0360(2008/025)

Maor, I., Shalev, A., & Mizrahi, A. (2016). Distinct Spatiotemporal Response Properties of Excitatory Versus Inhibitory Neurons in the Mouse Auditory Cortex. Cerebral Cortex, 26(11), 4242–4252. https://doi.org/10.1093/cercor/bhw266

Maor, I., Shwartz-Ziv, R., Feigin, L., Elyada, Y., Sompolinsky, H., & Mizrahi, A. (2020). Neural Correlates of Learning Pure Tones or Natural Sounds in the Auditory Cortex. Frontiers in Neural Circuits, 13. https://doi.org/10.3389/fncir.2019.00082

Marsh, R. A., Nataraj, K., Gans, D., Portfors, C. V., & Wenstrup, J. J. (2006). Auditory responses in the cochlear nucleus of awake mustached bats: Precursors to spectral integration in the auditory midbrain. Journal of Neurophysiology, 95(1), 88–105. https://doi.org/10.1152/jn.00634.2005

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(4), 760–773. https://doi.org/10.1037/0735-7044.117.4.760

Mazzucato, L. (2022). Neural mechanisms underlying the temporal organization of naturalistic animal behavior. ELife, 11. https://doi.org/10.7554/elife.76577

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

Meng, X., Winkowski, D. E., Kao, J. P. Y., & Kanold, P. O. (2017). Sublaminar Subdivision of Mouse Auditory Cortex Layer 2/3 Based on Functional Translaminar Connections. The Journal of Neuroscience, 37(42), 10200–10214. https://doi.org/10.1523/jneurosci.1361-17.2017

Merzenich, M. M., Kaas, J. H., & Roth, G. L. (1976). Auditory cortex in the grey squirrel: Tonotopic organization and architectonic fields. Journal of Comparative Neurology, 166(4), 387–401. https://doi.org/10.1002/cne.901660402

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://doi.org/10.1038/nn.2272

Mizrahi, A., Shalev, A., & Nelken, I. (2014). Single neuron and population coding of natural sounds in auditory cortex. In Current Opinion in Neurobiology (Vol. 24, Issue 1, pp. 103–110). https://doi.org/10.1016/j.conb.2013.09.007

Monk, K. J., Allard, S., & Hussain Shuler, M. G. (2020). Reward Timing and Its Expression by Inhibitory Interneurons in the Mouse Primary Visual Cortex. Cerebral Cortex, 30(8), 4662–4676. https://doi.org/10.1093/cercor/bhaa068

Musall, S., Haiss, F., Weber, B., & von der Behrens, W. (2017). Deviant Processing in the Primary Somatosensory Cortex. Cerebral Cortex, 27(1), 863–876. https://doi.org/10.1093/cercor/bhv283

Namboodiri, V. M. K., Huertas, M. A., Monk, K. J., Shouval, H. Z., & Shuler, M. G. H. (2015). Visually cued action timing in the primary visual cortex. Neuron, 86(1), 319–330. https://doi.org/10.1016/j.neuron.2015.02.043

Narayanan, N. S., & Laubach, M. (2009). Delay activity in rodent frontal cortex during a simple reaction time task. Journal of Neurophysiology, 101(6), 2859–2871. https://doi.org/10.1152/jn.90615.2008

Natarajan, N., Batts, S., & Stankovic, K. M. (2023). Noise-Induced Hearing Loss. Journal of Clinical Medicine, 12(6), 2347. https://doi.org/10.3390/jcm12062347

Nelken, I. (2004). Processing of complex stimuli and natural scenes in the auditory cortex. Current Opinion in Neurobiology, 14(4), 474–480. https://doi.org/10.1016/j.conb.2004.06.005

Nelken, I. (2008). Processing of complex sounds in the auditory system. In Current Opinion in Neurobiology (Vol. 18, Issue 4, pp. 413–417). https://doi.org/10.1016/j.conb.2008.08.014

Nelken, I. (2014). Stimulus-specific adaptation and deviance detection in the auditory system: experiments and models. Biological Cybernetics, 108(5), 655–663. https://doi.org/10.1007/s00422-014-0585-7

Nelken, I., & Bar-Yosef, O. (2009). Neurons and objects: The case of auditory cortex. Frontiers in Neuroscience, 2(JUL), 107–113. https://doi.org/10.3389/neuro.01.009.2008

Nelken, I., Bizley, J. K., Nodal, F. R., Ahmed, B., Schnupp, J. W. H., & King, A. J. (2004). Large-scale organization of ferret auditory cortex revealed using continuous acquisition of intrinsic optical signals. Journal of Neurophysiology, 92(4), 2574–2588. https://doi.org/10.1152/jn.00276.2004

Nelken, I., Bizley, J., Shamma, S. A., Shamma, S. A., Wang, X. Q., & Wang, X. Q. (2014). Auditory cortical processing in real-world listening: The auditory system going real. Journal of Neuroscience, 34(46), 15135–15138. https://doi.org/10.1523/JNEUROSCI.2989-14.2014

Nelken, I., Fishbach, A., Las, L., Ulanovsky, N., & Farkas, D. (2003). Primary auditory cortex of cats: Feature detection or something else? Biological Cybernetics, 89(5), 397–406. https://doi.org/10.1007/s00422-003-0445-3

Nelken, I., Rotman, Y., & Yosef, O. B. (1999). Responses of auditory-cortex neurons to structural features of natural sounds. Nature, 397(6715), 154–157. https://doi.org/10.1038/16456

Ng, C. W., Plakke, B., & Poremba, A. (2009). Primate auditory recognition memory performance varies with sound type. Hearing Research, 256(1–2), 64–74. https://doi.org/10.1016/j.heares.2009.06.014

Ning, W., Bladon, J. H., & Hasselmo, M. E. (2022). Complementary representations of time in the prefrontal cortex and hippocampus. Hippocampus, 32(8), 577–596. https://doi.org/10.1002/hipo.23451

Nixon, J. S., & Tomaschek, F. (2021). Prediction and error in early infant speech learning: A speech acquisition model. Cognition, 212, 104697. https://doi.org/10.1016/j.cognition.2021.104697

O'sullivan, C., Weible, A. P., & Wehr, M. (2019). Auditory cortex contributes to discrimination of pure tones. ENeuro, 6(5), 1–14. https://doi.org/10.1523/ENEURO.0340-19.2019

Ohl, F. W., Wetzel, W., Wagner, T., Rech, A., & Scheich, H. (1999). Bilateral ablation of auditory cortex in Mongolian gerbil affects discrimination of frequency modulated tones but not of pure tones. Learning and Memory, 6(4), 347–362. https://doi.org/10.1101/lm.6.4.347

Okada, K., Matchin, W., & Hickok, G. (2018). Neural evidence for predictive coding in auditory cortex during speech production. Psychonomic Bulletin and Review, 25(1), 423–430. https://doi.org/10.3758/s13423-017-1284-x

Ozmeral, E. J., Eddins, A. C., Frisina, D. R., & Eddins, D. A. (2016). Large cross-sectional study of presbycusis reveals rapid progressive decline in auditory temporal acuity. Neurobiology of Aging, 43, 72–78. https://doi.org/10.1016/j.neurobiolaging.2015.12.024

Pachitariu, M., Stringer, C., Dipoppa, M., Schröder, S., Rossi, L. F., Dalgleish, H., Carandini, M., & Harris, K. (2016). Suite2p: Beyond 10,000 neurons with standard two-photon microscopy. BioRxiv, 061507. https://doi.org/10.1101/061507

Pancholi, R., Ryan, L., & Peron, S. (2023). Learning in a sensory cortical microstimulation task is associated with elevated representational stability. Nature Communications, 14(1). https://doi.org/10.1038/s41467-023-39542-x

Parras, G. G., Nieto-Diego, J., Carbajal, G. V., Valdés-Baizabal, C., Escera, C., & Malmierca, M. S. (2017). Neurons along the auditory pathway exhibit a hierarchical

organization of prediction error. Nature Communications, 8(1). https://doi.org/10.1038/s41467-017-02038-6

Peelle, J. E., & Wingfield, A. (2016). The Neural Consequences of Age-Related Hearing Loss. Trends in Neurosciences, 39(7), 486–497. https://doi.org/10.1016/j.tins.2016.05.001

Pendyam, S., Bravo-Rivera, C., Burgos-Robles, A., Sotres-Bayon, F., Quirk, G. J., & Nair, S. S. (2013). Fear signaling in the prelimbic-amygdala circuit: A computational modeling and recording study. Journal of Neurophysiology, 110(4), 844–861. https://doi.org/10.1152/jn.00961.2012

Pérez-Ortega, J., Alejandre-García, T., & Yuste, R. (2021). Long-term stability of cortical ensembles. ELife, 10. https://doi.org/10.7554/eLife.64449

Peron, S. P., Freeman, J., Iyer, V., Guo, C., & Svoboda, K. (2015). A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. Neuron, 86(3), 783–799. https://doi.org/10.1016/j.neuron.2015.03.027

Polley, D. B., Steinberg, E. E., & Merzenich, M. M. (2006). Perceptual learning directs auditory cortical map reorganization through top-down influences. Journal of Neuroscience, 26(18), 4970–4982. https://doi.org/10.1523/JNEUROSCI.3771-05.2006

Polley, D. B., Thompson, J. H., & Guo, W. (2013). Brief hearing loss disrupts binaural integration during two early critical periods of auditory cortex development. Nature Communications, 4, 1–13. https://doi.org/10.1038/ncomms3547

Ranson, A. (2017). Stability and Plasticity of Contextual Modulation in the Mouse Visual Cortex. Cell Reports, 18(4), 840–848. https://doi.org/10.1016/j.celrep.2016.12.080

Rauschecker, J. P. (1998). Cortical processing of complex sounds. Current Opinion in Neurobiology, 8(4), 516–521. https://doi.org/10.1016/S0959-4388(98)80040-8

Read, H. L., Winer, J. A., & Schreiner, C. E. (2002). Functional architecture of auditory cortex. Current Opinion in Neurobiology, 12(Vi), 433–440.

Recanzone, G. H., Schreiner, C. E., & Merzenich, M. M. (1993). Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. Journal of Neuroscience, 13(1), 87–103. https://doi.org/10.1523/jneurosci.13-01-00087.1993

Ress, D., & Chandrasekaran, B. (2013). Tonotopic organization in the depth of human inferior colliculus. Frontiers in Human Neuroscience, 7(SEP), 1–10. https://doi.org/10.3389/fnhum.2013.00586

Reznik, D., Guttman, N., Buaron, B., Zion-Golumbic, E., & Mukamel, R. (2021). Action-locked Neural Responses in Auditory Cortex to Self-generated Sounds. Cerebral Cortex, 31(12), 5560–5569. https://doi.org/10.1093/cercor/bhab179

Rhode, W. S., & Smith, P. H. (1986). Encoding timing and intensity in the ventral cochlear nucleus of the cat. Journal of Neurophysiology, 56(2), 261–286. https://doi.org/10.1152/jn.1986.56.2.261

Rokni, U., Richardson, A. G., Bizzi, E., & Seung, H. S. (2007). Motor Learning with Unstable Neural Representations. Neuron, 54(4), 653–666. https://doi.org/10.1016/j.neuron.2007.04.030

Romero, S., Hight, A. E., Clayton, K. K., Resnik, J., Williamson, R. S., Hancock, K. E., & Polley, D. B. (2020). Cellular and Widefield Imaging of Sound Frequency Organization in Primary and Higher Order Fields of the Mouse Auditory Cortex. Cerebral Cortex, 30(3), 1603–1622. https://doi.org/10.1093/cercor/bhz190

Roth, B. L. (2016). DREADDs for Neuroscientists. Neuron, 89(4), 683–694. https://doi.org/10.1016/j.neuron.2016.01.040

Rothschild, G., Cohen, L., Mizrahi, A., & Nelken, I. (2013). Elevated correlations in neuronal ensembles of mouse auditory cortex following parturition. Journal of Neuroscience, 33(31), 12851–12861. https://doi.org/10.1523/JNEUROSCI.4656-12.2013

Rothschild, G., & Mizrahi, A. (2015). Global Order and Local Disorder in Brain Maps. Annual Review of Neuroscience, 38(1), 247–268. https://doi.org/10.1146/annurev-neuro-071013-014038

Rothschild, G., Nelken, I., & Mizrahi, A. (2010). Functional organization and population dynamics in the mouse primary auditory cortex. Nature Neuroscience, 13(3), 353–360. https://doi.org/10.1038/nn.2484

Rouiller, E., de Ribaupierre, Y., Morel, A., & de Ribaupierre, F. (1983). Intensity functions of single unit responses to tone in the medial geniculate body of cat. Hearing Research, 11(2), 235–247. https://doi.org/10.1016/0378-5955(83)90081-3

Rubin, J., Ulanovsky, N., Nelken, I., & Tishby, N. (2016). The Representation of Prediction Error in Auditory Cortex. PLoS Computational Biology, 12(8), 1–28. https://doi.org/10.1371/journal.pcbi.1005058

Rule, M. E., O'Leary, T., & Harvey, C. D. (2019). Causes and consequences of representational drift. In Current Opinion in Neurobiology (Vol. 58, pp. 141–147). Elsevier Ltd. https://doi.org/10.1016/j.conb.2019.08.005

Rummell, B. P., Klee, J. L., & Sigurdsson, T. (2016). Attenuation of responses to self-generated sounds in auditory cortical neurons. Journal of Neuroscience, 36(47), 12010–12026. https://doi.org/10.1523/JNEUROSCI.1564-16.2016

Rybalko, N., Šuta, D., Nwabueze-Ogbo, F., & Syka, J. (2006). Effect of auditory cortex lesions on the discrimination of frequency-modulated tones in rats. European Journal of Neuroscience, 23(6), 1614–1622. https://doi.org/10.1111/j.1460-9568.2006.04688.x

Sadagopan, S., & Wang, X. (2008). Level invariant representation of sounds by populations of neurons in primary auditory cortex. Journal of Neuroscience, 28(13), 3415–3426. https://doi.org/10.1523/JNEUROSCI.2743-07.2008

Sadagopan, S., & Wang, X. (2009). Nonlinear spectrotemporal interactions underlying selectivity for complex sounds in auditory cortex. Journal of Neuroscience, 29(36), 11192–11202. https://doi.org/10.1523/JNEUROSCI.1286-09.2009

Schneider, D. M., Nelson, A., & Mooney, R. (2014). A synaptic and circuit basis for corollary discharge in the auditory cortex. Nature, 513(7517), 189–194. https://doi.org/10.1038/nature13724

Schneider, D. M., & Woolley, S. M. N. (2013). Sparse and Background-Invariant Coding of Vocalizations in Auditory Scenes. Neuron, 79(1), 141–152. https://doi.org/10.1016/j.neuron.2013.04.038

Schoonover, C. E., Ohashi, S. N., Axel, R., & Fink, A. J. P. (2021). Representational drift in primary olfactory cortex. Nature, 594(7864), 541–546. https://doi.org/10.1038/s41586-021-03628-7

Schreiner, C. E., Froemke, R. C., & Atencio, C. A. (2011). Spectral Processing in Auditory Cortex. In J. A. Winer & C. E. Schreiner (Eds.), The Auditory Cortex (pp. 275–308). Springer US. https://doi.org/10.1007/978-1-4419-0074-6_13

Schwartz, Z. P., Buran, B. N., & David, S. V. (2020). Pupil-associated states modulate excitability but not stimulus selectivity in primary auditory cortex. Journal of Neurophysiology, 123(1), 191–208. https://doi.org/10.1152/jn.00595.2019

Shackleton, T. M., Skottun, B. C., Arnott, R. H., & Palmer, A. R. (2003). Interaural time difference discrimination thresholds for single neurons in the inferior colliculus of guinea pigs. Journal of Neuroscience, 23(2), 716–724. https://doi.org/10.1523/jneurosci.23-02-00716.2003

Shamash, P., Carandini, M., Harris, K., & Steinmetz, N. (2018). A tool for analyzing electrode tracks from slice histology. BioRxiv, 447995. https://www.biorxiv.org/content/10.1101/447995v1%0Ahttps://www.biorxiv.org/content/10.1101/447995v1.abstract

Singer, Y., Teramoto, Y., Willmore, B. D. B., King, A. J., Schnupp, J. W. H., & Harper, N. S. (2018). Sensory cortex is optimised for prediction of future input. ELife, 7, 1–31. https://doi.org/10.7554/eLife.31557

Smeal, R. M., Keefe, K. A., & Wilcox, K. S. (2008). Differences in excitatory transmission between thalamic and cortical afferents to single spiny efferent neurons of rat dorsal striatum. European Journal of Neuroscience, 28(10), 2041–2052. https://doi.org/10.1111/j.1460-9568.2008.06505.x

Souffi, S., Lorenzi, C., Varnet, L., Huetz, C., & Edeline, J. M. (2020). Noise-Sensitive but More Precise Subcortical Representations Coexist with Robust Cortical Encoding of Natural Vocalizations. Journal of Neuroscience, 40(27), 5228–5246. https://doi.org/10.1523/JNEUROSCI.2731-19.2020

Stefanics, G., Hangya, B., Hernádi, I., Winkler, I., Lakatos, P., & Ulbert, I. (2010). Phase entrainment of human delta oscillations can mediate the effects of expectation on reaction speed. Journal of Neuroscience, 30(41), 13578–13585. https://doi.org/10.1523/JNEUROSCI.0703-10.2010

Stiebler, I., & Ehret, G. (1985). Inferior colliculus of the house mouse. I. A quantitative study of tonotopic organization, frequency representation, and tone-threshold distribution. Journal of Comparative Neurology, 238(1), 65–76. https://doi.org/10.1002/cne.902380106

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J. P., Yoon, K. E., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. Proceedings of the National Academy of Sciences of the United States of America, 106(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106

Stosiek, C., Garaschuk, O., Holthoff, K., & Konnerth, A. (2003). In vivo two-photon calcium imaging of neuronal networks. Proceedings of the National Academy of Sciences of the United States of America, 100(12), 7319–7324. https://doi.org/10.1073/pnas.1232232100

Strand, E. (1995). Treatment of motor speech disorders in children. Seminars in Speech and Language, 16(2), 126–139. https://doi.org/10.1055/s-2008-1064115

Suga, N., O'Neill, W. E., Kujirai, K., & Manabe, T. (1983). Specificity of combination-sensitive neurons for processing of complex biosonar signals in auditory cortex of the mustached bat. Journal of Neurophysiology, 49(6), 1573–1626. https://doi.org/10.1152/jn.1983.49.6.1573

Summerfield, C., & De Lange, F. P. (2014). Expectation in perceptual decision making: Neural and computational mechanisms. Nature Reviews Neuroscience, 15(11), 745–756. https://doi.org/10.1038/nrn3838

Sun, W., Tang, P., Liang, Y., Li, J., Feng, J., Zhang, N., Lu, D., He, J., & Chen, X. (2022). The anterior cingulate cortex directly enhances auditory cortical responses in air-puffing-facilitated flight behavior. Cell Reports, 38(10), 110506. https://doi.org/10.1016/j.celrep.2022.110506

Suri, H., & Rothschild, G. (2022). Enhanced Stability of Complex Sound Representations Relative to Simple Sounds in the Auditory Cortex. ENeuro, 9(4). https://doi.org/10.1523/ENEURO.0031-22.2022

Suri, H., Salgado-Puga, K., Wang, Y., Allen, N., Lane, K., Granroth, K., Olivei, A., Nass, N., & Rothschild, G. (2023). A Cortico-Striatal Circuit for Sound-Triggered Prediction of Reward Timing. BioRxiv. https://doi.org/10.1101/2023.11.21.568134

Surlykke, A., & Moss, C. F. (2000). Echolocation behavior of big brown bats, Eptesicus fuscus , in the field and the laboratory. The Journal of the Acoustical Society of America, 108(5), 2419–2429. https://doi.org/10.1121/1.1315295

Sutter, M. L., & Schreiner, C. E. (1991). Physiology and topography of neurons with multipeaked tuning curves in cat primary auditory cortex. Journal of Neurophysiology, 65(5), 1207–1226. https://doi.org/10.1152/jn.1991.65.5.1207

Sutter, M. L., & Shamma, S. A. (2011). The Relationship of Auditory Cortical Activity to Perception and Behavior. In The Auditory Cortex (pp. 617–641). Springer US. https://doi.org/10.1007/978-1-4419-0074-6_29

Svoboda, K., & Yasuda, R. (2006). Principles of Two-Photon Excitation Microscopy and Its Applications to Neuroscience. Neuron, 50(6), 823–839. https://doi.org/10.1016/j.neuron.2006.05.019

Taaseh, N., Yaron, A., & Nelken, I. (2011). Stimulus-specific adaptation and deviance detection in the rat auditory cortex. PLoS ONE, 6(8). https://doi.org/10.1371/journal.pone.0023369

Tallal, P., Miller, S., & Fitch, R. H. (1995). Neurobiological Basis of Speech: A Case for the Preeminence of Temporal Processing. The Irish Journal of Psychology, 16(3), 194–219. https://doi.org/10.1080/03033910.1995.10558057

Tallot, L., & Doyère, V. (2020). Neural encoding of time in the animal brain. Neuroscience and Biobehavioral Reviews, 115(June 2019), 146–163. https://doi.org/10.1016/j.neubiorev.2019.12.033

Tischbirek, C. H., Noda, T., Tohmi, M., Birkner, A., Nelken, I., & Konnerth, A. (2019). In Vivo Functional Mapping of a Cortical Column at Single-Neuron Resolution. Cell Reports, 27(5), 1319-1326.e5. https://doi.org/10.1016/j.celrep.2019.04.007

Town, S. M., Wood, K. C., & Bizley, J. K. (2018). Sound identity is represented robustly in auditory cortex during perceptual constancy. Nature Communications, 9(1). https://doi.org/10.1038/s41467-018-07237-3

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

Tremblay, P., Baroni, M., & Hasson, U. (2013). Processing of speech and non-speech sounds in the supratemporal plane: Auditory input preference does not predict sensitivity to statistical structure. NeuroImage, 66, 318–332. https://doi.org/10.1016/j.neuroimage.2012.10.055

Tunes, G. C., de Oliveira, E. F., Vieira, E. U. P., Caetano, M. S., Cravo, A. M., & Reyes, M. B. (2022). Time encoding migrates from prefrontal cortex to dorsal striatum during learning of a self-timed response duration task. ELife, 11, 1–19. https://doi.org/10.7554/eLife.65495

Uchida, Y., Sugiura, S., Nishita, Y., Saji, N., Sone, M., & Ueda, H. (2019). Age-related hearing loss and cognitive decline — The potential mechanisms linking the two. Auris Nasus Larynx, 46(1), 1–9. https://doi.org/10.1016/j.anl.2018.08.010

Ulanovsky, N., Las, L., Farkas, D., & Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. Journal of Neuroscience, 24(46), 10440–10453. https://doi.org/10.1523/JNEUROSCI.1905-04.2004

Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. Nature Neuroscience, 6(4), 391–398. https://doi.org/10.1038/nn1032

Vivaldo, C. A., Lee, J., Shorkey, M. C., Keerthy, A., & Rothschild, G. (2023). Auditory cortex ensembles jointly encode sound and locomotion speed to support sound perception during movement. PLoS Biology, 21(8), e3002277. https://doi.org/10.1371/journal.pbio.3002277

Wang, M., Liao, X., Li, R., Liang, S., Ding, R., Li, J., Zhang, J., He, W., Liu, K., Pan, J., Zhao, Z., Li, T., Zhang, K., Li, X., Lyu, J., Zhou, Z., Varga, Z., Mi, Y., Zhou, Y., … Chen, X. (2020). Single-neuron representation of learned complex sounds in the auditory cortex. Nature Communications, 11(1). https://doi.org/10.1038/s41467-020-18142-z

Wang, Q., Ding, S. L., Li, Y., Royall, J., Feng, D., Lesnar, P., Graddis, N., Naeemi, M., Facer, B., Ho, A., Dolbeare, T., Blanchard, B., Dee, N., Wakeman, W., Hirokawa, K. E., Szafer, A., Sunkin, S. M., Oh, S. W., Bernard, A., … Ng, L. (2020). The Allen Mouse Brain Common Coordinate Framework: A 3D Reference Atlas. Cell, 181(4), 936-953.e20. https://doi.org/10.1016/j.cell.2020.04.007

Wang, X. (2013). The harmonic organization of auditory cortex. Frontiers in Systems Neuroscience, 7(DEC), 1–11. https://doi.org/10.3389/fnsys.2013.00114

Wang, X., Lu, T., Snider, R. K., & Liang, L. (2005). Sustained firing in auditory cortex evoked by preferred stimuli. Nature, 435(7040), 341–346. https://doi.org/10.1038/nature03565

Wearden, J. (2005). Origins and development of internal clock theories of psychological time. Psychologie Francaise, 50(1), 7–25.

Weible, A. P., Liu, C., Niell, C. M., & Wehr, M. (2014). Auditory cortex is required for fear potentiation of gap detection. Journal of Neuroscience, 34(46), 15437–15445. https://doi.org/10.1523/JNEUROSCI.3408-14.2014

Whitton, J. P., Hancock, K. E., & Polley, D. B. (2014). Immersive audiomotor game play enhances neural and perceptual salience of weak signals in noise. Proceedings of the

National Academy of Sciences of the United States of America, 111(25). https://doi.org/10.1073/pnas.1322184111

Wiener, M., Lohoff, F. W., & Branch Coslett, H. (2011). Double dissociation of dopamine genes and timing in humans. Journal of Cognitive Neuroscience, 23(10), 2811–2821. https://doi.org/10.1162/jocn.2011.21626

Wiener, M., Matell, M. S., & Coslett, H. B. (2011). Multiple mechanisms for temporal processing. Frontiers in Integrative Neuroscience, 5(July), 1–3. https://doi.org/10.3389/fnint.2011.00031

Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. Trends in Cognitive Sciences, 13(12), 532–540. https://doi.org/10.1016/j.tics.2009.09.003

Wollberg, Z., & Newman, J. D. (1972). Auditory Cortex of Squirrel Monkey: Response Patterns of Single Cells to Species-Specific Vocalizations. Science, 175(4018), 212–214. https://doi.org/10.1126/science.175.4018.212

Xiong, Q., Znamenskiy, P., & Zador, A. M. (2015). Selective corticostriatal plasticity during acquisition of an auditory discrimination task. Nature, 521(7552), 348–351. https://doi.org/10.1038/nature14225

Yaron, A., Hershenhoren, I., & Nelken, I. (2012). Sensitivity to Complex Statistical Regularities in Rat Auditory Cortex. Neuron, 76(3), 603–615. https://doi.org/10.1016/j.neuron.2012.08.025

Yu, L., Hu, J., Shi, C., Zhou, L., Tian, M., Zhang, J., & Xu, J. (2021). The causal role of auditory cortex in auditory working memory. ELife, 10, 1–19. https://doi.org/10.7554/ELIFE.64457

Yumoto, N., Lu, X., Henry, T. R., Miyachi, S., Nambu, A., Fukai, T., & Takada, M. (2011). A neural correlate of the processing of multi-second time intervals in primate prefrontal cortex. PLoS ONE, 6(4), 3–9. https://doi.org/10.1371/journal.pone.0019168

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

Zhang, L. I., Bao, S., & Merzenich, M. M. (2001). Persistent and specific influences of early acoustic environments on primary auditory cortex. Nature Neuroscience, 4(11), 1123–1130. https://doi.org/10.1038/nn745

Zhou, X., Panizzutti, R., de Villers-Sidani, É., Madeira, C., & Merzenich, M. M. (2011). Natural restoration of critical period plasticity in the juvenile and adult primary auditory cortex. Journal of Neuroscience, 31(15), 5625–5634. https://doi.org/10.1523/JNEUROSCI.6470-10.2011

Ziv, Y., Burns, L. D., Cocker, E. D., Hamel, E. O., Ghosh, K. K., Kitch, L. J., Gamal, A. El, & Schnitzer, M. J. (2013). Long-term dynamics of CA1 hippocampal place codes. Nature Neuroscience, 16(3), 264–266. https://doi.org/10.1038/nn.3329

Znamenskiy, P., & Zador, A. M. (2013). Corticostriatal neurons in auditory cortex drive decisions during auditory discrimination. Nature, 497(7450), 482–485. https://doi.org/10.1038/nature12077