Spring 5-15-2023

Mapping Brain Development and Decoding Brain Activity with Diffuse Optical Tomography

Kalyan Tripathy

WASHINGTON UNIVERSITY IN ST. LOUIS

Division of Biology and Biomedical Sciences
Neurosciences

Dissertation Examination Committee:
Joseph Culver, Chair
Tamara Hershey
Jin-Moo Lee
Bradley Schlaggar
Christopher Smyser

Mapping Brain Development and Decoding Brain Activity with Diffuse Optical Tomography

by

Kalyan Tripathy

A dissertation presented to
Washington University in St. Louis
in partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2023
St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

Acknowledgments

The work described in this dissertation would not have been possible without exceptional support from many mentors, collaborators, friends, and family. Words cannot do justice to all the gratitude that I owe them, and it would fill another dissertation to even just name everyone who has had an impact on my development as a scientist so far, but I would like to begin by thanking some of the people who contributed most directly to my graduate training.

First, I must thank my primary research mentor, Dr. Joseph Culver. There are a lot of things I admire about Joe, perhaps most of all how much he genuinely cares about his students. Being a wildly successful scientist imposes many demands on his schedule, but he always had my back and was very generous with his time when I really needed his guidance, even if that meant meeting late into the evening or working on a paper draft through the weekend. He always encouraged me and coached me to pursue opportunities to present my research, learn about others' work, and network through domestic and international scientific conferences that greatly enriched my graduate training experience. He is also very welcoming of different backgrounds, as evidenced by his multidisciplinary and culturally diverse lab. Having joined the lab with a degree in biology and primarily wet lab research experience, I would have never thought that I had it in me to build brain imaging hardware in a machine shop or write thousands of lines of code, but Joe always encouraged me to develop new skills and I am extremely grateful for that. It was always enlightening to witness Joe's mathematical and engineering approach to problem solving. I still marvel at how he draws inspiration for developing brain imaging technology from diverse sources like ski boots and orthotics, and hope that I can similarly think outside the box in my future research endeavors. In addition, it made a world of difference to my health and

happiness in grad school that Joe fosters an incredibly inclusive, supportive, and friendly lab environment. Working with him was genuinely fun – besides the science itself being stimulating, meetings always included jokes and appreciation awards and sometimes even costumes, we went to several concerts as a lab that expanded my musical horizons, and I even got to play music with Joe as we formed a lab band called the Tomo Sapiens. So thank you for everything, Joe – this gig was truly a blast and I am already looking forward to the reunion tour.

I was fortunate to also rotate with and be co-mentored by Dr. Bradley Schlaggar before he was recruited away from Wash U, and our regular meetings over a couple of years significantly shaped my thinking as a neuroscientist. In fact, the section of Chapter 5 that is currently titled "Models of plasticity" was briefly named "An ode to Brad", until I realized that this was putting far too much pressure on me to write it more perfectly than I was capable of doing. I am extremely grateful for all the time that Brad spent mentoring me over the years, including advising me in the research lab, sharing genuine guidance regarding training and career decisions, and inspiring me with his compassion towards patients as I shadowed him in the clinic. Brad was actually also one of the people who recommended I consider rotating in Joe's lab. During the subsequent co-mentorship, it was always exciting to hear Brad and Joe bounce ideas off each other and see some of their conversations grow into new grant proposals and research projects. I aspire to one day bring Brad's combination of deep clinical and scientific insight to research collaborations as a physician-scientist. And I hope I continue to hear Brad's voice echo in my head sometimes, advising me to critically consider methodological rigor in data analysis, to contemplate the basic neuroscience underlying imaging findings, and to think compassionately about patients in the clinic and research lab alike.

The other professors on my thesis committee – Dr. Tamara Hershey, Dr. Jin-Moo Lee, and Dr. Christopher Smyser – were also extremely supportive. During thesis updates, they always provided insightful, constructive feedback that helped me think more deeply about my data, refine my presentations, move projects forward, and stay on track to graduate. Tammy was also incredibly approachable as my program director and committee chair. She always asked a lot of thoughtful scientific questions during update meetings, but I also appreciated how proactively and caringly she always checked in about how I was really doing after updates were officially over. Jin-Moo always provided a welcome mix of genuine encouragement about research progress and realistic advice about next steps. For example, I remember how he helped me reshape an overly ambitious initial proposal to conduct a clinical study of children with strabismus into the more feasible thesis proposal that I ended up pursuing in this dissertation. Chris provided valuable regular guidance as a collaborator on the pediatric imaging project. I appreciate how patiently he stuck by Joe and me and continued to advise us as we stumbled through troubleshooting the challenges of developing a new optical neuroimaging systems for toddlers. I look forward to the possibility of working with all of my committee members again in future research projects, as well as hopefully working with Dr. Lee and Dr. Smyser in the clinic as a medical student during the next two years.

In addition to my official thesis committee, several other professors provided invaluable mentorship during my thesis research. Dr. Adam Eggebrecht deserves special mention for working closely with me as a collaborator and unofficial advisor on various projects. Adam was always incredibly generous with his time as he taught me fundamentals principles of biomedical optics, helped troubleshoot numerous instrumentation and data analysis issues, and provided insightful feedback on research updates, presentations, and manuscripts. He also played some

wicked bass lines for the Tomo Sapiens. I also learned a lot through collaborations with several other members of the Culver Lab. Andrew Fishell helped me find my footing during my rotation in the lab. Zachary Markow was a fantastic collaborator who brought incredible mathematical insight to the decoding projects as well as invaluable assistance with troubleshooting technical difficulties. Morgan Fogarty really helped salvage and accelerate the movie data analysis efforts through painstaking validation. Arefeh Sherafati shared helpful advice and resources for handling motion artifact and other data processing challenges. Sean Rafferty and Edward Richter helped immensely with developing new imaging systems, and Mariel Schroeder, Alexa Svoboda, Tracy Burns-Yocum, and Katie Mansfield helped immensely with data collection. Lindsey Brier, Rachel Rahn, Hunter Banks, Inema Orukari, and Paul Lin provided frequent entertainment and occasional help from the mouse arm of the lab. All the other members of the Culver lab were amazing colleagues as well. And several collaborators in the Optical Radiology Lab, the Neuroimaging Laboratories, and elsewhere at Wash U also kindly lent their time to various helpful and enjoyable intellectual conversations, including Dr. Abraham Snyder, Dr. Adam Bauer, Dr. Muriah Wheelock, Dr. Broc Burke, Dr. Jason Trobaugh, Dr. Samuel Achilefu, Dr. Mark Anastasio, and Dr. Ana María Arbeláez.

I would also like to thank my mentors from previous research endeavors for all of the time and effort they invested in training me prior to my thesis work. Dr. Virginia Lee and Dr. John Trojanowski offered me my first ever research position when I was a freshman at the University of Pennsylvania, and I enjoyed working with them so much that I stayed in their lab for 5 years. I will always be thankful to John and Virginia, as well as my other mentors (especially Dr. Scott Pesiridis, Dr. Linda Kwong, and Dr. Adam Walker) and colleagues in their lab at Penn, for helping me fall in love with research so much as to pursue a PhD in addition to medical school.

Dr. Randall Bateman (together with his kind and helpful lab members including post-doc Dr. Norelle Wildburger) almost convinced me to keep studying neurodegenerative diseases during my rotation with him, before I decided I should switch things up and study neurodevelopment instead. But I still loved hearing how Randy thought about biological research questions, and he certainly helped inspire me to pursue human subjects research during graduate school and perhaps for the rest of my career. Dr. Steve Petersen mentored me alongside Brad during my rotation in their lab and also taught me to think critically about neuroscience questions and neuroimaging data. The other members of the Petersen-Schlaggar lab, including Dr. Evan Gordon, Tim, Tunde, Ashley, and Ben were also very helpful as I learned to code and think about imaging data. I am also grateful for all the time, research insight, and training advice I received along the way from other professors, including Dr. Deanna Barch, Dr. David Holtzman, Dr. Lawrence Tychsen, Dr. Beau Ances, the members of my qualifying exam committee – Dr. Nico Dosenbach, Dr. Chad Sylvester, and Dr. Larry Snyder – and all the other professors who taught me during graduate school.

Outside of the research environment, I owe a debt of gratitude to all my other teachers through high school, college, medical school, and graduate school for nurturing my love of science and academia in the classroom. During grad school, I also had the pleasure of helping to teach neuroscience to first-year medical students as well as several gifted high school students. I learned a lot about neuroscience and about teaching from the course directors and lecturers as well as from my students themselves. I would like to especially thank Dr. Krikor Dikranian, Dr. David Van Essen, and Dr. Timothy Holy for being particularly enthusiastic and generous with their time during these mentored teaching experiences.

My graduate research and training have also relied on financial and administrative support from various sources. The Medical Scientist Training Program (MSTP) and the Neuroscience Program of the Division of Biological and Biomedical Sciences (DBBS) at Wash U admitted me into their structured training programs and provided me with financial support. I received advanced interdisciplinary training and a year of financial support through a Cognitive, Computational, and Systems Neuroscience (CCSN) Fellowship from the McDonnell Center for Systems Neuroscience. And while my citizenship status precluded me from applying for an NIH training grant, my work was nevertheless funded by research grants awarded to my mentors Joe Culver and Brad Schlaggar by the NIH (including grants U01EB027005 and R01NS090874) and the Bill and Melinda Gates Foundation (grant OPP1184813), among other sources. Besides the financial support, I am also extremely grateful for all the program directors and administrators who granted me all these wonderful training opportunities and guided me through fulfilling my associated responsibilities – Dr. Wayne Yokoyama, Brian Sullivan, Christy Durbin, Liz Bayer, Linda Perniciaro, Clarissa Craft, Tammy and Tim, Larry and Erik, Sally Vogt, Laura Williams, Carmen Horn, and others. And on the topic of invaluable administrative support, none of my thesis committee meetings could have been scheduled without incredibly patient assistance from Dana Wilhelm, Francesca Ryan, Beth Beato, Mary Dick, and Dr. Chris Smyser himself.

Before ending with some personal thank-yous, I must give a huge shout out to all the research participants who generously provided their time and invaluable data. This included several especially dedicated participants who returned for numerous repeat scans in order to serve as highly sampled individuals, and many wonderful children and families who really livened up my day while participating in the pediatric neuroimaging study.

My journey through grad school would not have been the same without the company and support of many amazing friends. I can't name them all here, but I also can't go without naming some of them. Lindsey and Dillan have been my medical school classmates, my graduate school classmates, my lab mates, my band mates, my travel partners, and constant sources of hilarious entertainment and vital emotional support. My other MSTP classmates and Neuroscience classmates have similarly enriched my graduate education with rich intellectual discourse. My other bandmates – Chris, Chuner, Ethan, and Rachel from The Package, and Adam, Adam, and Joe from the Tomo Sapiens – and many other musical collaborators have also helped me find meaningful outlets for creative expression. Amanda, Turja, Ameya, Akhilesh, Sai, and Sandeep, among others, have continued to be wonderful friends as they were through college, both visiting me in St. Louis and checking in from afar. I have also made many close friends here in St Louis over the years, including Sarah, Alex, Tirth, Lindsey, Kristina, and Lee, who have all provided invaluable emotional support. And I couldn't have made it through the pandemic or the last several months of frantic writing without the constant support from Anna. All these people have given me a family and a home away from home.

In Indian culture, we believe that nothing is more important than the blessings of our elders and the support of our families. I am so fortunate to have had incredible family support throughout my life and education – like from my cousin Leena Nani, who was the first in our family to move stateside and gave me a home to visit for the holidays for the first few years after I moved here; my other cousin Ruchi Nani back in Bengaluru who has been mailing me handwritten letters about how much I mean to her on Rakshabandhan (the festival celebrating the brother-sister relationship) every year for the last 30 years; and my grandmother in Bhubaneshwar who prays for the entire family's wellbeing every day and put in a special word for me today as she does

every time I have a big exam or assignment due. It would take another thesis to mention everyone – during my last visit home three years ago, I met no less than a hundred loving relatives between two casual family reunions – but I have to give special thanks to three people who have been there for me through every single day that I can remember: my father, mother, and sister. My father is in fact the one who first sparked my interest in neuroscience through his own fascination with the brain as a neurosurgeon, and he also inspired me deeply with his journey getting there from humble beginnings growing up in a poor household in a small town in India. He showed me what it meant to work hard, remain resolute in the face of adversity, and maintain one's cool and a sense of humor through it all, which proved to be invaluable skills over the course of grad school. He also showed me how to work a crowd, which was another skill I attempted to hone over grad school during scientific presentations and singing performances alike. Neither he nor I would be anywhere without my mother though. She too inspired me with her sincerity and dedication, from medical school through motherhood to her recent rise as a locally celebrated singer. Choosing to stay at home as she raised her children, she taught us principles that are as important in scientific research as they are in the rest of life, like integrity, humility, and respect for others. She also strongly emphasized the value of education throughout our childhood and advocated avidly for us to pursue the best learning opportunities available to us, even if that meant dipping into savings or debt or came at the cost of going years at a time without seeing her children in person. And she still calls us almost every single day to check in on how we are doing and remind us that she loves us from the other side of the world. Finally, I must thank my little sister and lifelong best friend, Amy. I feel incredibly lucky to have shared several of my grad school experiences with her, including the computational neuroscience class that we took together while she was an undergraduate at Wash U (which actually helped

shape the second chapter of my thesis), the couple of summers we spent living together again after years of being apart from family, and the couple of months that we worked from home together this past year and helped each other weather the emotional storms of the pandemic. Besides grad school, Amy has also been there for me through every other journey so far, cheering me on through the highs, providing emotional support through the lows, and inspiring me all the way with her own diligence, resilience, brilliance, and love. I owe immeasurable gratitude to her and to all my family for keeping me grounded, motivated, and happy.

Finally, I would like to thank you, the reader, for your time and interest in learning about my story and my work. Scientists will always explore and artists will always create, but it is for the audience that performers put on their best show.

Kalyan Tripathy

Washington University in St. Louis

May 2023

Dedicated to Mom, Dad, Amy,
and all the other dear family, friends, and mentors
who have inspired me, supported me, and believed in me.

ABSTRACT OF THE DISSERTATION

Mapping Brain Development and Decoding Brain Activity with Diffuse Optical Tomography

by

Kalyan Tripathy

Doctor of Philosophy in Biology and Biomedical Sciences

Neurosciences

Washington University in St. Louis, 2023

Professor Joseph P. Culver, Chair

Functional neuroimaging has been used to map brain function as well as decode information from brain activity. However, applications like studying early brain development or enabling augmentative communication in patients with severe motor disabilities have been constrained by extant imaging modalities, which can be challenging to use in young children and entail major tradeoffs between logistics and image quality. Diffuse optical tomography (DOT) is an emerging method combining logistical advantages of optical imaging with enhanced image quality. Here, we developed one of the world's largest DOT systems for high-performance optical brain imaging in children. From visual cortex activity in adults, we decoded the locations of checkerboard visual stimuli, e.g. localizing a 60° wedge rotating through 36 positions with an error of 25.8±24.7°. Using animated movies as more child-friendly stimuli, we mapped reproducible responses to speech and faces with DOT in awake, typically developing 1-7 year-old children and adults. We then decoded with accuracy significantly above chance which movie a participant was watching or listening to from DOT data. This work lays a valuable foundation for ongoing research with wearable imaging systems and increasingly complex algorithms to map atypical brain development and decode covert semantic information in clinical populations.

PREFACE

Over the course of my PhD, I have had the fortune of working with many amazing mentors, peers, and students on a range of collaborative, interdisciplinary research projects. This dissertation focuses on my primary efforts yielding first author manuscripts.

Chapter 2 has already been published as:

> Tripathy, K., Markow, Z.E., Fishell, A.K., Sherafati, A., Burns-Yocum, T.M., Schroeder, M.L., Svoboda, A.M., Eggebrecht, A.T., Anastasio, M.A., Schlaggar, B.L., Culver, J.P., 2021. Decoding visual information from high-density diffuse optical tomography neuroimaging data. Neuroimage 226, 117516. https://doi.org/10.1016/j.neuroimage.2020.117516

That chapter hence contains material copyrighted by the authors that is reused here in accordance with the publisher's copyright license.

The other chapters currently contain no copyrighted material.

Kalyan Tripathy

Washington University in St. Louis

May 23rd 2021

“It just takes some time”

- Jimmy Eat World

Chapter 1: The promise of high-density diffuse optical tomography for developmental cognitive neuroscience and neural decoding

The evolution of functional neuroimaging methods has fueled exciting progress in neuroscience research, revealing both how information is mapped in the brain and how it can be decoded from measurements of neural activity [1–4]. However, mainstream imaging modalities entail tradeoffs between image quality and the logistics of image acquisition, constraining applications like mapping early brain development in typical and atypical populations or creating augmentative communication interfaces for children with cerebral palsy [5,6]. High-density diffuse optical tomography (HD-DOT) is an emerging functional neuroimaging method that combines logistical advantages of optical imaging with improved image quality [6]. This dissertation focuses on developing HD-DOT for mapping brain function in awake young children and on evaluating decoding of brain activity measured with HD-DOT. To provide the necessary framework, the opening chapter discusses the significance of studying early functional brain development and neural decoding, the challenges posed by extant neuroimaging methods, and the way in which

HD-DOT functions to address limitations of other imaging approaches, culminating in a roadmap for the remainder of the dissertation.

1.1 Importance and challenges of mapping brain function in early development

The first few years of life comprise a critical period for the development of myriad behaviors, during which atypical experiences or biological processes can cause lasting negative neurological outcomes, while heightened neuroplasticity can conversely be harnessed to counteract atypical development [7–12]. However, we have a limited understanding of the changes in brain function underlying these behavioral and clinical observations in humans due to a relative dearth of functional neuroimaging research in early childhood. Animal studies have certainly provided invaluable insight into the development of sensory processing and many behavioral responses [13–16], but not all human behaviors are easily modeled in animals. Human language (distinguished from other forms of communication by its complex structure and compositional nature) is a prime example of a behavior that cannot be readily studied in animals [17], exhibits dramatic changes during the first few years of life [18,19], and yet has not been the focus of a proportional number of imaging studies during this epoch. In fact, a systematic review of over 20 years of task fMRI research on language development identified only 1 study in children younger 4 years of age [20], which imaged passive listening during sleep [21]. This void

in the literature can be attributed to the limitations of common neuroimaging modalities as well as challenges with child compliance during traditional imaging tasks.

Functional magnetic resonance imaging (fMRI), the mainstay of human functional neuroimaging, most commonly tracks changes in a blood oxygenation level dependent (BOLD) signal evoked by neural activity, exploiting differences in the magnetic properties of oxygenated and deoxygenated hemoglobin [5]. However, measuring this signal requires participants to lie inside the cramped bore of a large magnet, surrounded by loudly vibrating electromagnetic coils – an altogether intimidating environment for young children. Furthermore, any head motion relative to the magnetic fields creates problematic artifacts in the data. Therefore, an abundance of motion artifact and the challenges of ensuring child comfort and compliance have plagued fMRI studies of early development [22]. A growing number of fMRI studies have imaged sleeping infants to minimize compliance and motion issues and to probe the functional organization of the developing brain at rest [23–25], but sleep alters brain networks [26,27] and also does not allow investigation of how the brain is actively engaged during various task states. Reports of fMRI studies in awake children below the age of 4 years are rare [20,25,28].

Alternative methods for functional neuroimaging in humans offer logistical advantages conducive to studying children but at the expense of significantly inferior image quality. For instance, functional near-infrared spectroscopy (fNIRS) involves shining harmless near-infrared light into a participant's head and detecting scattered photons as they return to the head surface after interacting en route with hemoglobin in the superficial cerebral vasculature [29]. Differences in the optical properties of oxygenated and deoxygenated hemoglobin are used to

track changes in blood oxygenation as a surrogate marker of neural activity using a wearable imaging cap amid a silent, open imaging environment. FNIRS is conducive to imaging children as the imaging cap can be worn while children sit comfortably next to their caregivers or in their laps. However, fNIRS has traditionally used sparse arrays of small numbers of light sources and detectors (collectively referred to as optodes), resulting in image quality issues such as low spatial resolution, signal localization artifacts varying with optode geometry , and contamination of cerebral signals by those from superficial tissues like the scalp and skull [30,31]. In fact, developmental research studies using fNIRS frequently compare differences between ages, populations, and task conditions with reference to optode positions on the scalp [32,33], which are not a particularly biologically relevant unit of comparison and can vary considerably between studies using different imaging systems.

In addition to these issues with imaging technology, developmental studies must also address challenges with task compliance in children. Common task paradigms for mapping brain function in adults are often too mundane and repetitive to keep children cooperative and engaged [22]. It has been shown that presenting movies can increase compliance and reduce head motion in young children [34], and this strategy has been utilized by several fMRI studies of children mostly over 4 years of age [28]. Furthermore, studies in adults have shown that free viewing of movies also evokes reproducible patterns of brain activity that can be used to map responses to a variety of movie features in parallel through feature regression [35–37]. Combining this naturalistic viewing paradigm with an imaging modality that offers a naturalistic, open imaging environment like fNIRS could facilitate child comfort and compliance in studies mapping brain

function across early childhood. However, the complex, distributed response patterns evoked by movies [35] are unlikely to be extensively or reliably captured by traditional fNIRS with its sparse sampling and associated image quality issues. The field of developmental cognitive neuroscience could hence benefit greatly from an imaging modality that offers the child-friendly logistics of fNIRS paired with image quality closer to that of fMRI.

1.2 Significance and limitations of neural decoding research

While much neuroscience research has focused on mapping sensory stimuli, behavioral responses, or cognitive processes within the brain, a growing body of work has underscored the value of the reverse analysis known as neural decoding [2,4], which attempts to decipher information from measurements of brain activity (Figure 1).

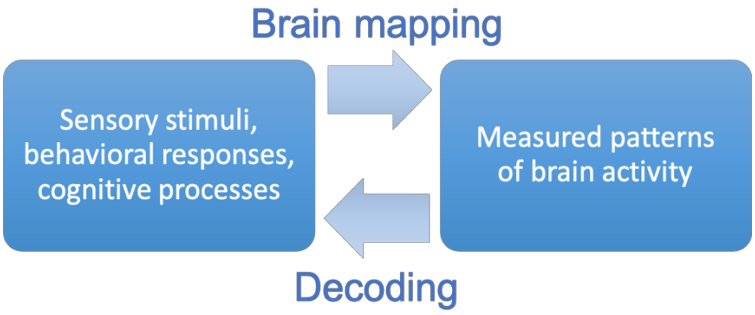

Figure 1.1: Conceptual schematic of brain mapping and decoding

Some studies using extant recording methods have accurately decoded elaborate information from their sampling of brain activity. For example, fMRI studies have been able to reconstruct images and dynamic visual scenes from visual cortex activity [38–41], while studies using electrocorticography (ECoG) have been able to synthesize intelligible speech from various brain regions including areas associated with hearing, articulation, and language [42,43]. Compelling potential applications of this work include providing a means of augmented communication to patients who are unable to speak or move effectively as a result of various neurological conditions like strokes, neurodegenerative diseases, and neurodevelopmental disorders. One could also imagine how neural decoding could be applied more broadly to change how we all interface with technology and with each other.

However, translation of decoding research beyond the laboratory has been constrained by various logistical challenges associated with the neural recording methods used. MRI scanners require large magnets and large amounts of liquid helium to run, and hence are notoriously cumbersome, expensive, technically demanding to operate and maintain, and confined to large clinical and research facilities. Due to their strong magnetic fields, MRI scanners are also contraindicated for individuals with certain implanted medical devices or other forms of metal in the body, cannot be used around objects containing ferromagnetic materials, and can interact unfavorably with electronics [44,45]. Meanwhile, ECoG requires implantation of electrode arrays through invasive neurosurgery. Other non-invasive and more portable imaging technologies like fNIRS have been used in some decoding studies, but sparse sampling has restricted the granularity and accuracy with which decoding can be accomplished. Far from

reconstructing naturalistic visual and auditory information, fNIRS decoding research has largely been limited to binary or 4-way classifications of task states with variable accuracy [46–51]. The logistical advantages of fNIRS need to be combined with greater space-bandwidth in order to better facilitate real world applications of neural decoding.

1.3 Principles and potential of high-density diffuse optical tomography (HD-DOT)

HD-DOT is an emerging functional neuroimaging technique that uses dense arrays of light sources and detectors to capture many overlapping measurements of light absorption and scattering that can be used to localize changes in blood oxygenation more accurately than fNIRS [30,52,53]. HD-DOT thereby combines logistical advantages of optical imaging with enhanced image quality.

Some of the principles underlying HD-DOT imaging are the same as for fNIRS, as detailed elsewhere [6,29] and summarized briefly here. Near-infrared light can travel effectively through most biological tissues enabling non-invasive imaging. As a result of photon scattering, light incident on the head tends to travel in an arc. Furthermore, light can arc through superficial tissues like the scalp and skull into the cerebral cortex and back out to the head surface where light levels can be measured. The main substance that absorbs near-infrared wavelengths of light passing through tissues is hemoglobin in blood, and oxygenated and deoxygenated hemoglobin have different absorption spectra. Therefore, emerging light levels measured at the head surface

convey information about the oxygenation state of blood in the vasculature of the cerebral cortex (though also in blood vessels of the scalp and skull). Changes in oxygenated hemoglobin levels caused by the local hemodynamic response to neural activity can thus be measured using noninvasive optical imaging.

HD-DOT improves image quality relative to fNIRS in several ways. Traditional fNIRS systems generally have 1-20 sources and 1-20 detectors each separated by a constant distance of 2-4 cm to obtain 1-50 source-detector pairs of optical measurements or channels [54]. In contrast, a recently developed HD-DOT system used 96 sources and 92 detectors to capture over 1200 usable measurements across a range of source-detector separations from 1.3 to 4.7 cm [53]. The large number of overlapping measurements enables more accurate localization of signals with increased lateral resolution [30], increased signal-to-noise ratio [55,56], and decreased artifacts associated with sparse optode geometry [30]. The range of source-detector distances provides depth sensitivity, as short-range measurements sample more superficial tissues while longer range measurements additionally convey information from deeper structures like the cerebral cortex [52,57]. The mean signal from short-range measurements can thus be regressed out of the long-range ones to make them more brain-specific, a strategy known as superficial signal regression [58]. Tomographic reconstruction of HD-DOT data based on biophysical modeling of light transmission through the tissues of the head further facilitates localization of signals specifically to the cortex in three-dimensional maps of brain activity [52,58,59].

As a result of these improvements, HD-DOT has proven an effective means of imaging both localized and distributed patterns of cortical activity, producing maps comparable to those

obtained with fMRI during visual stimulation, language tasks, and resting state functional connectivity analysis [53,60]. As a limitation, sulcal depths, subcortical nuclei, and other deep structures lie beyond the field of view, which is restricted by light penetration to <20mm below the head surface [6]. However, many task responses and functional brain networks can be mapped within the superficial cortex [6,53]. Importantly, in contrast to fMRI, HD-DOT achieves its image quality while maintaining an open and silent imaging environment, a lack of contraindications, compatibility with other devices, significantly lower equipment and operational costs, and scope for portability and wearability [6]. So far, these benefits have been utilized by HD-DOT studies to map brain activity in various populations and settings not conducive to fMRI, including awake infants [56], 7-10 year-old children in low resource settings [61], stroke patients at the bedside [62,63], and Parkinson's disease patients with electrodes implanted for deep brain stimulation [53].

The logistics of HD-DOT also ought to be conducive to imaging awake young children between infancy and 7 years of age, as well as to pursuing potential applications of neural decoding. However, prior to the work in this dissertation, HD-DOT had yet to be developed or applied for imaging 1-6 year-old children. In addition, the feasibility and performance of decoding brain activity measured with HD-DOT had yet to be evaluated.

1.4 Specific objectives of the dissertation

In the context of the prior literature and unaddressed questions introduced above, this dissertation focuses on the following objectives:

1. Evaluating the feasibility and performance of neural decoding with HD-DOT data (Chapter 2): For a first study of decoding, we attempted to classify the positions of checkerboard stimuli from HD-DOT recordings of brain activity in participants viewing them. This paradigm allowed us to control the stimulus state, accurately reference ground truth when evaluating decoding accuracy, and capitalize on the accurate retinotopic mapping previously established with HD-DOT for decoding [52,60]. We quantified the performance of decoding on a single-trial basis across multiple imaging sessions, participants, and levels of stimulus complexity.

2. Developing and applying HD-DOT imaging methods for mapping brain function in awake young children during naturalistic movie viewing (Chapter 3): We built a new HD-DOT system geared towards enhancing both image quality and child comfort. Furthermore, building on recent work mapping responses during movie viewing in adults with HD-DOT [64], we characterized a library of children's movie clips and refined associated data processing pipelines, validating them as effective tools for mapping brain function using the new system in adults. Finally, we combined the optimized imaging system, stimulus library, and analysis pipelines to map responses to speech in a group of awake 1-6 year-old children.

3. Evaluating the feasibility and performance of decoding naturalistic auditory and visual information from HD-DOT data (Chapter 4): Given both effective decoding of visual stimulus position and effective mapping of responses to audiovisual movies with HD-DOT, we proceeded to study decoding of more complex, naturalistic, and multisensory

stimuli. We repurposed the template matching decoding algorithm developed in Chapter 2 and the movie viewing data set collected in adults in Chapter 3 to investigate whether auditory and visual movie content could be decoded from HD-DOT data. We further assessed how factors like data quality, trial duration, and template number influenced decoding.

4. Synthesizing the above avenues of investigation and discussing implications and future directions for optical neuroimaging method development, developmental cognitive neuroscience research on models of plasticity, and potential applications of neural decoding (Chapter 5)

“Roobaroo roshni hai”
(“I am face to face with the light”)

- Prasoon Joshi, A. R. Rahman

Chapter 2: Decoding visual information from high-density diffuse optical tomography neuroimaging data

2.1 Introduction

While much cognitive neuroscience research has focused on mapping brain regions that are activated while participants perform tasks, recent work has emphasized the value of the reverse analysis known as decoding, i.e., deducing task information from recordings of brain activity [2,4]. Potential applications of decoding range from building brain-computer interfaces that could restore movement or communication in paralyzed patients [46,65] to recreating scenes from brain activity measured during visual experiences [38–40], imagination [41,66], memory tasks [67], or dreams [68]. However, decoding applications are constrained by the limitations of established neuroimaging modalities. For instance, many decoding research studies so far have relied on electrocorticography (ECoG), which is invasive, or functional magnetic resonance imaging (fMRI), which is not conducive to imaging in certain populations and applications due to its logistics [2,29,43]. Functional near-infrared spectroscopy (fNIRS) addresses some of these

issues as a noninvasive, portable, optical imaging method that supports imaging in an open environment [29]. However, the low channel count of traditional fNIRS limits its spatial resolution, coverage, and image quality, thereby constraining the complexity and precision of decoding [30,46,48]. High-density diffuse optical tomography (HD-DOT) is an emerging optical neuroimaging modality that uses a dense array of light sources and detectors to capture thousands of overlapping measurements, overcoming several limitations of traditional fNIRS and more closely matching fMRI in extensive brain mapping studies [30,52,53,58,60,69]. However, decoding of brain activity as measured by HD-DOT has yet to be evaluated. Here, we examine the feasibility of decoding visual stimulus information from evoked brain activity measured with HD-DOT, assessing accuracy, reproducibility, and granularity obtainable.

Prior research using other neuroimaging modalities has made a compelling case for the clinical and neuroscientific utility of decoding [2,4,70]. For one, decoding could provide a means of augmented communication for patients who cannot speak or move effectively as a result of various neurological disorders including, but not limited to, stroke, neurodegenerative diseases, and developmental conditions like cerebral palsy. The complexity of decoding-based communication has so far ranged from obtaining yes-no responses using traditional fNIRS noninvasively [46] to generating comprehensible speech from ECoG arrays in epilepsy patients undergoing neurosurgical treatment [43]. Neuroimaging signals have also been decoded to drive motor prosthetics for amputees and paralyzed patients [70]. Success with decoding motor signals has ranged from noninvasive but coarse decoding of binary movement direction using fNIRS in healthy participants [51] to detailed control of robotic arms using intracortical electrodes in

paraplegic patients [65]. Furthermore, cognitive neuroscience studies have decoded naturalistic images [38], movies [39,40], speech [71,72], music [73], dreams [68], and semantic content [74] from fMRI data and provided a powerful window into human brain function [4].

While promising, these applications of decoding are currently constrained by the limitations of mainstream neuroimaging modalities. The advanced decoding achieved using ECoG has remarkable performance, but implantation of the electrode arrays requires invasive neurosurgery and is therefore limited to small clinical populations [43]. Meanwhile, fMRI is noninvasive, but is still challenging for certain populations, such as young children, and contraindicated in others, such as patients with implanted metallic devices. In addition, it is not feasible for patients to regularly use cumbersome, technically demanding, and prohibitively expensive MRI scanners for applications such as longitudinal communication.

In contrast, fNIRS is much more portable and cost-effective, allowing widespread use and longitudinal imaging in more natural settings [29]. In addition, fNIRS is without major contraindications: the instrumentation is compatible with implants, and the open imaging environment is well suited to imaging children with a parent nearby for comfort. However, the sparse optode arrays and low channel counts used in traditional fNIRS compromise spatial resolution, coverage, and image quality, which are likely to limit the information available for accurate and detailed decoding [30,61].

HD-DOT utilizes a comparatively dense arrangement of light sources and detectors to collect ten-fold to a hundred-fold more measurements than traditional, sparse fNIRS. The numerous

overlapping channels of HD-DOT enable higher spatial resolution, tomographic reconstruction, superficial signal regression, and a reduction in signal localization artifacts [30,58,61]. As a result, HD-DOT has been found to have a similarly high signal-to-noise ratio and high spatial concordance compared to fMRI in brain mapping studies using phase-encoded visual stimuli [60,69], hierarchical language tasks [53,75], and resting state functional connectivity analysis [53]. By combining logistical advantages of optical neuroimaging with improved space-bandwidth product (i.e., the field-of-view divided by the point-spread-function, or roughly the number of independent voxels) and image quality, HD-DOT holds the potential to support detailed, accurate decoding in naturalistic environments and advance the clinical and neuroscience applications of decoding.

In the current study, we evaluate the feasibility and performance of visual decoding with HD-DOT, focusing on decoding the positions of checkerboard stimuli. The visual system is a particularly useful model for proof of principle, due to the elaborate, consistent, and well-characterized organization of visual features in neuroanatomical space [76]. Indeed, early fMRI decoding research also focused on visual decoding and gradually progressed from decoding fundamental visual features such as stimulus position [77] and line orientation [78] to eventually performing intricate reconstructions of naturalistic images and movies [38–40]. Furthermore, prior work with HD-DOT has validated retinotopic mapping with HD-DOT against the gold standard of fMRI [52,60], making retinotopic decoding a promising initial goal.

Herein, we first establish, in healthy adults, the feasibility of binary visual decoding with HD-DOT using a template matching strategy, and we evaluate its sensitivity, specificity, and

accuracy across a range of thresholds using receiver operating characteristic (ROC) analysis. We then assess the reproducibility of this binary decoding across multiple imaging sessions and multiple participants. Finally, we extend the analysis of decoding performance to non-binary cases. We use phase-encoded retinotopic stimuli to evaluate the feasibility, accuracy, and replicability of 18-way and 36-way classifications of stimulus position across different parts of the visual field at individual time points without block-averaging test data. These studies reveal that HD-DOT can allow sensitive, specific, and reproducible retinotopic decoding, encouraging future studies of more complex decoding paradigms and applications in atypical populations.

2.2 Methods

2.2.1 HD-DOT imaging

This study aims to investigate the decoding performance of optical neuroimaging using a HD-DOT imaging array with an increased space-bandwidth product relative to traditional fNIRS. Imaging was performed using a previously characterized continuous-wave HD-DOT system that illuminates the back and sides of the head with 750nm and 850nm light through a grid of 96 LED sources interlocked with 92 APD detectors, collectively yielding over 1200 usable source-detector measurements per wavelength. The weight of the fiber optics is supported by an extruded aluminum frame and two suspended wooden rings. Fiber tips contact the head via a custom-built cap that spaces the optodes 13 mm apart (hence with first-nearest through fourth-nearest source-detector separations of 1.3, 3.0, 3.9 and 4.7 cm) across the posterior and lateral

surfaces of the head [53]. During each cap fit, the participant's hair was first parted and tied if necessary to minimize obstruction of light transmission within the field-of-view. The participant was then asked to sit in a chair placed below the HD-DOT cap, hold the front straps of the cap, and comb the cap's optodes through their hair and up against the scalp surface. Anatomical markers such as the tragus and inion were used to guide cap positioning across sessions and participants. Specifically, the vertical position of the cap was adjusted such that the lowest row of optodes contacted the head at the level of the inion, and the horizontal distance between reference optodes and the tragus was measured on each side and the positioning of the cap adjusted accordingly to ensure symmetry. The straps of the cap were then tightened and fastened. Real-time data quality metrics such as light and noise levels were used to guide any further optimization of the cap fit, e.g. combing any occluded optodes through any obstructing hair to ensure uniform optode-to-scalp coupling and light levels across the field of view. Photographs of the imaging cap placement were taken from both sides, from both upper and lower viewing angles, and used to confirm cap placement during data processing.

2.2.2 Data processing

Following data acquisition, data were pre-processed, reconstructed, and subjected to spectroscopy, as summarized below and detailed elsewhere [53].

Pre-processing: Raw light measurements were converted to log ratio time series using the temporal mean of each measurement as its relative baseline. Noisy channels with >7.5% variance across a run were excluded from further analysis, as this excessive variance was likely to reflect

non-physiological nuisance signals such as head motion as opposed to hemodynamic changes associated with brain activity [53]. High-pass filtering with a 0.02-Hz cutoff was performed to reduce long-term drift. Subsequently, superficial signal regression was performed by first averaging all first-nearest-neighbor measurements, which sample mostly the scalp, to use as an estimate of global systemic signals. This global superficial signal time trace was then subtracted out of every source-detector measurement time trace using linear regression. This approach has been shown to work in conjunction with HD-DOT to improve contrast-to-noise ratio, and the use of a single superficial signal regressor avoids both over-fitting and removal of brain activation signals [52,53,58]. Low-pass filtering (with a 0.5-Hz cutoff) removed residual pulse and other high-frequency noise, and data were then downsampled to 1 Hz [53].

Light modeling and reconstruction: Pre-processed data were reconstructed using an anatomically based light propagation model generated from the Montreal Neurological Institute (MNI) non-linear ICBM152 atlas [79–81], using Freesurfer for segmentation [82–84], NIRVIEW for mesh generation, finite element modeling with spring relaxation approaches for source and detector positioning, and the NIRFAST toolbox for determining a solution to the optical diffusion equation so as to model photon diffusion through the head [53,59,85]. The resulting sensitivity matrix A is the linear transformation for each time point between x, the vector of absorption coefficients at 750nm and 850nm within the brain volume, and y, the vector of relative changes in measured light levels at each wavelength detected at the head surface, as per the linear Rytov approximation:

$$(1)\ y = A\,x$$

This sensitivity matrix was inverted using Tikhonov regularization (λ_1=0.01) and spatially variant regularization (λ_2 = 0.1) [53]. The wavelength-dependent absorption and scattering coefficients (units mm^{-1}) for the five non-uniform tissue compartments were as follows: scalp ($\mu_{a,750}$ = 0.017; $\mu_{a,850}$ = 0.019; $\mu_{s,750}'$ = 0.74; $\mu_{s,850}'$ = 0.64), skull ($\mu_{a,750}$ = 0.012; $\mu_{a,850}$ = 0.014; $\mu_{s,750}'$ = 0.94; $\mu_{s,850}'$ = 0.84), cerebrospinal fluid ($\mu_{a,750}$ = 0.004; $\mu_{a,850}$ = 0.004; $\mu_{s,750}'$ = 0.3; $\mu_{s,850}'$ = 0.3), grey matter ($\mu_{a,750}$ = 0.018; $\mu_{a,850}$ = 0.019; $\mu_{s,750}'$ = 0.84; $\mu_{s,850}'$ = 0.67), and white matter ($\mu_{a,750}$ = 0.017; $\mu_{a,850}$ = 0.021; $\mu_{s,750}'$ = 1.19; $\mu_{s,850}'$ = 1.01) [60,64,86–88].

Spectroscopy: Relative changes in oxy- and deoxy- hemoglobin concentrations were calculated from the differential absorption image x at each time point through spectral decomposition:

$$(2)\quad \Delta C = E^{-1} x$$

where ΔC is a vector of oxy- and deoxy- hemoglobin concentration changes across the brain volume, and E is a matrix of extinction coefficients of oxy- and deoxy- hemoglobin.

2.2.3 Participants

While functional domains of the brain are generally common across individuals, there are readily quantified individual-specific differences in functional localization even after anatomical spatial normalization [89]. Because of this high inter-individual variability, it has been common for fMRI decoding studies to focus on 2-3 highly sampled research participants in order to amass

sufficient data per participant [38,39,90]. Here, we adopted a similar approach of collecting large amounts of data in a few participants to analyze decoding accuracy and reproducibility across multiple imaging sessions. We also collected smaller quantities of data across other participants to assess reproducibility across individuals. In total, 8 healthy adults participated in this study (age range 24-54 years, 7 female). In the first phase of the study (binary decoding), we collected extensive data in participant 1 across 10 imaging sessions, and evaluated reproducibility across n=5 individuals (participants 1-5). In the second phase of the study (non-binary decoding), we used pilot data from participant 1 to optimize decoding parameters, then analyzed 18-way and 36-way decoding extensively in participant 6, and finally evaluated reproducibility of the more complex 36-way decoding across n=3 individuals (participants 6-8). The participants' demographic information and the data that they contributed to each experiment are detailed in Table 2.1. We intentionally selected participants whom we expected to provide high-quality data, based on prior studies, as we aimed to evaluate the feasibility and performance of decoding without the confounds of poor data quality. Informed consent was obtained from all participants in accordance with the IRB protocol approved by the Human Research Protection Office at Washington University School of Medicine.

Table 2.1: Participant demographics and contributions to study

Participant Number	Age	Sex	Number of task runs contributed to study		
			Binary decoding	18-way decoding	36-way decoding
1	28	F	18	2	2
2	20	F	2	-	-
3	50	F	2	-	-
4	38	M	2	-	-
5	54	F	2	-	-
6	27	F	-	4	6
7	24	F	-	-	6
8	31	F	-	-	6

2.2.4 Stimulus protocols

Participants were imaged with HD-DOT while they performed multiple runs of three different types of checkerboard viewing tasks, adapted from prior retinotopic mapping studies [30,52,60,69]. The Psychophysics Toolbox 3 package for MATLAB was used to display the visual stimuli [91]. In all tasks, a checkerboard grid pattern was displayed, consisting of a black and white grid that reverses at 8 Hz in time against a 50% gray background, changing over time as follows:

Task 1 (binary decoding): Left- and right-sided checkerboard wedge stimuli provided an intuitive starting point to begin studying decoding as previous studies show that these two stimulus conditions produce distinct, reproducible, well-defined activation maps [52]. Participants were asked to maintain central fixation through multiple rounds of viewing a flickering checkerboard wedge for 10 seconds interspersed with rest periods of viewing just a fixation cross for 24 seconds. Longer inter-stimulus rest periods of 48s and 72s were used during a subset of the task runs to check whether this had any effects on stimulus response maps and decoding performance, but none were seen indicating that 24 seconds was sufficient to separate stimulus presentations. The checkerboard wedges extended over a polar angle of 70° and a radial angle of 2.5-10°. During each block, the checkerboard stimulus was presented on either the lower left quadrant or the lower right quadrant of the screen. Over the course of each task run, the stimulus was presented an equal number of times on each side in a pseudo-random order. The number of repetitions ranged from 5-8 depending on the duration of the inter-stimulus interval; fewer

repetitions were delivered when the inter-stimulus interval was extended beyond 24s to limit total run time.

Task 2 (18-way decoding): To create a more challenging decoding task with a greater number of targets, we presented participants with expanding/contracting checkerboard ring stimuli. Participants were asked to maintain central fixation while a flickering checkerboard ring expanded or contracted for 8 cycles through 18 positions on the screen at 2 seconds per position.

Task 3 (36-way decoding): For another evaluation of decoding with a greater number of targets, we also presented participants with rotating checkerboard wedge stimuli. Participants maintained central fixation while a flickering checkerboard wedge rotated through 10 revolutions, either clockwise or counterclockwise. The wedge subtended a polar angle of 60°, a radial angle of 2.5-10°, and rotated 10° at a time through 36 positions spanning 360° on the screen at 1 second per position (i.e., 10° per second, or 36 seconds per revolution).

2.2.5 Decoding by template matching

For an initial assessment of the feasibility of decoding HD-DOT data, we wanted to use a simple classification algorithm and hence chose a template matching strategy. For every decoding attempt, distinct task runs were used for training and testing. The training data were always derived from a single task run that included 5 to 10 rounds of stimulus presentation, block-averaged to construct oxyhemoglobin signal maps for different stimulus conditions. These mean activation maps served as “templates” of brain activity corresponding to each stimulus response,

in voxel space and spatially normalized to the MNI atlas as detailed in section 2.2. To create the templates at maximum signal-to-noise ratio, we used a data-driven, paradigm-specific time window to capture the peak signal. While the hemodynamic response is fixed, the signal measured is a convolution of the hemodynamic response function and the stimulus timing [92], with the latter varying across tasks. Longer presentations of the stimulus in each position delay the peak response [93]. For Task 1, in which stimulus presentations were most spread out, left-sided and right-sided template maps were constructed by block-averaging the signal across a 5-second time window starting 10 seconds after stimulus onset, as this coincided with the peak signal based on prior hemodynamic studies (Supplemental Figure S2.1A) [52,58]. For Task 2, the stimulus only spent 2 seconds in each position at a time, so each template was constructed by averaging data across blocks from two time points, 7 and 8 seconds after the checkerboard had passed through the position under consideration. This 7-8 second time window was empirically optimized using pilot data from participant 1 by evaluating decoding performance for a range of time delays and selecting the time delay that minimized decoding error (Supplemental Figure S2.1B). For Task 3, the stimulus rotated at 1 second per position, so each template was constructed by block-averaging data from a single time point across the 10 revolutions of the stimulus. Here, the optimal time delay was found to be 6 seconds (Supplemental Figure S2.1C-D).

Decoding with 2 templates: Initially, stimulus state was decoded at every time point to evaluate the feasibility of decoding at individual time points (Figure 2.1). This demonstrates the temporal quality of the decoding. To perform a statistical analysis of binary decoding performance

(Figures 2.2-2.4), we restricted the data to clearly independent trials, reasoning that individual time points were not all independent due to the temporal blurring of the hemodynamic response. Each stimulus presentation and each inter-stimulus rest period in the test data were treated as separate trials. A trial response was defined as the oxyhemoglobin signal map during the single frame 12 seconds after the onset of the stimulus presentation or the inter-stimulus interval (to coincide with the middle of the time window used to construct the stimulus response templates), in the same MNI atlas voxel space as the templates. We then evaluated binary decoding performance on a trial-wise basis. Each trial response was compared to both the left-sided and the right-sided templates by calculating a spatial Pearson correlation coefficient $r_{n,m}$ between the n^{th} template T_n and the oxyhemoglobin response map for the m^{th} trial S_m as follows, where both T_n and S_m were zero-mean-centered (i.e., mean(T_n)=0, mean(S_m)=0):

$$(3)\ r_{n,m} = \frac{(T_n)\cdot(S_m)}{|T_n||S_m|}$$

These correlation coefficients were compared to an adjustable threshold to determine the decoding output (D_m) for each trial m. The threshold value was initially set ad hoc for Figure 2.1 and later optimized using ROC analysis as described in section 2.6 and Figure 2.2. If all values of $r_{n,m}$ on the m^{th} trial fell below the threshold, the decoding output was set to $D_m = 0$ (i.e., decoded as rest). If $r_{n,m}$ rose above the threshold for only the n^{th} template, the decoding output was set to $D_m = n$. If $r_{n,m}$ rose above the threshold for multiple templates in one trial, the decoding output was determined by the template with the maximum correlation value $r_{n,m}$ during that trial:

$$(4)\ D_m = \underset{n}{\operatorname{argmax}}(r_{n,m})$$

Decoding with 18 and 36 templates: The template matching approach was similar for the non-binary decoding experiments, except that decoding involved a larger number of templates (either 18 or 36) and a decoding output was computed for every time point in the test run. Specifically, a spatial Pearson correlation $r_{n,t}$ was calculated between the n^{th} template T_n and the oxyhemoglobin response map S_t at the t^{th} time point for each of the 18 or 36 templates and every single time point in the test data:

$$(5)\ r_{n,t} = \frac{(T_n)\cdot(S_t)}{|T_n||S_t|}$$

Here, T_n and S_t were again first zero-mean-centered (i.e., mean(T_n)=0, mean(S_t)=0). The decoding output D_t at each time point was determined by the template number n that had the maximum correlation with the oxyhemoglobin signal map at time t:

$$(6)\ D_t = \underset{n}{\operatorname{argmax}}(r_{n,t})$$

2.2.6 Evaluation of binary decoding performance

After calculating the full vector of decoding outputs (D_m), we evaluated decoding performance by comparing D_m to the ground truth of the study design. ROC analysis was conducted in order to quantify the sensitivity and specificity of decoding across the full range of possible thresholds [94,95]. At each threshold value, D_m was compared to the true stimulus state for each trial to

classify every trial as a true or false positive or negative. For each threshold, the sensitivity, specificity and Youden *J* statistic [96] were then calculated as follows:

$$(7)\ sensitivity = \frac{true\ positives}{true\ positives + false\ negatives}$$

$$(8)\ specificity = \frac{true\ negatives}{true\ negatives + false\ positives}$$

$$(9)\ J = sensitivity + specificity - 1$$

Sweeping the correlation threshold across the full range of possible values (from -1 through 0 to +1 in increments of 0.01) allowed us to plot ROC curves for each possible pairwise classification, i.e., one ROC curve for distinguishing left- versus right- sided checkerboard stimulus presentations, one for decoding left-sided and rest trials, and one for differentiating right-sided and rest trials [94,95]. Areas under the curves were computed as a metric for decoding performance. The maximum value of the Youden *J* statistic was used to determine the optimal threshold value balancing sensitivity and specificity of decoding.

Reproducibility of binary decoding was assessed across 18 task runs in one highly sampled participant by changing which task run was used for template construction, pooling all the remaining runs as test data, and then repeating the decoding and ROC procedure for every possible template.

As an additional measure of reproducibility, ROC analysis was also repeated using only a single task run as test data (rather than pooling test data across runs), but for every possible pairing of

template and test run. This approach was used to evaluate reproducibility both across sessions in one highly sampled participant as well as across multiple participants.

2.2.7 Evaluation of 18-way and 36-way decoding performance

For the 18-way and 36-way decoding experiments, decoding error E_t was calculated at every time point t in the test task run as the absolute value of the difference between the decoded stimulus position (D_t) and the actual stimulus position (S_{t+l}), after sliding the entire true stimulus position time course by a lag time l to allow for the delayed hemodynamic response and match the time window of template construction (i.e., l = 8s for Task 2 and l = 6s for Task 3):

$$(10)\ E_t = |D_t - S_{t+l}|$$

This error value was then averaged across all time points in the task run to calculate mean decoding error.

In order to assess whether the resulting mean decoding error was significantly different from chance-level performance, permutation tests were conducted. For example, to evaluate 36-way decoding performance across sessions within a participant, decoding and mean error calculation procedures were repeated for all 30 possible pairings of training and test data set across 6 task runs in one participant to generate a distribution of mean decoding error. The template labels in the training data were then randomly shuffled 10 times for each of the 30 decoding attempts, and the same process of template matching and error calculation was repeated each time to generate a null distribution of decoding error values derived using 300 shuffled template sets. This null

distribution was compared to the error distribution for true decoding attempts to evaluate the statistical significance of our decoding performance.

To further study decoding performance across different parts of the visual field, mean decoding error was also calculated for each possible position of the stimulus rather than simply averaging across all positions.

Finally, to evaluate the reproducibility of this detailed decoding both within and across individuals, a total of 3 participants performed the rotating checkerboard task 6 times each over the course of 2 imaging sessions involving 2 separate cap fits per participant. The decoding and error calculation process was repeated for every possible pairing of training and test data sets, and the resulting matrix of error values was used to evaluate reproducibility of decoding within and between sessions and participants. Permutation testing was again conducted to test statistical significance, comparing the observed mean decoding error to the error distribution resulting from decoding with 10 random permutations of the training data set for every decoding attempt.

2.2.8 Data and code sharing

To ensure transparency, facilitate reuse of our data, and encourage comparative analyses by other groups, our data will be made available upon request and can be obtained by contacting the corresponding author. Code for pre-processing and reconstructing HD-DOT data is available through Github (https://github.com/WUSTL-ORL/NeuroDOT_Beta), and additional code specific to decoding experiments can also be obtained by contacting the corresponding author.

2.3 Results

2.3.1 Feasibility of decoding visual stimulus position from HD-DOT data

We first assessed the feasibility of decoding HD-DOT data using a simple stimulus position classification problem. Participants were imaged with HD-DOT while they performed multiple runs of a block-design visual task, in which a flickering checkerboard wedge was presented to either the left or the right visual hemi-field interspersed with rest periods of no checkerboard stimulus [52]. We anticipated that these three stimulus conditions (left-sided checkerboard, right-sided checkerboard, and no checkerboard) would produce spatially separable HD-DOT activations. A template matching strategy was therefore used to decode checkerboard position. Using one <10-minute task run as training data, we mapped templates of the oxyhemoglobin response to the left- and right- sided checkerboard stimuli. In a second independent task run used as test data, the stimulus condition was decoded as either left-sided, right-sided, or rest at each time point, based on the spatial correlation between the HD-DOT signal map at that time and each of the templates (as explained in section 2.5). Time traces of the Pearson correlation with each template, the decoded stimulus, and the true stimulus condition for a typical decoding attempt in one participant illustrate the general agreement between decoded and actual stimulus (Figure 2.1).

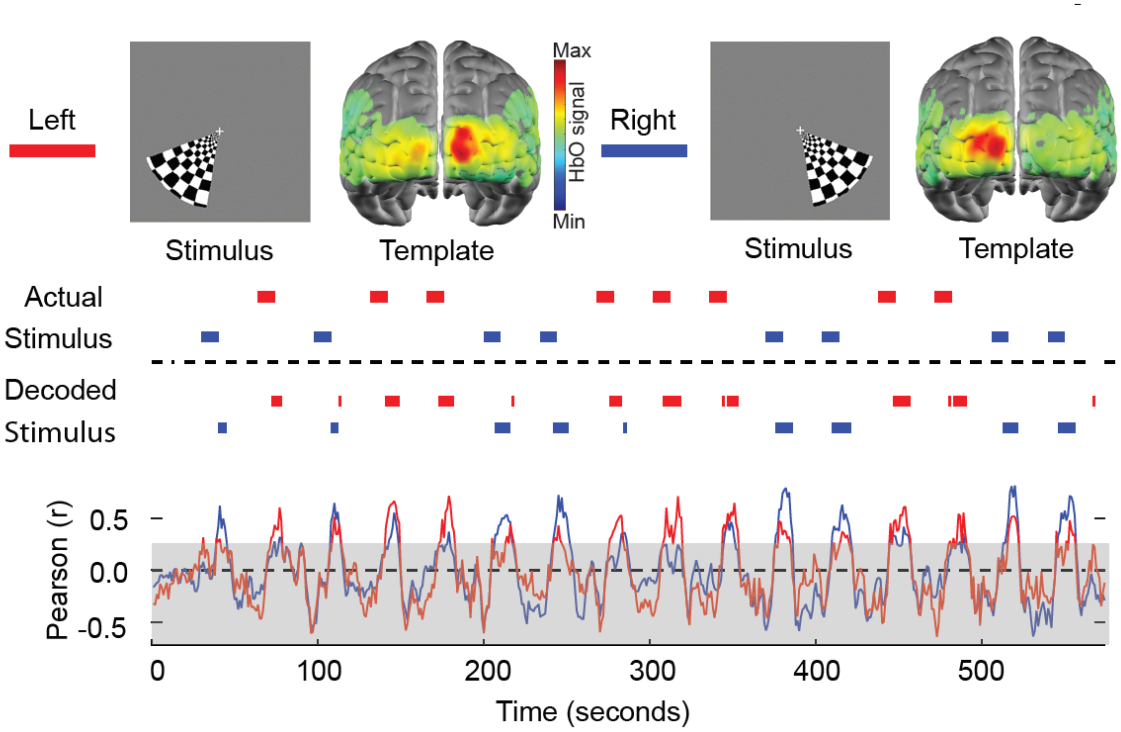

Figure 2.1: Feasibility of visual decoding with HD-DOT data using a template matching strategy Results are illustrated for a typical decoding attempt in one participant. The participant viewed a checkerboard wedge flickering in either the left or the right visual hemifield, interspersed with rest periods, while being imaged by HD-DOT. Training data were block-averaged to define "templates" of expected brain activity for each stimulus condition (top panels). Spatial Pearson correlation coefficients were then computed between each template and the oxyhemoglobin signal map at each time point in the test data (bottom panel). The resulting correlation values were compared between templates and to a designated threshold value (here selected to be 0.25 on an ad hoc basis, but later optimized by ROC analysis in Figure 2.2) to determine the decoded stimulus condition at each time point, and this was compared to the actual stimulus state (middle panels).

2.3.2 Quantitative assessment of binary decoding accuracy

In order to objectively quantify decoding performance, we converted our three-way decoding problem into three binary classifications (left-sided versus no checkerboard, right-sided versus no checkerboard, and left-sided versus right-sided stimulus), and performed ROC analysis for each case [94,95]. Furthermore, to ensure that each test data point was temporally independent, we calculated decoding accuracy on a trial-wise basis rather than across every time point. Each stimulus presentation and inter-stimulus rest period in the test data was treated as an independent event or "trial". To maintain a large sample of data points for a robust analysis amid this trial-wise evaluation of sparse block-design task data, test data were pooled across multiple task runs and imaging sessions where available. For example, one highly sampled participant performed the same checkerboard-viewing task 18 times over the course of 10 imaging sessions, so data from a single task run was used to construct templates, and the other 17 runs were pooled as test data. A plot of correlations for all of this participant's trial responses to the left and right templates reveals a clustering based on stimulus condition (Figure 2.2A).

The thresholds separating different decoding outputs were swept across the full range of possible values, and true and false positive rates of decoding were calculated in each case to construct three ROC curves (Figure 2.2B). The large areas under all these curves (AUCs > 0.98) reflect the high sensitivity and specificity of decoding. The maximum value of the Youden *J* statistic [96] for each ROC curve was used to guide the selection of a data-driven optimal threshold value, and sensitivity, specificity, and overall accuracy were calculated at this threshold (Table 2.2). We

also evaluated the feasibility and performance of decoding the deoxyhemoglobin signal in place of the oxyhemoglobin signal (Table 2.3).

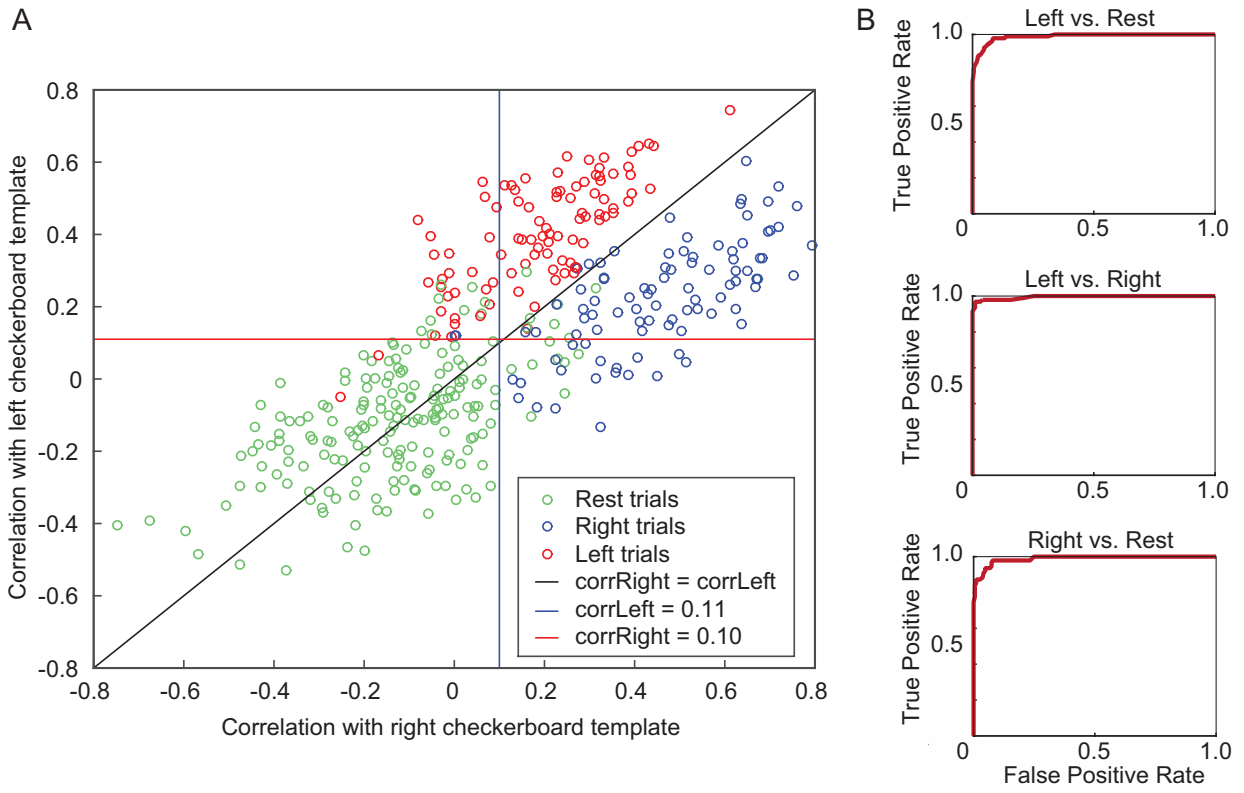

Figure 2.2: Sensitivity and specificity of binary retinotopic decoding with HD-DOT data
(A) Taking data from 18 task runs in one participant, a single task run was used for construction of templates, and independent test trials were pooled across the remaining 17 runs. Plotting all trials by their correlations with each of the two templates reveals a clustering by trial type. Clusters can be separated by thresholds, which can be swept across the full range of possible values and optimized in a receiver operating characteristic (ROC) analysis. (B) ROC curves for the 3 possible binary classifications.

Table 2.2: Measures of decoding accuracy in a highly sampled participant from the ROC analyses in Figure 2.2B

ROC	AUC	Sensitivity*	Specificity*	Accuracy*
Left vs. Rest	0.99	0.98	0.91	0.93
Left vs. Right	0.99	0.96	0.99	0.97
Right vs. Rest	0.99	0.98	0.91	0.93

*Sensitivity, specificity, accuracy calculated at optimized threshold based on Youden *J* statistic [96]

Table 2.3: Measures of decoding accuracy using the deoxyhemoglobin signal in a highly sampled participant

ROC	AUC	Sensitivity*	Specificity*	Accuracy*
Left vs. Rest	0.99	0.99	0.95	0.96
Left vs. Right	0.99	0.95	0.96	0.95
Right vs. Rest	0.99	0.96	0.95	0.95

*Sensitivity, specificity, accuracy calculated at optimized threshold based on Youden *J* statistic [96]

2.3.3 Reproducibility of decoding across imaging sessions

One approach that fMRI studies have taken to assess reproducibility is to collect large quantities of data on individual participants across multiple task runs and imaging sessions and then compare the similarity of results between all possible pairings of the data sets [89]. Adapting this approach to our HD-DOT decoding study, we used data from the participant who performed the same checkerboard-viewing task 18 times over the course of 10 different imaging sessions, spanning more than 1 year. We repeated the ROC analysis described in section 3.2 17 more times, changing which of the 18 task runs was used for template construction each time and again pooling all the remaining data as test data, to produce 18 ROC curves for each of the 3 binary classification problems (Figure 2.3A). The consistently high AUC values (mean AUCs = 0.98 for left vs. rest, 0.99 for left vs. right, 0.99 for right vs. rest) reflect the high sensitivity and specificity of decoding. As an alternate approach to evaluating reproducibility within this data, we also repeated our decoding and ROC analysis using single task runs for template construction again but then using only single task runs as test data (rather than pooling test data across runs), for every possible pair of template and test run (Figure 2.3B). While the number of test trials is lower here than in Figure 2.3A by design, we still find consistently high AUC values across each matrix of >300 possible pairings of template and test data (mean AUCs = 0.98 for left vs. rest, 0.99 for left vs. right, 0.99 for right vs. rest).

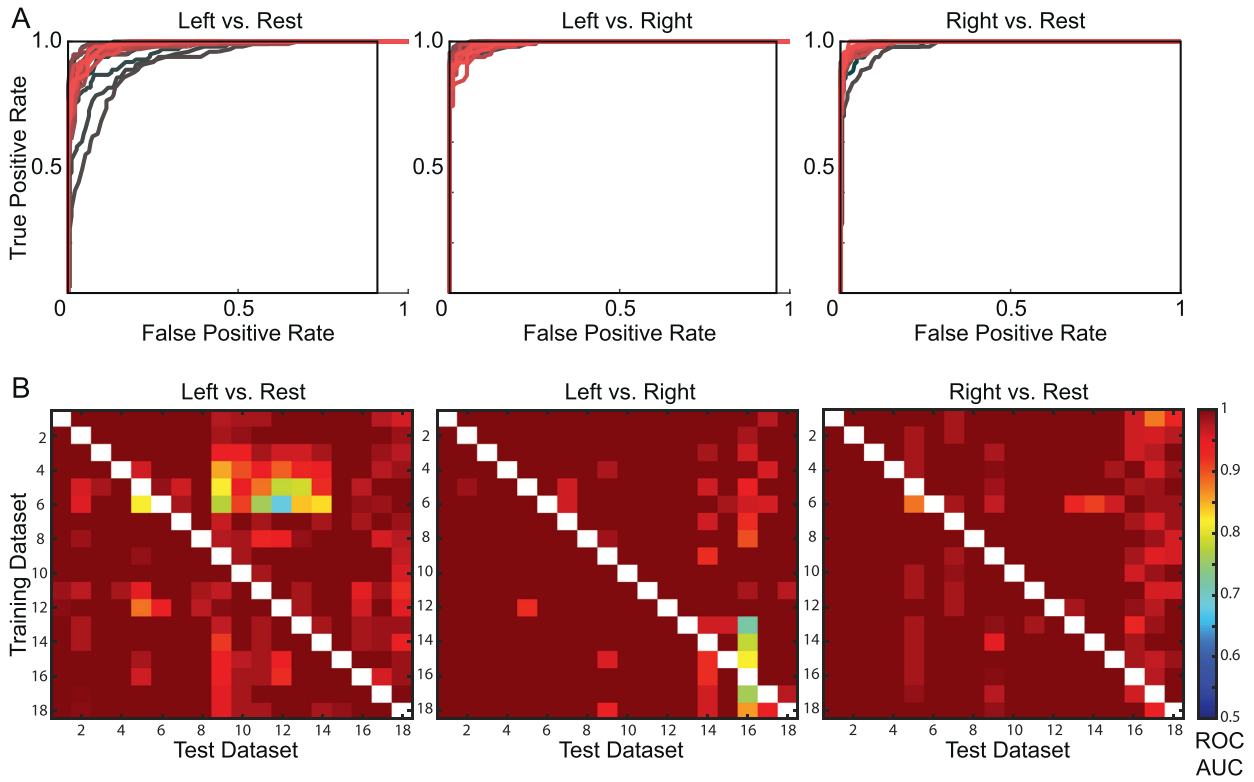

Figure 2.3: Reproducibility of binary retinotopic decoding with HD-DOT across sessions in a highly sampled participant

(A) The data from the participant who performed the checkerboard viewing task 18 times were used to evaluate the reproducibility of HD-DOT decoding across imaging sessions. The single task run used as training data for template construction was changed 17 times, and each time all the remaining runs were pooled as test data to conduct an ROC analysis, producing 18 ROC curves for each of the three binary classification problems. Different shades between red and black were used to plot different ROC curves to make the individual curves more discernible. (B) ROC analysis was also conducted to evaluate decoding with every possible pairing of a single training task run and a single test task run (i.e., without pooling test data across multiple runs). Areas under the curve (AUCs) for all these ROC analyses are plotted in matrices for each of the three possible pairwise classifications, illustrating the reproducibility of accurate binary retinotopic decoding with HD-DOT.

2.3.4 Reproducibility of binary decoding across participants and inter-subject decoding

We also sought to investigate how replicable our decoding was across multiple different participants, and hence took HD-DOT data from four additional healthy adults performing the same checkerboard-viewing task twice each. We conducted ROC analysis of decoding performance within all five participants as described in section 3.2, here using one task run in one participant as training data to decode the other task run in that participant (Figure 2.4A).

In addition, we conducted a reproducibility assay similar to that described in section 3.3 except across participants, evaluating the feasibility of inter-individual decoding. For this experiment we used one task run from one participant as training data to decode a second task run from each of the participants, constructing ROC curves to assess performance for each decoding attempt. This analysis was iterated to evaluate decoding for every possible combination of training and test participant (Figure 2.4B). Mean AUC values were 0.72 for left vs. rest, 0.82 for left vs. right, and 0.84 for right vs. rest, with a majority of the values ranging from 0.75 to 1.0 both along the matrices' diagonals and across off-diagonal elements. These results indicate that both within-subject decoding and between-subject decoding were effective across multiple participants. A one-sample Wilcoxon signed rank test was used to confirm that inter-participant decoding performance was significantly better than chance ($p = 0.011$ for left vs. rest; $p = 2.9 \times 10^{-4}$ for left vs. right; $p = 2.5 \times 10^{-4}$ for right vs. rest). Lower AUC values were observed for some pairings of

training and test data sets with participants 4 and 5. Raw data quality may be one contributor to this variance (Supplemental Figure S2.2).

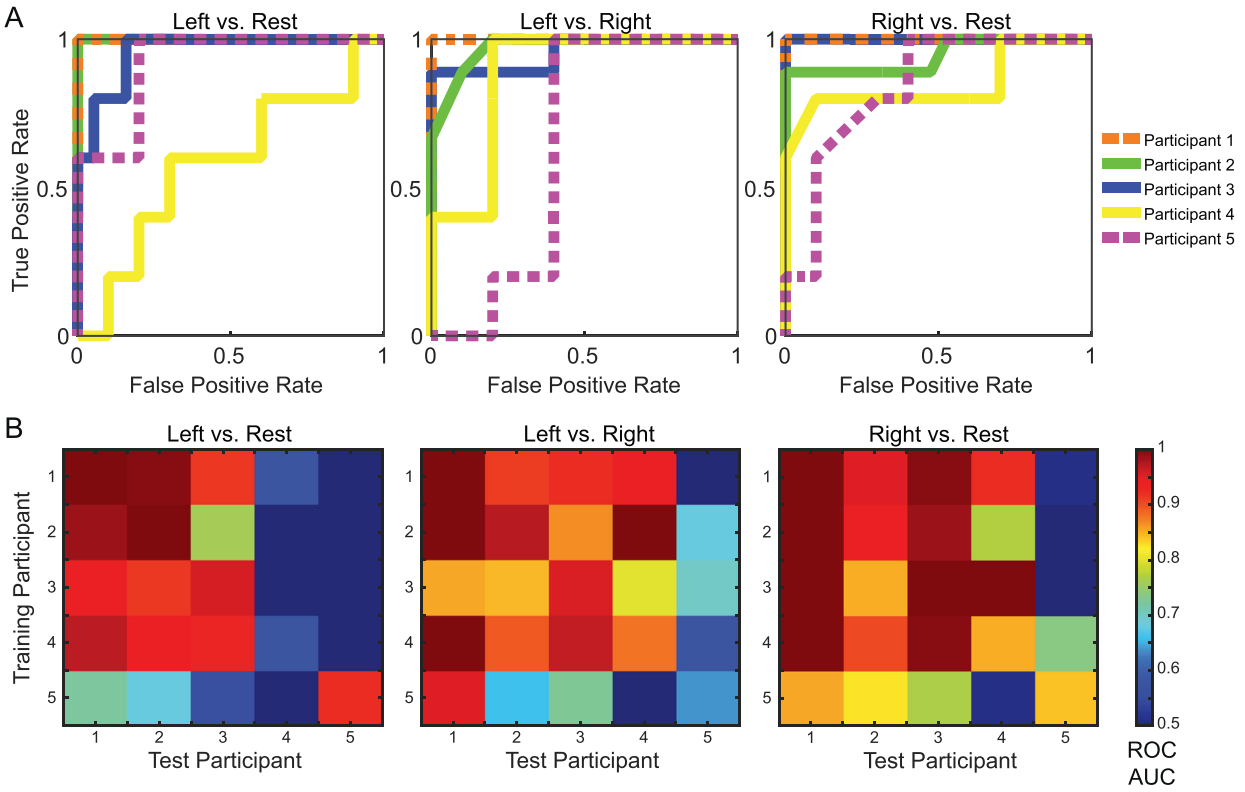

Figure 2.4: Reproducibility of binary HD-DOT decoding across n=5 participants
(A) Data was taken from 4 additional participants who performed the same checkerboard stimulus viewing task twice each. Decoding and ROC analysis were conducted for each participant using one task run for template construction and another task run as test data; resulting ROC curves are shown for all 5 participants. (B) Decoding and ROC analysis were also conducted using templates from one task run in one participant and test data from another run in any participant, across every possible pairing of training and test participants. AUC values along the diagonals of these matrices illustrate reproducibility and variability of within-participant decoding across multiple individuals, while off-diagonal values indicate the feasibility of inter-individual decoding.

2.3.5 Complex, non-binary decoding: 18- and 36- way classification of moving stimulus position

One of the main advantages of HD-DOT over sparser fNIRS imaging systems is the improved spatial resolution and image quality afforded by the high-density arrays of light sources and detectors with their thousands of overlapping measurements [30,53]. In order to harness the spatial resolution of HD-DOT and assess the feasibility of more elaborate decoding than the binary classification established so far, we imaged participants while they observed several patterns of moving checkerboard stimuli. As there were 18 possible sizes for the expanding/contracting ring stimuli and 36 possible positions of the rotating checkerboard wedges, we constructed sets of 18 and 36 templates, respectively, for these two task paradigms by block-averaging single <7-minute task runs. We then used a template matching strategy to decode stimulus position at each time point in independent runs of the same task without any block-averaging of the test data. Graphs of Pearson correlation values for all the templates and plots of the actual and decoded stimulus positions are shown for a typical contracting ring task run (Figure 2.5A) and a typical rotating wedge task run (Figure 2.6A, Supplemental Figure S2.3).

The similarity between the actual and decoded stimulus traces indicates that decoding each time point was feasible even in these more complex 18-way and 36-way classification problems. After correcting for the hemodynamic time delay, visible in the 6-8 second lag between the actual and decoded stimulus traces (Figures 2.5A and 2.6A), the discrepancy between the actual and

decoded stimulus positions was calculated for each time point and these values were averaged across the run as a quantitative measure of accuracy. For example, across all 30 of the 36-way rotating stimulus decoding attempts in one high-performing participant, the mean error (± standard deviation) in decoding stimulus position was 18.0±17.4° (compared to the mean error for decoding at chance which would have been 90°). To assess whether decoding performance in this participant was significantly better than chance, we conducted a permutation test across all the data collected in this participant as described in Section 2.7. The separation between the true decoding error distribution and the null distribution illustrates the statistical significance of the results (Figure 2.5B, $p<0.0083$, and Figure 2.6B, $p<0.0033$). Interestingly, breaking down the mean absolute error calculation for each stimulus position revealed no significant variation in decoding error with stimulus eccentricity (Figure 2.5C), but greater error when the checkerboard wedge rotated through the upper visual hemifield than the lower half of space (Figure 2.6C).

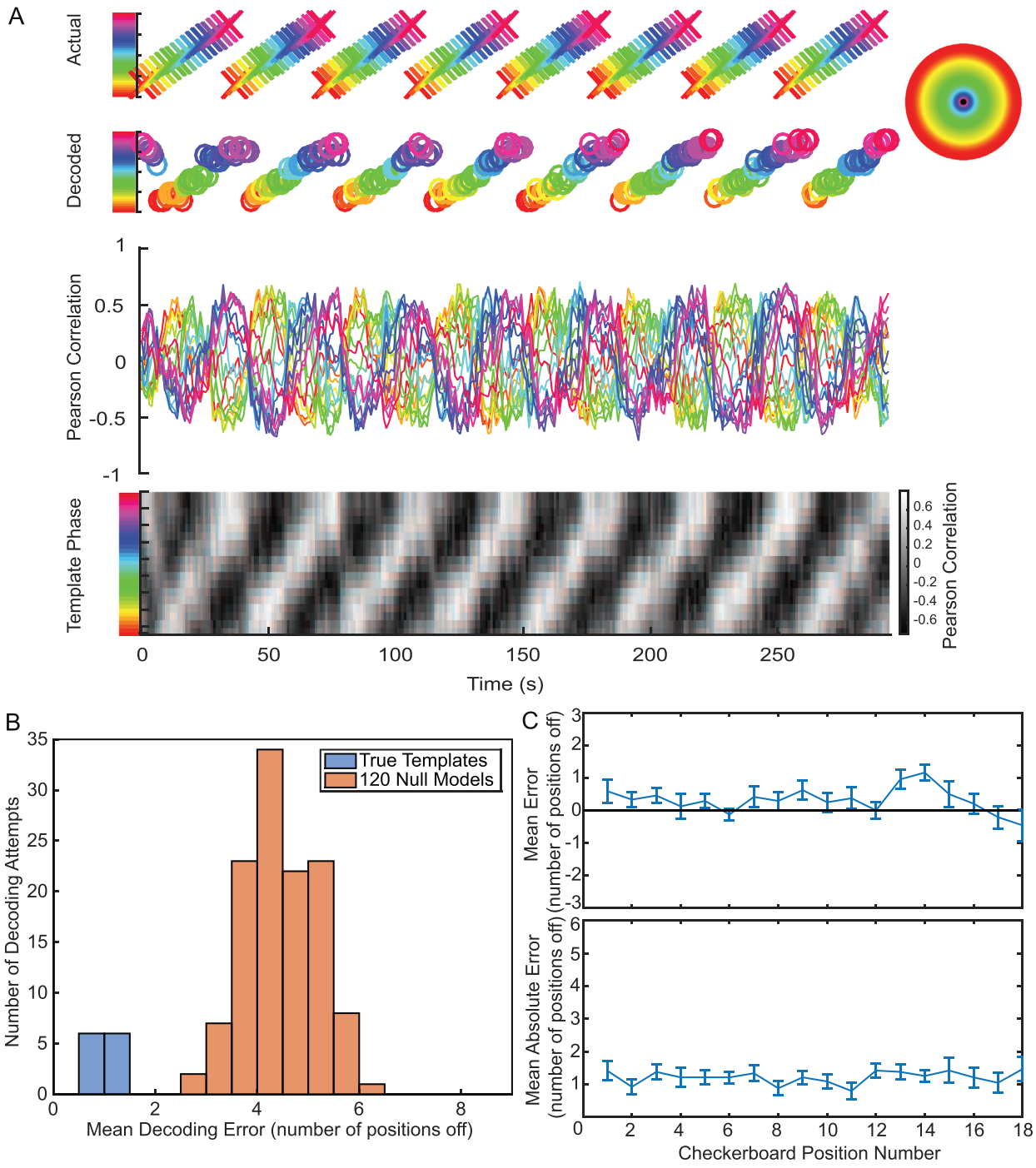

Figure 2.5: 18-way classification of visual stimuli with varying eccentricity within a single participant

A participant was imaged using HD-DOT while watching a flickering checkerboard ring over 8 cycles of either periodic expansion or contraction through 18 concentric positions on a screen. A template matching strategy was again used to decode stimulus location at each time point in a test dataset, here using 18 template maps – 1 for each stimulus phase. (A) Each stimulus phase is assigned a color (as per the color wheel) and a position along the vertical axis in the plots of actual and decoded stimulus positions. Pearson correlation coefficients were calculated between each of the 18 templates and the oxyhemoglobin signal map at every time point in the test data, and are plotted on the two graphs at the bottom of this panel. The decoded stimulus corresponds to the template with the maximum correlation at each time point. (B) A permutation test was performed to evaluate the significance of this decoding using every one of the 12 possible pairings of the 4 task runs collected in this participant for training and testing. The mean decoding error obtained using true template sets was compared with a null distribution generated using 10 random permutations of the training data for every true decoding attempt ($p<0.0083$). (C) Mean decoding error across the 3 test runs for a single training run is plotted here as a function of true stimulus position, showing little variation in decoding performance with eccentricity.

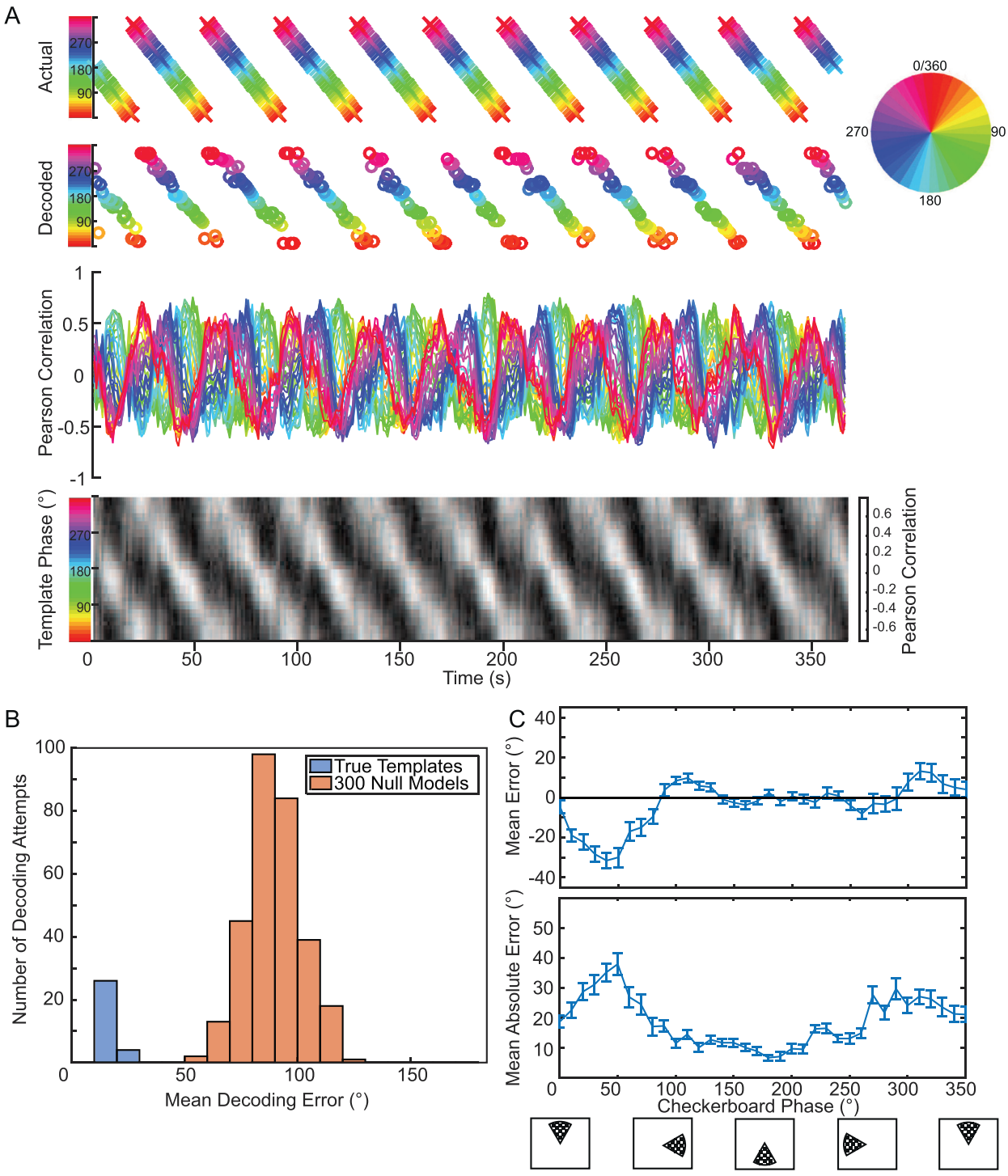

Figure 2.6: 36-way classification of rotating visual stimulus position within a single participant

A participant was imaged using HD-DOT while watching a flickering checkerboard wedge rotating 10 times through 36 positions over 36 seconds per revolution. A template matching strategy was used to decode stimulus location, here generating a set of 36 templates (one for each phase of the stimulus) from one training task run and decoding stimulus position at every time point in an independent test task run. (A) Each stimulus phase is assigned a color (as per the color wheel) and a position along the vertical axis in the plots of actual and decoded stimulus positions. Pearson correlation coefficients were calculated between each of the 36 templates and the oxyhemoglobin signal map at every time point in the test data, and are plotted on the two graphs at the bottom of this panel. The decoded stimulus corresponds to the template with the maximum correlation at each time point. (B) A permutation test was used to evaluate the statistical significance of the decoding performance. Mean decoding error was calculated for every one of the 30 possible pairings of training and test data set across the 6 task runs performed by this participant. The resulting error distribution was compared with a null distribution generated using 10 random permutations of the training data for each true template set ($p<0.0033$). (C) Mean decoding error across the 5 test runs for a single training run is plotted as a function of true stimulus position, showing better decoding performance as the wedge rotates through the lower half of visual space.

2.3.6. Reproducibility of non-binary HD-DOT decoding across participants

For the decoding run depicted in Figure 2.6A, both the training and test data were collected from the same participant and during the same imaging session. To further investigate the robustness of this more elaborate decoding, we imaged three different participants as they each viewed the rotating checkerboard stimulus sequence six times spread over the course of two imaging sessions involving two separate cap fits. We then repeated our decoding and error calculation analysis, training with one task run and testing with another independent run, for every possible pairing of training and test data set across all the participants and imaging sessions (Figure 2.7A-B). Although performance varied between participants, we observed reproducible decoding across task runs and imaging sessions, with a mean within-participant decoding error (± standard deviation) of 25.8±24.7° relative to chance performance at 90° ($p<0.0011$). We also observed effective inter-participant decoding (i.e., when training and test data were from different individuals), with performance varying between sessions and participants, but mean inter-participant decoding error passing permutation testing for statistical significance ($p<0.0005$).

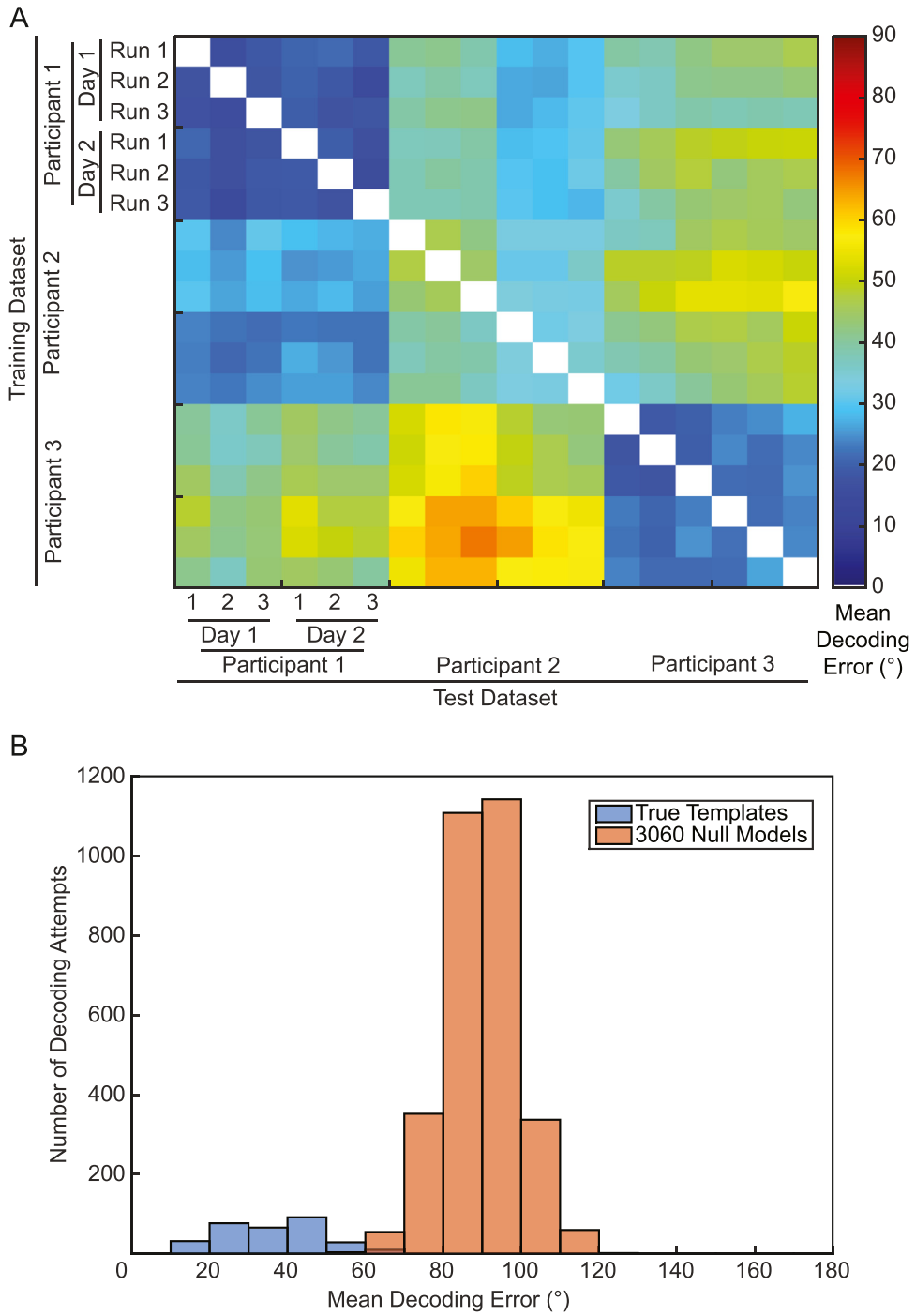

Figure 2.7: Reproducibility of complex, non-binary visual decoding with HD-DOT across n=3 participants

Three participants viewed the same rotating checkerboard stimulus paradigm six times across two imaging sessions each. Decoding performance was then evaluated using every possible pairing of template and test task run across participants. (A) The mean absolute error for each decoding attempt illustrates both the reproducibility and variability of within-subject and inter-individual decoding across the imaging sessions and participants. (B) Permutation testing comparing decoding performance across all sessions and participants to a null distribution.

2.4 Discussion

To establish the feasibility of decoding with HD-DOT, we have used a well-validated visual stimulation protocol and have evaluated the accuracy and replicability of decoding stimulus position across multiple imaging sessions and research participants for a range of decoding complexity. Using a straightforward template matching strategy with training and test data from separate 5-10 minute-long task runs, we found that HD-DOT data can be used to accurately decode stimulus position at individual time points without needing to block-average test data. We performed ROC analysis to quantify the accuracy of binary visual decoding, obtaining mean AUC values >0.97 across 10 imaging sessions (including 18 task runs) in one highly sampled subject, and obtaining mean (intra- and inter-participant decoding) AUC values of 0.7-0.85 across a group of 5 participants. More challenging 18-way and 36-way decoding experiments also showed strong decoding performance across multiple imaging sessions and participants. For these latter studies, the phase of a 60° wide checkerboard wedge rotating 10° per second through 360° was decoded with a mean error of 18.0±17.4° in our best participant across multiple task runs and cap fits. This decoding performance varied with data quality but remained significantly above chance based on permutation testing across imaging sessions in two additional participants. Inter-participant decoding, with training and test data taken from different individuals, was feasible for both the binary and the more complex decoding.

2.4.1 Study design in relation to prior decoding research

Some of the most elaborate decoding research so far has used either ECoG or fMRI to sample brain activity and reconstruct things like intelligible speech and detailed visual experiences [39,40,43]. However, ECoG is too invasive and fMRI too cumbersome for translation beyond clinical populations and laboratory studies into widespread use. While fNIRS and electroencephalography (EEG) overcome these challenges as noninvasive and portable imaging methods that have been used in previous decoding research, their low spatial resolution limits the space-bandwidth product available for decoding. On the one hand, decoding could have a meaningful impact even with a low bit rate; for instance, studies in patients with locked-in syndrome have shown the potential of fNIRS to enable patients with few other means of communication to respond yes or no to questions [46]. However, by increasing the optical channel count and density relative to traditional fNIRS, HD-DOT combines logistical advantages of optical neuroimaging with image quality that is closer to that of fMRI, providing motivation to study the feasibility and performance of more detailed decoding with HD-DOT.

Retinotopic decoding was chosen as the focus of the current study for several reasons. Firstly, using an externally controlled stimulus, rather than studying internal thought or other higher-order cognitive functions, provided definite knowledge of ground truth and a means to manipulate it. This approach enabled a fairly simple experimental design to objectively quantify decoding accuracy. In addition, it had already been shown that the flickering checkerboard stimuli used here evoke repeatable and distinct brain activity patterns at different positions in the visual field that can be mapped by HD-DOT [52,60,69], making decoding the positions of these

stimuli a reasonable goal. Finally, using the visual system as a model for initial proof of principle follows the wisdom of neuroscience literature in general, exemplified by Hubel and Wiesel's seminal plasticity studies [13–15,97], and follows the arc of the successful fMRI decoding literature in particular [2,77,78,98].

2.4.2. Accuracy of binary visual decoding with HD-DOT

The feasibility of decoding with HD-DOT was evident in how closely our decoding results mirrored actual stimulus conditions. We employed ROC analysis to further evaluate decoding accuracy, objectively quantifying true and false positive rates across a range of decoding threshold values [94,95]. The consistently large area under the ROC curves highlights the high sensitivity, specificity, and accuracy of the single-trial level decoding presented (Figures 2.2B and 2.3, Table 2.2). Decoding the deoxyhemoglobin signal in place of oxyhemoglobin also yielded similarly high performance (Table 2.3). While head-to-head comparison to previous fNIRS studies is difficult due to differences in study design, a study of binary audiovisual decoding in infants attained trial-wise decoding accuracy in the 55-70% range [48]. While our focus on simple stimuli and adult participants may have partly facilitated our higher decoding performance, another likely contributor to our gain in accuracy is the high channel count of HD-DOT (supporting ten-fold to a hundred-fold more measurements than most fNIRS arrays), which improves the space-bandwidth product of our data, increasing the amount of information that can be leveraged for decoding. These various contributing factors could be separated in future research.

2.4.3 Reproducibility of binary visual decoding with HD-DOT across sessions and participants

Both the research and clinical applications of decoding hinge on its reliability, so we sought to assess the replicability of our results across multiple data sets. Recent fMRI studies have highlighted the advantages of conducting research with highly sampled individuals [89,99], so we chose to first investigate reproducibility in one participant across 18 task runs conducted over 10 different cap fits. Across combinations of training and test data, we obtained mean ROC AUC values >0.97, demonstrating the reproducibility of accurate binary visual decoding that can be achieved within a single participant, even when data is collected across multiple imaging sessions and cap fits over an extended period of time. This consistency of decoding across sessions illustrates the reproducibility of HD-DOT signals as well as the reliability of our cap fit procedure, which leverages anatomical landmarks, the structural integrity and flexibility of the imaging cap, and the use of feedback from real-time data quality visualizations [53]. In addition, this observed reproducibility is encouraging for the continued pursuit of research and applications wherein a patient would rely on HD-DOT as a means of communication; it is technically feasible, for instance, that a HD-DOT decoder could be trained over one or more days, and then be used for decoding for months to years afterwards. This potential use case also further supports the study of highly sampled individuals.

Interestingly, there were some cases in which a run served as a consistently good test data set but not as reliable a training data set (e.g. run 6 for left vs. rest classification), and vice versa (e.g.

run 16 was a consistent training data set but not always a good test data set for left vs. right decoding). These occasional asymmetries likely reflect differences in processing for training and test data. Templates were constructed from training data by block-averaging across a 5 second time window and across all presentations of a stimulus. Meanwhile, for test data, a single time point was used to assess decoding for each trial. Furthermore, inconsistencies in decoding performance for some specific pairings of data sets and not others may be a result of minor variations in the positioning of the cap between sessions. Future studies could investigate these potential sources of variance and attempt to further improve the consistency of decoding.

Data was taken from four additional participants and subjected to a similar analysis of decoding accuracy and reproducibility across individuals (Figure 2.4). Here, the within-participant ROC curves (Figure 2.4A) and the AUC values along the diagonal of the reproducibility matrices (Figure 2.4B) reflect high within-subject reliability in most of the participants, while off-diagonal matrix values illustrate that with HD-DOT data it is feasible to train a decoder on one individual and successfully infer the visual stimulus seen by another (Figure 2.4B). The inter-individual variability in decoding outcomes, with some participants yielding higher decoding accuracy than others, is likely partly related to variation in data quality between individuals associated with a combination of factors such as anatomical variability and participant compliance with regard to both maintaining central fixation and minimizing motion. Indeed, as illustrated in Supplemental Figure S2.2, participants with poor decoding performance had poor data quality to begin with, as assessed by signal-to-noise ratio and template laterality. This underscores the importance of monitoring data quality.

Nevertheless, the reproducibility of HD-DOT decoding across multiple sessions and individuals and the feasibility of inter-individual decoding suggest that it should be possible to pool training data across multiple imaging sessions and participants in future studies. While in this study we were able to train our decoder sufficiently with data from a single 5-10 minute task run, more complex decoding tasks will likely require significantly more data than could be acquired in one continuous session. For instance, prior visual decoding studies using fMRI data and more complex algorithms to reconstruct novel naturalistic stimuli outside of the training set collected 2-3 hours of training data in each of 3 participants to train their decoders without over-fitting and also block-averaged hours' worth of test data [39,40]. Furthermore, though subject-specific training data may often facilitate the most accurate decoding, the decoding performance we observed across sessions and individuals suggests that a decoder could be pre-trained with group data to enhance efficiency of training and performance of decoding in some cases. While inclusion of low quality training data would likely impair decoding performance, including high quality data from other individuals could even improve decoding; e.g. decoding left vs. right in participant 4 is even more effective using the high quality training data from participants 1 and 2 than the data from participant 4 theirself. The success of a pre-trained decoder could further be enhanced by anatomical and functional co-registration of group-level templates to a subject-specific space and by updating the decoder with additional subject-specific training data. Future studies can systematically assess the quantity of HD-DOT data required for an optimized pre-training approach and its effects on decoding performance.

2.4.4 Performance of more spatially and temporally detailed retinotopic decoding with HD-DOT

Finally, we reasoned that the increased space-bandwidth product afforded by our high-density imaging array may enable us to distinguish a larger number of targets than prior optical decoding studies that have mostly performed binary or 4-way decoding [47–51]. As a result, we increased the complexity of our decoding and performed the 18-way and 36-way classification experiments with the moving checkerboard ring and wedge stimuli. We found that we were able to localize a 60° wide checkerboard wedge rotating 10° at a time through 360° with a mean within-participant error of 18.0±17.4° across all sessions in our highest performing participant (Figure 2.6B) and 25.8±24.7° across all participants. While this decoding performance is significantly better than chance (90°, $p<0.0033$), the error indicates that the decoder cannot independently resolve all of the 36 positions separated by 10°. A better estimate of the number of independent radial positions that could be decoded is the full cycle (360°) divided by the mean error, or approximately 20 positions in our highest performing participant and fewer for other participants (~14 positions on average).

We observed several intriguing patterns in our evaluation of 36-way decoding across multiple participants (Figure 2.7). Decoding performance within a participant appears to be consistent across runs and sessions, and is best in participant 1 and worst in participant 2. We interpret this as reflecting that the repeatability of precise retinotopic activations is greatest in participant 1 and lowest in participant 2, perhaps as a result of differences in participant compliance with central

fixation and differences in data quality. Another possible contributing factor is anatomical variability between participants in the precise folding of visual cortex. For example, in some participants, responses to stimuli in all four quadrants of the visual field can be mapped by HD-DOT, but in others only responses to the lower visual field hemi-field can be captured with high signal-to-noise [69]. The fidelity of activations appears to be even more critical for test data than training data based on the asymmetries observed in the matrix. For instance, testing with participant 1 after training with participant 2 yields better results than both testing with participant 2 after training with participant 1 as well as using both training and test data from participant 2. These observations are consistent with the fact that template maps are created by averaging training data across blocks, which boosts signal-to-noise ratio, while every time point is evaluated separately in the test data, which makes consistency particularly important. However, additional factors evidently influence inter-participant decoding; for instance, decoding of test data from participant 1 is better using templates from participant 2 (who had the worst within-participant decoding performance) than using templates from participant 3. It is likely that inter-individual differences in anatomy and cap positioning contribute to such trends. For example, participants 1 and 2 may have had more similar head shapes, occipital cortex anatomy, and cap fits, such that retinotopic activations were more similar between them than to those seen in participant 3, potentially explaining why inter-participant decoding was more successful with participants 1 and 2.

In addition, the visual stimulation paradigms used allowed us to study how decoding performance varied across the visual field. It emerged that there was a discrepancy between

decoding accuracy in the lower and upper hemifields with generally lower decoding error for stimuli in the lower hemifield (Figure 2.6C). This result is consistent with prior reports of variations in HD-DOT signal-to-noise ratio across different parts of visual cortex [69]. These observations likely stem from the anatomy of visual areas V1, V2, and V3, organized retinotopically around the calcarine fissure with the upper hemifield represented deeper below the cranial surface and hence less accessible to photonic measurements [100,101]. Structural and functional neuroanatomical variability between individuals likely explains why this effect of stimulus position on decoding error was more pronounced in some individuals than others [90,102–106].

Nonetheless, the feasibility of this detailed decoding even with training and test data from different cap fits and individuals reflects the potential of HD-DOT for supporting applications of neural decoding. Furthermore, the success of this spatially detailed decoding encourages exploring other forms of more elaborate decoding, such as deciphering naturalistic stimulus information. In fact, some of the error recorded in the current decoding experiments may have stemmed from lapses in compliance with central fixation (as it is challenging to keep gaze fixated on a central crosshair while bright, distracting patterns flicker in the periphery), so a more naturalistic visual stimulus may even improve decoding performance. Decoding of more complex stimuli could also leverage signals from additional areas beyond early visual cortex to potentially distinguish a larger number of targets.

Aside from the increased spatial detail of decoding revealed by the 18- and 36-way classification experiments, the success of decoding stimulus location at individual time points without having

to average across multiple trials is also a step towards developing the real-time decoding that would enable efficient brain-computer interfaces for clinical use. The current study aimed to evaluate the feasibility of decoding while using established methods for processing and reconstructing optical data, including steps that use the full run of data such as zero-phase high-pass filtering, superficial signal regression, and mean-subtraction. Future studies could reassess the performance of single-trial and single-time-point decoding using a modified processing pipeline that changes these steps to respect causal relations, only using prior data for any given time point. Such analysis would provide an assessment of pseudo-real-time decoding. Based on the accuracy and timing of this proposed pseudo-real-time decoding, subsequent studies could decode HD-DOT data in real time while participants are being imaged and deliver feedback to investigate closed-loop HD-DOT-based brain-computer interfaces.

2.4.5 Limitations, potential solutions, and future directions

It is important to note that our study intentionally enriched for participants and data sets with high raw data quality, as evaluated by light levels, signal-to-noise, pulse signal, and head motion. Our rationale for this decision was multifactorial. Firstly, as we were evaluating a new method of analyzing HD-DOT data, it was important to ensure that the quality of the data itself was not a confounding factor undermining the study. Secondly, the myriad possible sources of noise (instrumentation, cap fit, head motion, participant attention, confounding physiology, etc.) render a reasonable analysis of these factors and their effects on decoding beyond the scope of the current study. Furthermore, data quality is a moving target, with newer systems and strategies

mitigating the effects of head motion and improving signal-to-noise [107–109]. Data quality is also expected to vary dramatically as decoding is applied to different tasks, domains, and populations. Therefore, additional research will be required to evaluate generalizability across a broader population and range of data quality. Future studies with larger numbers of participants would also be better powered to more systematically evaluate the relationship between different components or measures of data quality and decoding performance. Overall, we anticipate average decoding performance might be poorer among a general population. However, on the upside, most of the potential sources of noise discussed are addressable. For instance, anatomical variability could be at least partly addressed through subject-specific head modeling [60]. Participant compliance could be improved through training and feedback for both maintaining central fixation [110] and minimizing head motion [34,111]. And given the fluid nature of data quality and the progression of the fNIRS research field towards ensuring higher quality data, it may be possible to further improve decoding performance through advances in hardware such as wireless systems with high signal-to-noise [107,108] and algorithms for rigorous monitoring and optimization of data quality [109]. As a result, the field of optical neuroimaging will likely see improvement in the performance and scope of decoding in future research.

The current study was restricted to visual decoding, but it also provides a framework to begin exploring decoding of other modalities with HD-DOT. For example, other sensory systems and the motor system also organize information systematically in neuroanatomical space [112,113], and these maps could be leveraged to decode motor imagery or auditory signals with HD-DOT to replicate and expand on prior work with fMRI and fNIRS [51,72,73]. In these applications

too, HD-DOT could combine the strengths of fNIRS, such as portability and a quiet scanning environment, with higher spatial resolution and image quality to potentially perform the more detailed decoding that fMRI has supported.

It is left for future studies to determine if the decoding established here using a simple task paradigm and normative adult population can be extended to more challenging tasks and populations, as have been explored with other imaging modalities. In particular, decoding naturalistic stimuli outside the decoder's training set, as previous fMRI studies have done, will require more sophisticated decoding strategies than the straightforward template matching approach used here. Future HD-DOT studies may hence explore decoding using stimulus feature encoding regression models or convolutional neural networks for classification [38–40]. Decoding more complex tasks will also rely on capturing more subtle signals distributed across broader areas of cortex than the robust signal localized to visual cortex that is evoked by checkerboard stimuli. However, HD-DOT has been shown to be capable of functional brain mapping with high signal-to-noise ratio during tasks such as verb generation, covert reading, and viewing multimodal naturalistic stimuli [53,64], which is encouraging for decoding covert and complex signals pertaining to language and other domains beyond vision. Decoding in other populations, such as infants and patients with neurological disorders who have participated in prior fNIRS decoding studies [46,48], will present additional data acquisition challenges such as increased levels of motion and the need for imaging at the bedside. However, previous HD-DOT brain mapping studies have already established the feasibility of imaging at the bedside in neonates and stroke patients [62,114,115], while newer lightweight [116] or fiber-less [107]

designs will further increase portability and reduce motion artifacts facilitating decoding studies in such populations.

Finally, potential real-world applications of HD-DOT decoding such as brain-computer interfaces in neurological populations would build on the combination of advances in imaging hardware, real-time data processing methods, and decoding algorithms. Although motor prosthetic control and augmented communication commonly garner more attention, one of the earliest demonstrations of a brain-computer interface was based on decoding EEG responses to checkerboard visual stimuli [117]. In that study, participants guided a cursor through a maze by shifting their center of fixation relative to a checkerboard stimulus, which produced visual evoked potentials that were decoded and used to update the location of the cursor on the screen. A similar paradigm could be used to apply our retinotopic decoding in a visual HD-DOT-based brain-computer interface in healthy participants for proof of principle. This framework for real-time processing and feedback could then be combined with other parallel progress in HD-DOT decoding, such as motor and semantic decoding, and with hardware advances, such as increasingly wearable HD-DOT systems, to develop HD-DOT-driven prosthetics and communication for patients with motor disabilities.

There are indeed many steps between our study of explicit visual decoding and long-term goals such as brain-computer interfaces for neurological patients. However, the current validation of the feasibility, accuracy, and reproducibility of detailed, single-trial visual decoding with HD-DOT provides a solid foundation to build upon in future research.

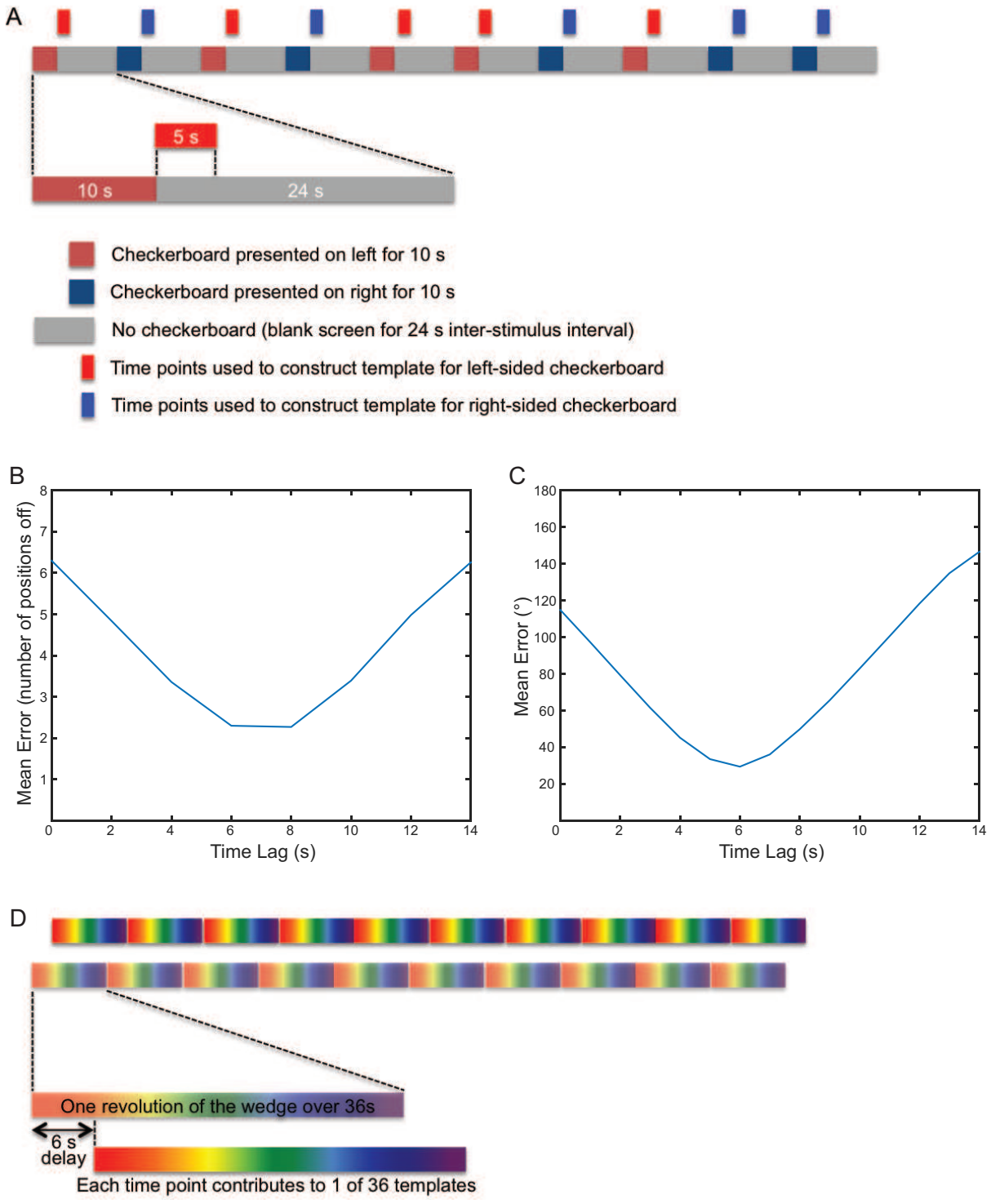

Figure S2.1: Stimulus design and template timing

(A) Left-versus-right checkerboard wedge viewing task paradigm: in this block design task (used in Figures 2.1-2.4), 10 second checkerboard presentation windows were interleaved with set inter-stimulus intervals (24s long in this schematic). The checkerboard stimulus could be presented to either the left or the right visual hemifield in a pseudo-random order. Responses to each stimulus were block-averaged to construct two "templates" of the typical hemodynamic responses to the two stimulus positions. The template maps were sampled across a 5 second window beginning 10 seconds after stimulus onset, in accordance with the peak of the hemodynamic response to these stimuli (Zeff et al. 2007; Gregg et al. 2010). (B) In order to determine the optimal time delay for template construction for the 18-way decoding experiment, mean decoding error was evaluated for a range of time delays using a single pairing of training and test data sets collected in participant 1. Error was minimized at a time delay of 8s. (C) An analogous error minimization analysis for the 36-way decoding experiment yielded an optimal time delay of 6s. (D) Phase-encoded checkerboard stimulus viewing task paradigms: In the rotating checkerboard paradigm (used for Figures 2.6-2.7) depicted here, participants watched a flickering checkerboard wedge rotate 10 times through 36 positions spanning 360 degrees at 36 seconds per revolution (1 second per position). After allowing the empirically optimized 6 second delay, every time point contributed to one of a set of 36 templates (one for each position of the checkerboard) averaged across the revolutions of the stimulus. In the expanding/contracting checkerboard ring paradigm (used for Figure 2.5), the stimulus instead progressed through 18 concentric positions, spending 2 seconds at each position, and hence a set of 18 templates was constructed sampling 2 successive time points (with the empirically optimized 7-8 second delay) from each cycle for each template

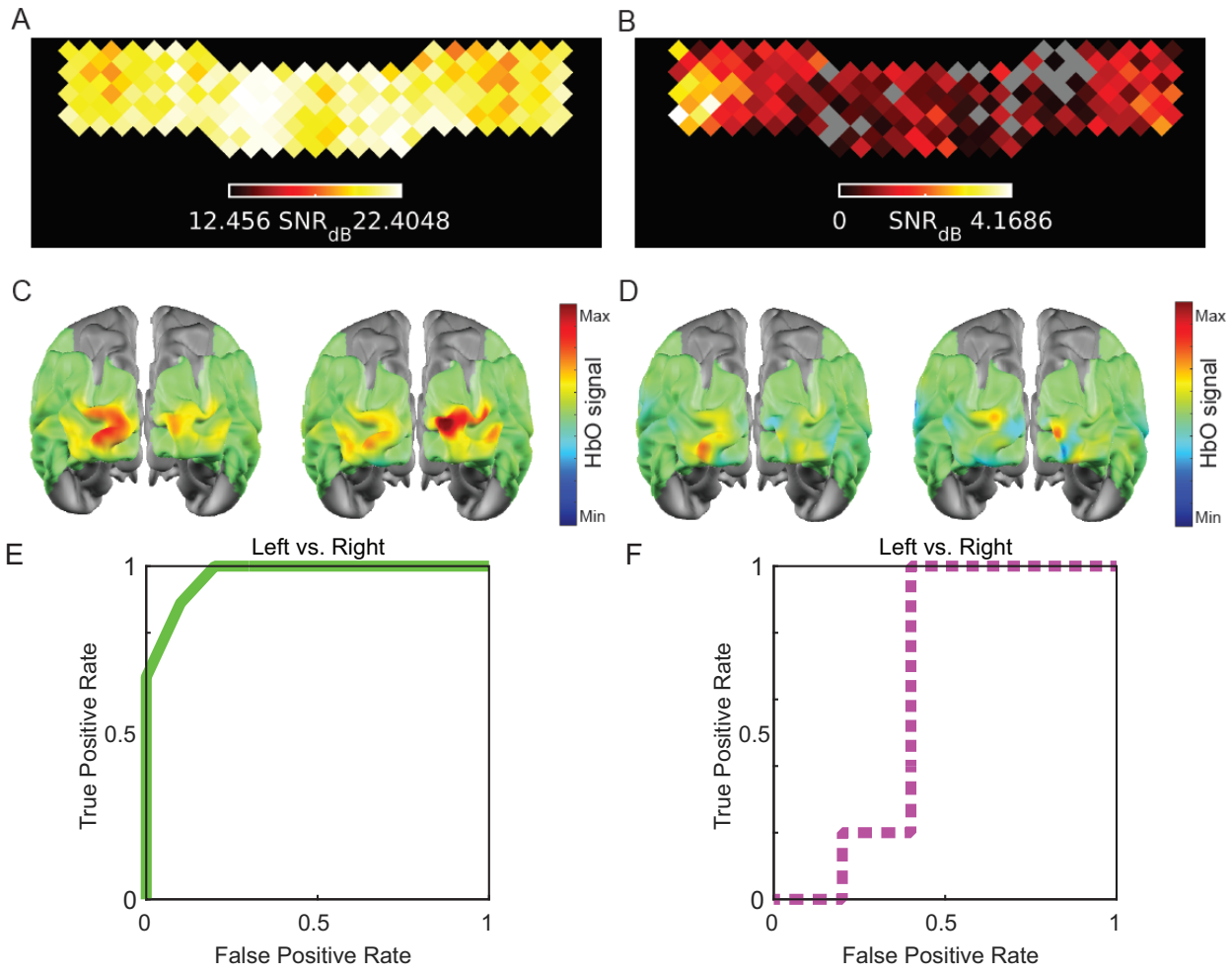

Figure S2.2: Raw data quality is associated with decoding performance
High data quality, including high cardiac-pulse-to-background signal-to-noise (measured as the ratio of signal power in the 0.5-2 Hz range to bandwidth-scaled median power in the flanking frequency bands) across the cap (A), is associated with well-lateralized template maps (C) and high decoding performance as indicated by ROC analysis (E). In contrast, poor data quality including low signal-to-noise across the cap (B) is associated with less robust template maps (D) and lower decoding performance (F).

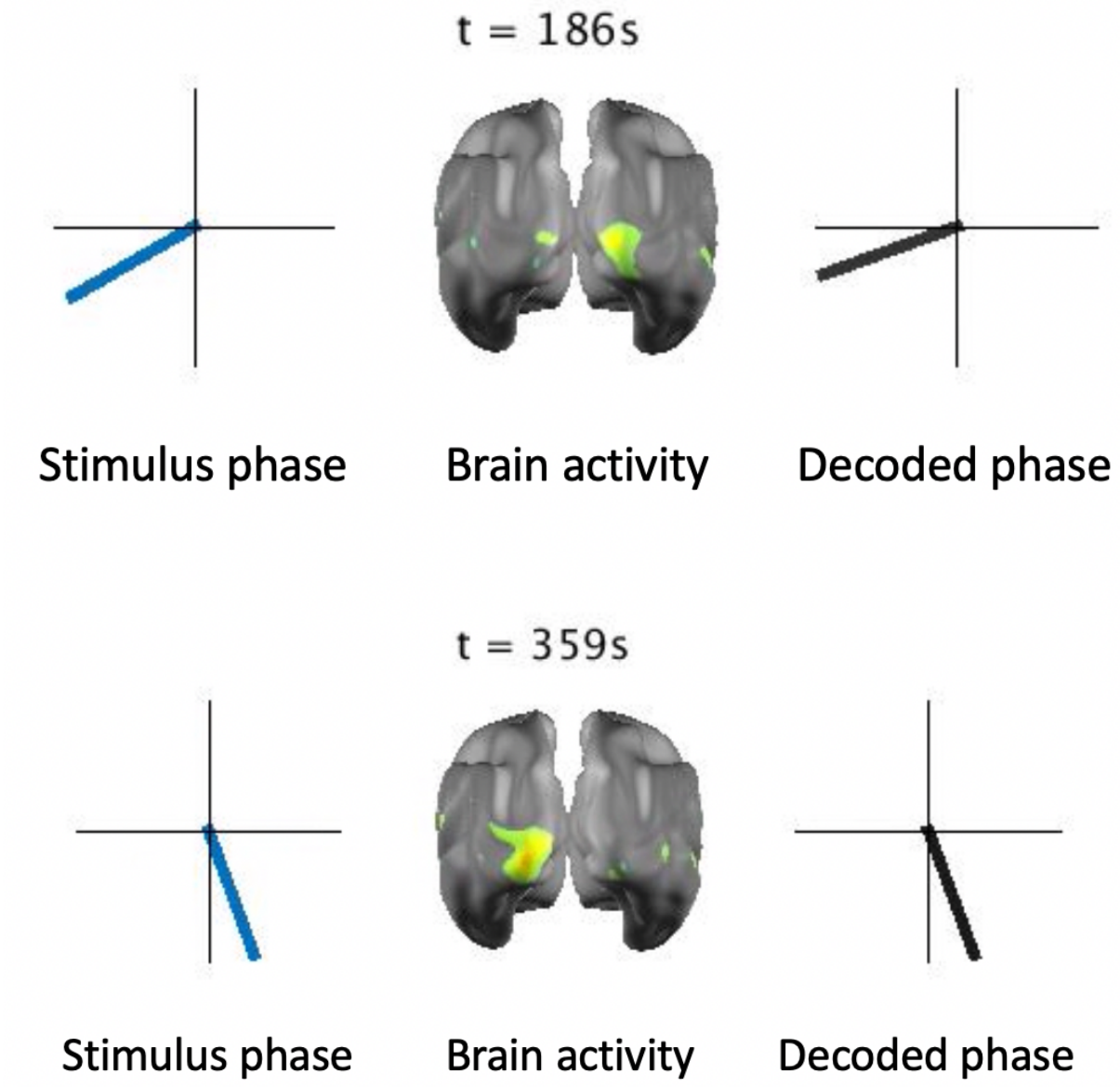

Figure S2.3: 36-way classification of rotating visual stimulus position
Stills from a movie providing an alternative presentation of the data in Figure 2.6A, comparing true stimulus phase (left) to the decoded phase (right) determined from the HD-DOT map of brain activity (center) at various time points in an exemplary decoding run

“Kho na jaaiye, taare zameen par”
(“Lest they get lost, the stars on our Earth”)

- Prasoon Joshi, Shankar-Ehsaan-Loy

Chapter 3: Mapping brain function in awake young children during naturalistic viewing with high-density diffuse optical tomography

3.1 Introduction

The first several years of human life are packed with exciting behavioral changes, but the underlying changes in brain function have been challenging to study. Since the pioneering work of Hubel and Wiesel using the cat visual system as a model, decades' worth of animal experiments have provided valuable insight into the critical periods for brain development and heightened neuroplasticity that occur early in life [14–16,118]. But not all behaviors can be effectively investigated in animal models. For instance, no other species learns to comprehend and speak languages quite like humans do over early childhood. Furthermore, while laboratory models provide opportunities for powerful basic science research, translation into clinical neuroscience studies in patient populations requires methods for measuring brain function in humans. However, human studies of functional brain development have been constrained by available imaging methods and the entailed tradeoffs between image quality and logistics. For

example, functional magnetic resonance imaging (fMRI) has been used to map brain function in adults in exquisite detail, but the confined, loud, and solitary scanning environment and low tolerance for head motion are not conducive to scanning young children while awake [22]. As a result, a systematic review of over two decades of task fMRI research on language development found only one study in children younger than 4 years, which imaged sleeping toddlers during passive listening [20,21]. Functional near-infrared spectroscopy (fNIRS) is an alternative imaging modality that uses light to track brain function in an open, silent, child-friendly environment [29]. However, the traditionally sparse sampling of fNIRS results in poor image quality [30,56,61]. Indeed, developmental fNIRS studies have often had to describe differences between ages, groups, and task conditions with reference to measurement locations on the scalp rather than localizing them to brain tissue [32,33].

High-density diffuse optical tomography (HD-DOT) is an emerging neuroimaging approach that uses dense arrays of light sources and detectors to combine logistical advantages of optical imaging with tomographic reconstruction, increased resolution, reduced localization artifacts, and other image quality improvements relative to traditional fNIRS [30,52,56,61]. HD-DOT has been benchmarked against the current functional neuroimaging gold standard of fMRI [53,60] and has potential to be a powerful tool for developmental neuroscience research. However, prior HD-DOT studies have focused on adults [53], children older than 7 years [61], or infants [56,114,115], omitting the intermediate age range of 1-7 years. Imaging these young children while awake requires optimization of instrumentation for this age group as well as accompanying child-friendly task paradigms. Movies can serve as rich stimuli for mapping brain function

[35,36,64] while improving compliance and reducing head motion in children [28,34]. Here, we further developed optical imaging methods previously applied in adults, characterized movie viewing paradigms and data analysis methods initially designed by fMRI studies, and combined these tools to map responses to naturalistic speech in awake young children.

Specifically, we first designed and built a new HD-DOT system targeted towards imaging preschool-age children and incorporating various image quality improvements over previous systems, including an increased channel count, expanded field of view, and laser optics. After developing subject-specific head models for optimal reconstructions, we validated the performance of the imaging system based on raw data quality and reconstructed image quality for conventional task data from a group of adults. Next, we compiled and characterized a library of children's movie clips and associated data analysis pipelines, developing and validating these tools for functional brain mapping in adults. Finally, we used the optimized system, stimulus library, and analysis pipelines to map receptive language function during movie viewing in a pilot group of participants aged 23-80 months. We therein establish the feasibility of mapping brain function in young children while awake with HD-DOT. Our work lays an important methodological foundation for future research to further investigate typical and atypical trajectories of functional brain development through larger cross-sectional and longitudinal studies in various pediatric populations.

3.2 Methods

3.2.1 HD-DOT system design and build

The study aimed to first develop a new optical neuroimaging system geared towards imaging children while maximizing data quality. The instrument illuminated the head with 128 laser sources at each of two wavelengths, 685nm and 830nm (HL8338MG and HL6750MG, Thor Labs), well within American National Standards Institute consensus standards for near-infrared light exposure. Scattered light was measured by 125 avalanche photodiode (APD) detectors (C12703-01 SPL, Hamamatsu), and each one's signal was digitized by a dedicated analog-to-digital converter (ADC) at 96 kHz (A16R, Focusrite). Source, APD, and ADC boxes were stacked on a mobile extruded aluminum cart (custom design, MiniTec).

Light was conveyed between the participant's head and the sources and detectors themselves by bundles of fiber-optic cables (FTIIQ24425 for connecting 2 sources, one for each wavelength, to each source position on the cap, and FTIIG24426 for 1:1 detector connections, both from Fiberoptics Technology). The weight of the fibers was supported by and symmetrically distributed around a pair of concentric, suspended wooden halos (12" x 5/4" wood disks with a 8" cutout, Menards) and a mobile, extruded aluminum frame (custom design, MiniTec). Fiber tips were capped with 3-D printed, round-ended, soft plastic sleeves secured by heat-shrink tubing, so as to reduce the sharpness of optodes on the scalp. Fiber ends were held in place by an imaging cap in a grid of alternating source and detector positions that curved around the head surface.

The main body of the cap was made by molding four heated panels of thermoplastic Aquaplast (NC81339-1, North Coast Medical) to four surfaces (occipital, left and right temporal, and dorsal) of a model head. The model head was itself created by 3-D printing a previously acquired structural MR image from an adult participant [60] that had been scaled down to match the average diameter of a 4 year-old head. The Aquaplast panels were allowed to cool and harden before lining their inner surfaces with soft foam. The occipital panel was attached to each of the two temporal panels by a series of zip ties, while the dorsal panel was kept separate. The chosen thermoplastic material and this multi-panel design ensured a combination of structural integrity across long-term usage as well as flexibility to fit the cap to multiple head shapes and sizes. A uniform grid of holes with 11mm spacing was drilled across the formed panels, creating first-nearest through fifth-nearest source-detector separations of 1.1, 2.5, 3.3, 3.9 and 4.6 cm. The holes were then lined with 3-D printed, cylindrical plastic guides, secured by rubber bands, to hold fibers perpendicular to the head surface. The cap panels were suspended by string from the wooden halos, and then fibers were inserted in their designated positions through the guides. Fibers were held in place by elastic foam rings (punched out of a sheet, 8785K82, McMaster Carr) that also provided a spring action to press fiber tips against the participant's scalp. Fiber ends projected 5mm beyond the inner surface of the cap so that they could be combed through hair like the teeth of a hairbrush. Velcro straps and buckles were added along the edges of the cap so that the panels could be fastened in fixed positions relative to each other and pressed close against the head during imaging sessions.

The cap was suspended above a motorized salon chair that could be lowered or raised according to the participant's height. The height of the cap could also be changed by adjusting the frame but was generally held fixed across imaging sessions, except for the dorsal panel that was lowered and raised each time a participant entered and left the cap. An LCD display was positioned at eye level 90cm away from the cap to present visual stimuli and two speakers placed at the opposite end of the room were used to deliver auditory stimuli with stereo sound. A screen-based eye tracker (X3-120, Tobii Pro) was mounted below the display for a subset of imaging sessions. Seven cameras were mounted around the aluminum frame to capture the position of the cap on the participant's head during each imaging session from upper and lower viewing angles on the left, right, and back of the cap as well as head-on. The aluminum frame was cushioned and decorated with colored foam to have the appearance of a sandcastle, the aluminum cart was covered with blue curtains, and the walls and doors of the room were covered with stickers to give the entire imaging suite an under-the-sea theme. There was also ample space for one or more additional chairs to accommodate caregivers and/or experimenters sitting near children for comfort and supervision during imaging sessions.

The system was operated through three computers from an adjacent control room. Light sources were controlled by a dedicated source computer, using software created in-house and connected to a 20MHz digital input/output card (782608-01, National Instruments) and two BrainBoxes for multiplexing and demultiplexing of signals, as detailed for previous systems (Eggebrecht 2014). This setup enabled sources to be illuminated in highly temporally precise encoding patterns, allowing time and frequency coding of the source responsible for every detected signal. Detected

and digitized signals were sent from the ADCs to a separate detector computer, on which another custom-coded software program presented real-time displays of light and noise levels and also allowed for initiation and termination of data collection. Finally, a third stimulus computer used the Psychophysics Toolbox 3 package for MATLAB (Brainard, 1997) to present stimuli to participants via an audiovisual mixer connected to the two speakers and display. Synchronizing signals were sent from the source computer via the demultiplexing boxes and from the stimulus computer via the audiovisual mixer to the detector computer for time-locked alignment of source, detector, and stimulus information.

3.2.2 Cap fit procedure

The cap fit process featured slight modifications to the protocol used in previous studies in order to work with the expanded field of view [53,119]. At the beginning of each cap fit, any participant with long hair had their hair parted down the middle and back and then tied in left and right pigtails. The participant was then asked to sit in the imaging chair and slide their head back into the cap, and the pigtails (if present) were threaded between the lateral and dorsal cap panels to minimize obstruction of light transmission between optodes and the scalp. The height of the chair was adjusted to position the participant's ears just below the side panels of the cap. The participant was then directed to hold a pair of Velcro straps at the front of the cap and move the cap and their head side-to-side so as to comb the optodes through their hair and against the surface of the scalp. The tragus on either side of the head and fiducial markers on the cap were used as points of reference to ensure the cap was positioned symmetrically and consistently. The

front Velcro straps were then fastened. The dorsal imaging pad was lowered by an experimenter and combed through the hair on the top of the head, and additional Velcro straps were then used to secure the top panel relative to the rest of the cap. Real-time data quality visualizations and tactile feedback from the participant were then utilized to guide any further combing of individual occluded optodes with lowered light levels through obstructing hair. The final position of the cap on the participant's head was recorded from seven camera angles.

3.2.3 Participants

While our ultimate goal was to image young children, we needed to first validate the performance of our new imaging system, task paradigm, and analysis pipeline. A group of 5 young healthy adults (age range 20-33 years) served as a compliant, high-performing, well-characterized reference population to ensure that our imaging tools worked as expected. In the first phase of the study (HD-DOT system validation), we mapped responses to visual stimuli, auditory stimuli, and a motor task in 4 adult participants each. In the second phase (movie paradigm validation), we mapped responses to a library of animated movies in 3 highly sampled adults. In phase three (child imaging), we attempted to image 31 children (age range 18-80mo) while they were presented with movies and spoken word lists. Among them, 30 children complied with the cap fit procedure and wore the cap for at least 5 minutes of imaging. Our inclusion criteria for data analysis in the current study were that participants provide at least three high quality movie viewing runs (including two using the same movie clip, to allow for comparisons between matched and mismatched movie clips as explained in section 2.6.7) or one

high quality word-hearing task run. Data quality was determined by light levels, measurement variance and retention, pulse signal-to-noise ratio, and indicators of head motion, as explained in section 2.6. Altogether, 15 children (age range 23-80mo) met our criteria for inclusion in final analyses (Supplemental Table 2). Providing only 1-2 movie runs (insufficient to allow the comparison of 2 matched and 2 mismatched clips as described in section 2.6.7) was the most common reason other children were excluded.

Informed consent was obtained from all adult participants. For children, informed consent was obtained from the participant's parent or legal guardian, and assent was obtained from the child. Consent procedures were conducted in accordance with the IRB protocol approved by the Human Research Protection Office at Washington University School of Medicine.

3.2.4 Stimulus protocols

Visual Stimulation: Participants were asked to sit still and maintain fixation on a crosshair at the center of the display while a wedge-shaped flickering checkerboard stimulus rotated around the screen. The black and white checkerboard pattern of the wedge reversed at 8 Hz against a constant 50% grey background. The wedge subtended a polar angle of 60°, a radial angle from 2.5° to 10°, and rotated 10° at a time through 36 positions spanning 360° at 1 second per position, for a total of 10 cycles per run [119].

Finger Tapping: Participants were asked to sit still and maintain fixation on a central crosshair until either the letter "L" or the letter "R" was presented on the screen. In response to the letter

"L", participants were instructed to tap the fingers of their left hand against their left thumb at 2Hz. In response to the letter "R", participants were instructed to perform the same motion with their right hand instead. The letters were presented for 10 seconds at a time in a pseudorandomized order, interspersed with rest periods of 24 seconds, for a total of 8 finger tapping blocks per side per run.

Word Hearing: Participants were asked to sit still and maintain central fixation throughout the task. For adults, a simple fixation cross was presented, while children were presented a cartoon gif to hold their attention. Meanwhile, spoken word lists were presented through the speakers at 1 word per second for 15 seconds per block, with 15 seconds of silence in between blocks, and a total of 6 blocks per run. Participants were instructed to passively listen to the spoken words [53].

Movie Viewing: After an initial fixation period of 15 seconds, during which a silent gif from the movie Finding Nemo was presented on the screen, participants were presented with a 5-6 minute-long audiovisual movie clip. The presented clip was selected from a set of 20 options taken from 5 different children's movies and TV shows: Finding Nemo, Moana, Frozen, Curious George, and Daniel Tiger. Adult participants were presented with all of the movie clips over the course of multiple imaging sessions: generally, 4 different clips were presented twice each at each of 5 sessions. Children were allowed to choose which movie/show they wanted to watch, and were presented with 2-4 alternating clips from that program until they requested that the session end or up to a maximum of 8 movie viewing runs interspersed with 2 word-hearing task runs per imaging session.

3.2.5 Data processing

All acquired data was processed using the pipeline described below, built on methods developed and detailed in previous studies [52,53,58,109,119]:

Pre-processing: Raw light levels were converted to log ratios across time relative to the temporal mean of each measurement as a baseline. Channels with >7.5% variance across a run were omitted from further analysis, to minimize contamination of more subtle hemodynamic signals by non-physiological nuisance signals such as head motion [53]. Data was high-pass filtered using a 0.02-Hz cutoff to reduce long-term drift. Superficial signal regression was performed by averaging all first-nearest-neighbor measurements, which sample mostly the scalp, to estimate global systemic signals, and then linearly regressing this out of every source-detector measurement time trace [52,58]. Low-pass filtering was done with a 0.5-Hz cutoff to remove residual pulse and other high-frequency noise. Data were finally downsampled to 1 Hz [53].

Light modeling and reconstruction: Pre-processed data were reconstructed using an anatomical light propagation model. For the initial adult comparison of head modeling approaches (Figure 3.2) and the child data analysis (Figure 3.7), a generic head model was used, made from one participant's structural MRI data. For all other analyses of adult data, a subject-specific head model was generated, using each participant's own structural MRI data available from an independent research project. Head models were created using Freesurfer to segment a T1 image [82–84], NIRVIEW to generate a mesh of the participant's head, finite element modeling with spring relaxation approaches to position sources and detectors, and the NIRFAST toolbox to

model the diffusion of photons through the head [53,59,85]. For subject-specific head models, photographs from imaging sessions were used to guide initial placement of the optode array on the head mesh. A reconstruction of word-hearing task data using a first iteration of the head model was then compared to a reference group fMRI word-hearing task activation map [109] to iteratively guide adjustment of the head model.

The final sensitivity matrix A is a linear transformation between x, the vector of absorption coefficients within the brain volume, and y, the vector of relative changes in light levels measured at the head surface, for each time point, in accordance with the linear Rytov approximation:

$$(1)\ y = A\,x$$

The sensitivity matrix was inverted using Tikhonov regularization with λ_1=0.05 and spatially variant regularization with λ_2 = 0.1. Flat-field reconstructions with these regularization parameters were used to designate voxels with sufficient sensitivity [57] and Freesurfer segmentation results were used to mask out signals from superficial tissues. The intersection across 3 highly sampled participants of flat-field reconstructions thresholded at 1% of their maxima was used to designate a common field of view for visualization purposes, as in previous studies [75]. For quantitative comparisons of different maps, a more stringent sensitivity cutoff was applied for each head model by thresholding its flat-field reconstruction at 10% of its maximum value to generate a conservative, subject-specific spatial mask.

Spectroscopy: Relative changes in oxy- and deoxy- hemoglobin concentrations were calculated from the differential absorption image x at each time point through spectral decomposition:

$$(2)\ \Delta C = E^{-1}x$$

where ΔC is a vector of oxy- and deoxy- hemoglobin concentration changes across the brain volume, and E is a matrix of extinction coefficients of oxy- and deoxy- hemoglobin [53,119].

Motion scrubbing: For every run of data, the global variance of temporal derivatives (GVTD) metric of head motion was calculated at each time point as the root-mean-square of the temporal derivatives of all measurements [109]:

$$(3)\ g_t = \sqrt{\frac{1}{N}\sum_{m=1}^{N}(y_{mt} - y_{m(t-1)})^2}$$

Where g_t is the GVTD value at time t, m indexes measurements, N is the total number of measurements, and y_{mt} is the optical density change for measurement m at time t. Time points for which g_t rose above a statistically determined threshold g_{thresh} were censored from further analysis to prevent confounding of results by motion artifact. The value of g_{thresh} varied between participants, sessions, and runs due to differences in the baseline resulting from physiological variation, but was set through a data-driven method [109]:

$$(4)\ g_{thresh} = \kappa + 10\sigma_L$$

where κ is the mode of GVTD values and σ_L is calculated across the number n_L of GVTD values less than κ as:

$$(5)\ \sigma_L = \sqrt{\frac{1}{n_L}\sum_{g_t<\kappa}(g_t - \kappa)^2}$$

Data quality evaluation: The quality of every data set was evaluated with regard to its light fall-off curve, the proportion of measurements retained below the 7.5% variance cutoff, GVTD time course, pulse signal-to-noise ratio (SNR) across the cap, and for adult imaging sessions also the quality of control word-hearing task activation maps. Sessions and runs with poor data quality based on these metrics were excluded from further analysis.

3.2.6 Task data analysis:

For the word-hearing tasks, a general linear model (GLM) was used to calculate beta values of stimulus response relative to rest for each run, using a canonical adult hemodynamic response function that had previously been empirically determined using HD-DOT [120]. A GLM was similarly used to analyze finger tapping data, but for this lateralized task, the contrast between left-sided and right-sided responses was mapped by calculating the difference in beta values between those two task conditions. As the continuous visual stimulation task was not conducive to an event design GLM analysis, responses were instead block-averaged across a 6 second time window beginning 7 seconds after the wedge had entered the lower left and lower right visual quadrants, based on prior studies using the same stimuli [119]. For all these tasks used to validate the system, responses were aggregated across runs and participants by calculating a fixed effects t-statistic map [109]. Response time courses were plotted by averaging responses in seed regions

of interest (selected for each task from their activation maps) across blocks, runs, and participants.

3.2.7 Movie viewing data analysis:

Movie-viewing data runs were aligned with each other and with feature regressors by calculating cross-correlations between the attached audio signals (which were recorded with the HD-DOT data through one of the ADC channels) and then shifting data time courses by the lag that maximized the correlation between the audio signals. This approach corrected for any variable stimulus delays caused by the experimenter or equipment, allowing highly timing-sensitive analyses like feature regression to be accurately time-locked [28].

To study inter-run synchronization, voxel-wise correlations were calculated between the oxyhemoglobin signals across time from the two movie runs of interest [35,37,64]. Random effects t-statistic maps were constructed to aggregate the results across multiple pairs of runs and allow comparisons between conditions, e.g. across all matched movie run pairs, or across mismatched movie run pairs as a negative control.

For feature regressor analysis, speech and face regressors were coded for every movie clip in the library by 3 independent experimenters, using a protocol adapted from previous studies [36,64]. Speech content was scored on a binary basis (present or not) in 1 second bins across each movie clip. The salience of faces in the scene was graded on a 4-point scale (0 = absent, 1 = present but in the background, 2 = present and salient, 3 = filling the screen or otherwise highly salient) for

frames sampled at 1 second intervals across each movie clip. Each coded regressor was averaged across the 3 experimenter's ratings, with a final reviewer resolving major disagreements at any time points attributable to human error (e.g. if one experimenter rated faces as absent and another scored faces as highly salient for the same frame). The resulting consensus regressors were then convolved with a canonical hemodynamic response function [120] and z-scored. Both regressors and HD-DOT data were filtered with the same high-pass cutoff of 0.02 and low-pass cutoff of 0.2 prior to correlation analysis.

To conduct univariate regressor analysis (e.g. simple speech mapping) for a run of data, voxel-wise correlations were calculated across time between the oxyhemoglobin signal and the regressor of interest, as in previous studies [64]. To conduct a more refined multivariate regressor analysis for parallel feature mapping, voxel-wise feature correlations were evaluated by calculating beta values using a GLM incorporating multiple features of interest (both speech and faces) and nuisance regressors (GVTD, audio envelope, luminance, temporal derivative of luminance, temporal derivative of squared luminance, hands, and bodies) in the design matrix.

In all cases, results were aggregated across runs by calculating random effects t-statistics and plotting them on an atlas surface. The resulting feature regressor maps were further evaluated by calculating spatial correlations, e.g. between feature regressor maps from different sets of movies to evaluate reproducibility and generalizability, or between speech regressor maps and word-hearing task activation maps to evaluate construct validity.

3.3 Results

3.3.1 Preschooler HD-DOT imaging system:

Our new HD-DOT system incorporated design features to maximize both image quality and child comfort (Figure 3.1A). Measures to improve image quality included a high channel count and density, an expanded field of view, and enhanced optics relative to previous systems [53,61]. The effort to make the imaging setup child-friendly targeted the cap itself as well as the associated infrastructure and scanning environment.

The cap contained 128 source locations interleaved with 125 detector locations in a grid with 11mm spacing between neighboring optodes, capturing 5202 measurements with <5cm of source-detector separation (Figure 3.1B). These measurements were distributed across the posterior, right and left lateral, and dorsal surfaces of the head (Figure 3.1C), so as to cover regions of occipital, temporal, parietal, and frontal cortex including visual, auditory, sensorimotor, and association areas (Figure 3.1D-E). To maximize the dynamic range of measurements, laser light sources were used with a "2-pass encoding pattern", i.e. every source was flickered in turn at first 1% brightness and then 50% brightness. This approach yielded light fall-off curves spanning seven orders of magnitude of optical power above the noise floor without clipping – a wider range of light levels than those effectively measurable with any single intensity illumination, as evidenced by data acquired with an optical phantom (Figure 3.1F-I). The sensitivity, detectivity, crosstalk, dynamic range, and frame rate of image acquisition are detailed in Table 3.1.

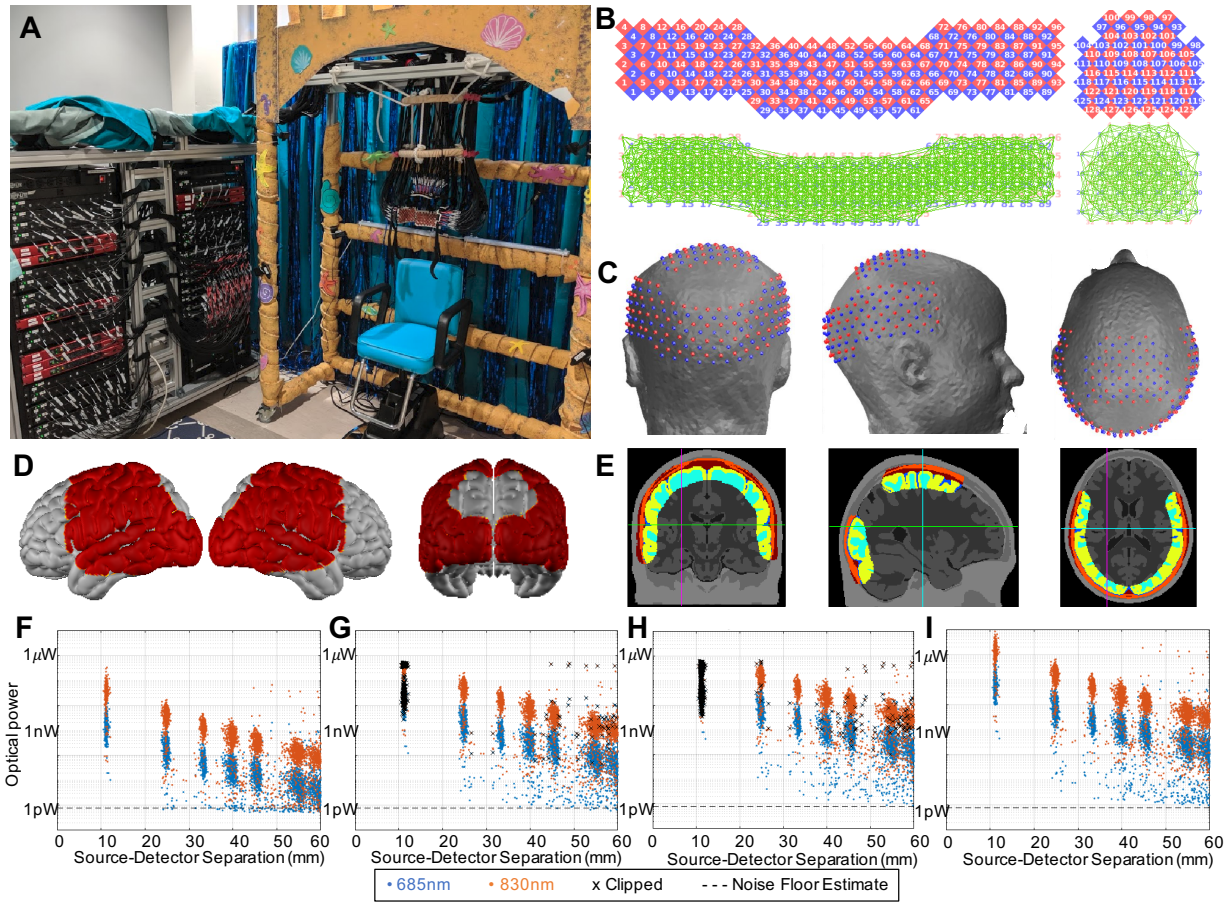

Figure 3.1: Preschooler HD-DOT system
(A) Image of the system, showing the cart stacked with sources and detectors, and the imaging cap suspended above a pediatric chair from a decorated frame. (B) Schematic of the optode array, depicting 128 sources in red, 125 detectors in blue, and thousands of overlapping source-detector pairs below in green. Not depicted are additional measurements connecting the dorsal pad and the rest of the cap between the sources and detectors along their adjacent edges. (C) Depiction of source and detector positions on a participant's head during a typical imaging session. (D) Surface projection of a flat field reconstruction thresholded at 10% of its maximum illustrates the coverage typically achieved, including regions of occipital, parietal, temporal, and frontal cortex. (E) Tomographic slices of a flat field reconstruction thresholded at 10% of its maximum illustrate the depth of sensitivity. (F) Illuminating sources at 1% of their maximum intensity allows for unsaturated nearest-neighbor measurements but low light levels that approach the noise floor for long-range measurements. (G-H) At 10% (G) and 50% (H) intensities, light levels remain higher above the noise floor for long distances, but are clipped for short ones. (I) A 2-pass encoding pattern alternating between 1% and 50% brightness supports unclipped measurements for short-range separations as well as measurements well above the noise floor for longer separations, maximizing dynamic range.

Table 3.1: Preschooler HD-DOT system characterization

System Specification	Goal	Actual (mean)
Sensitivity (V/W)	$> 1x10^6$	$6x10^6$
Detectivity (fW/[√Hz*mm^2])	< 350	10.6
Crosstalk (dB)	< -120	-131.5
Dynamic range (dB)	> 120	120.1
Frame rate (Hz)	> 3	10.4

The cap was sized to fit an average 4 year-old head, but the multi-panel design, flexible materials, and adjustable Velcro attachments allowed the cap to fit a range of head sizes and shapes, tested across participants ranging from 18 months to 33 years of age. As with previous caps, fiber tips extended beyond the inner surface of the cap so as to comb through hair, but they were covered with newly designed, rounded, soft plastic caps to reduce sharpness at points of contact. The large number of fibers further distributed pressure across the head. Lightweight fiber optic cables were used to convey light between the cap and imaging console, and their collective weight, tension, and torque were supported and evenly distributed around a system of two wooden halos and an extruded aluminum Mini-Tech frame, further minimizing the weight of the cap on the participant's head. The frame was padded with foam and decorated to look like a sandcastle to fit with the under-the-sea theme of the imaging suite, creating a child-friendly aesthetic. During imaging sessions, participants were seated in a motorized, adjustable, and cushioned pediatric salon chair with a five-point harness to secure younger children. The open imaging environment also included ample space for a caregiver and/or experimenter to sit next to children during imaging sessions to supervise them and keep them at ease.

3.3.2 Head modeling

Finite element algorithms model light propagation through the head between source-detector pairs, enabling source-detector measurements to be transformed into a neuroanatomical space [53,85]. While generic atlas-derived head models can be effective [81], we created subject-specific head models using participants' own structural MRI data when it was available, i.e. for

all adults who were imaged for validation of the system. The rationale for this approach was to account for individual variation in head shape and internal tissue structure and thereby ensure more accurate system validation. Initial positioning of the optode array in each head model was guided by photographs of the cap's true position on the participant's head captured from 7 camera angles during each imaging session (Figure 3.2A). Word-hearing task data was used as a functional localizer to guide further task-based adjustment of head models (Figure 3.2B) by comparison to reference fMRI maps [109]. In order to evaluate head model performance, independent runs of the word-hearing task that were not used for head model training were reconstructed using both generic and subject-specific head models. The resulting test maps were compared to reference word-hearing task maps from a prior fMRI study [109] by plotting overlap and calculating Dice coefficients across a range of thresholds. The subject-specific head models consistently yielded greater overlap with the reference data, quantified by higher Dice coefficients, than the generic head model (Figure 3.2C-D).

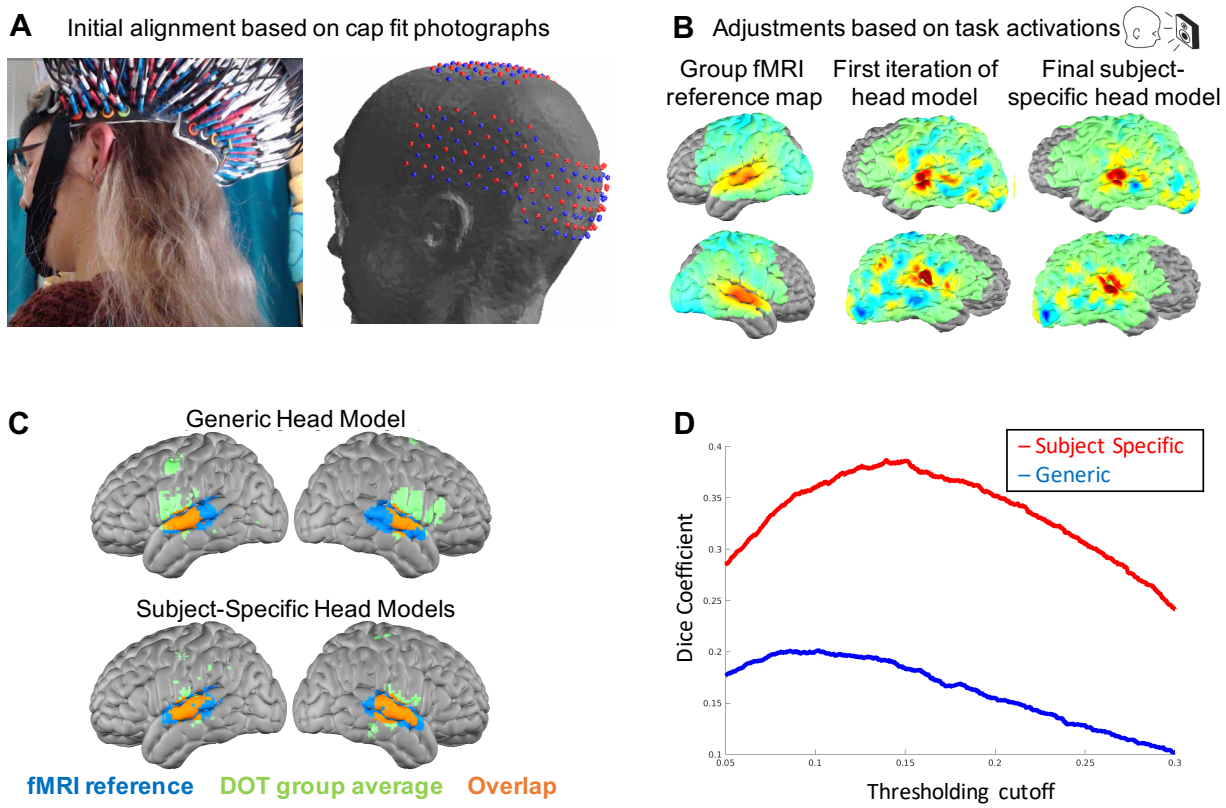

Figure 3.2: Head modeling
When anatomical MRI data was available for a participant (i.e. for all adults in validation phases of the study), it was used to create a subject-specific head model to attempt optimized reconstructions of HD-DOT data. (A) Cap fit photographs taken from 7 camera angles were used to guide positioning of the optode array on a mesh of the participant's head for the first iteration of the subject-specific head model. (B) One participant-specific word-hearing task run was used as training data by an experimenter to inform adjustments to the head model based on how the reconstructed data compared to a reference group fMRI word-hearing activation map. (C) Reconstructions of separate test data using subject-specific head models show increased overlap with reference fMRI data relative to reconstructions using a generic head model. (D) The Dice coefficient used to quantify overlap between group HD-DOT data and group fMRI data is greater across map thresholds for subject-specific head models than for a generic head model.

3.3.3 System validation

In vivo data was first collected in adults to validate the performance of the new system using simple tasks in a compliant and well-characterized reference population before pursuing studies of more complex paradigms like movie viewing or more challenging populations like children. We regularly observed a log-linear fall-off of light intensity with source-detector separation, with a clustering of measurements within ~2 orders of magnitude at each separation, and with measurements consistently above the noise floor out to fifth-nearest neighbor pairs (Figure 3.3A). The cardiac pulse waveform was clearly visible in measurement time traces (Figure 3.3B) and a corresponding peak around 1Hz was notable in Fourier spectra of measurements (Figure 3.3C). Plotting the ratio of signal power in the 0.5-2 Hz frequency band to the bandwidth-scaled median power in flanking frequency ranges revealed high pulse SNR across the cap, although SNR was frequently higher across the lateral and posterior panels than the dorsal panel (Figure 3.3D). Other indicators of high *in vivo* data quality that were assessed for every scan included low measurement variance, high retention of measurements after excluding those above a 7.5% variance threshold, and low head motion as quantified by the global variance of temporal derivatives. Task activation maps provided further evidence of image quality. We were able to map contralateral visual cortex activations in response to unilateral checkerboard stimuli (Figure 3.3E-F), contralateral motor cortex activation during finger tapping (Figure 3.3G), and bilateral auditory cortex activations during a word-hearing task (Figure 3.3H), as well as map corresponding time courses of oxyhemoglobin, deoxyhemoglobin, and total hemoglobin signals (Figure 3.3I-L).

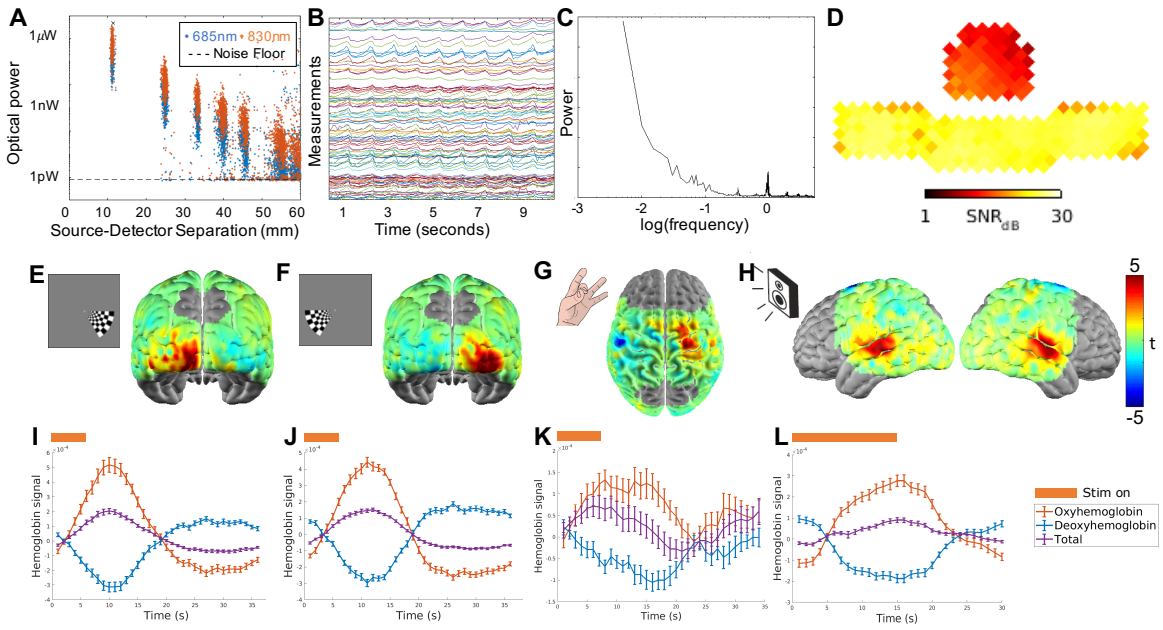

Figure 3.3: System validation with in vivo adult data
Young healthy adults served as a compliant, high performing reference group to help validate the performance of the new imaging system. (A) A representative light fall-off plot, showing a log-linear fall off of light levels with increasing source-detector separation. Optical power measurements are tightly clustered within 2 orders of magnitude at each possible distance, and are above the noise floor across 5 degrees of source-detector separation. (B) The cardiac pulse is clearly visible in measurement time traces. (C) Fourier spectra show a strong cardiac pulse peak at ~1Hz. (D) High mean 0.5-2Hz band-limited signal-to-noise ratio (pulse SNR) across the cap indicates high data quality. SNR tended to be higher across the posterior and lateral panels than the dorsal panel. (E-H) Activation maps from block-design tasks illustrate image quality across the cap. (E) A right-sided visual stimulus induces measurable activations in left visual cortex. (F) A left-sided visual stimulus induces measurable activations in right visual cortex. (G) A finger tapping task evokes contralateral motor cortex activity. (H) A word-hearing task evokes bilateral auditory cortex activations. (I-L) Corresponding plots of oxy-, deoxy- and total hemoglobin levels over time, showing measured hemodynamic responses during each of the tasks.

3.3.4 Characterization of children's stimulus library

While block-design tasks are simple to analyze and provide an effective means of mapping brain function with high SNR in a compliant population of adults, young children are less likely to cooperate with paradigms involving extended periods of central fixation and monotonous stimuli such as flickering checkerboards and spoken word lists. Movie viewing is a complex but more engaging task that has been used to map brain function in fMRI studies [28,35,36] and a recent HD-DOT study in adults [64], and has also been shown to reduce head motion in children [34]. We compiled a library of children's movie clips and an associated data analysis pipeline, and sought to validate their effectiveness as tools for functional brain mapping with our new HD-DOT system, again first using adults as a compliant reference population. With the new system and movie viewing paradigm, we were able to capture reproducible responses to movies over independent viewings of the same clip across large swaths of temporal, parietal, and occipital cortex in particular (Figure 3.4A). Mapping correlations between responses to mismatched movie clips did not yield the same inter-run synchronization (Figure 3.4B). However, any movie clip, or even a combination of clips, could be used to map responses to specific movie features through feature regressor analysis. For instance, by coding a speech regressor for each movie and computing correlations with the oxyhemoglobin signal at every voxel in our field of view, we constructed speech regressor maps that highlighted regions putatively involved in processing speech (Figure 3.4C). Speech regressor maps were similar across two completely independent sets of various movie clips (Figure 3.4C-D), and also appeared comparable to activation maps from session-matched, block-design, word-hearing task data (Figure 3.4E).

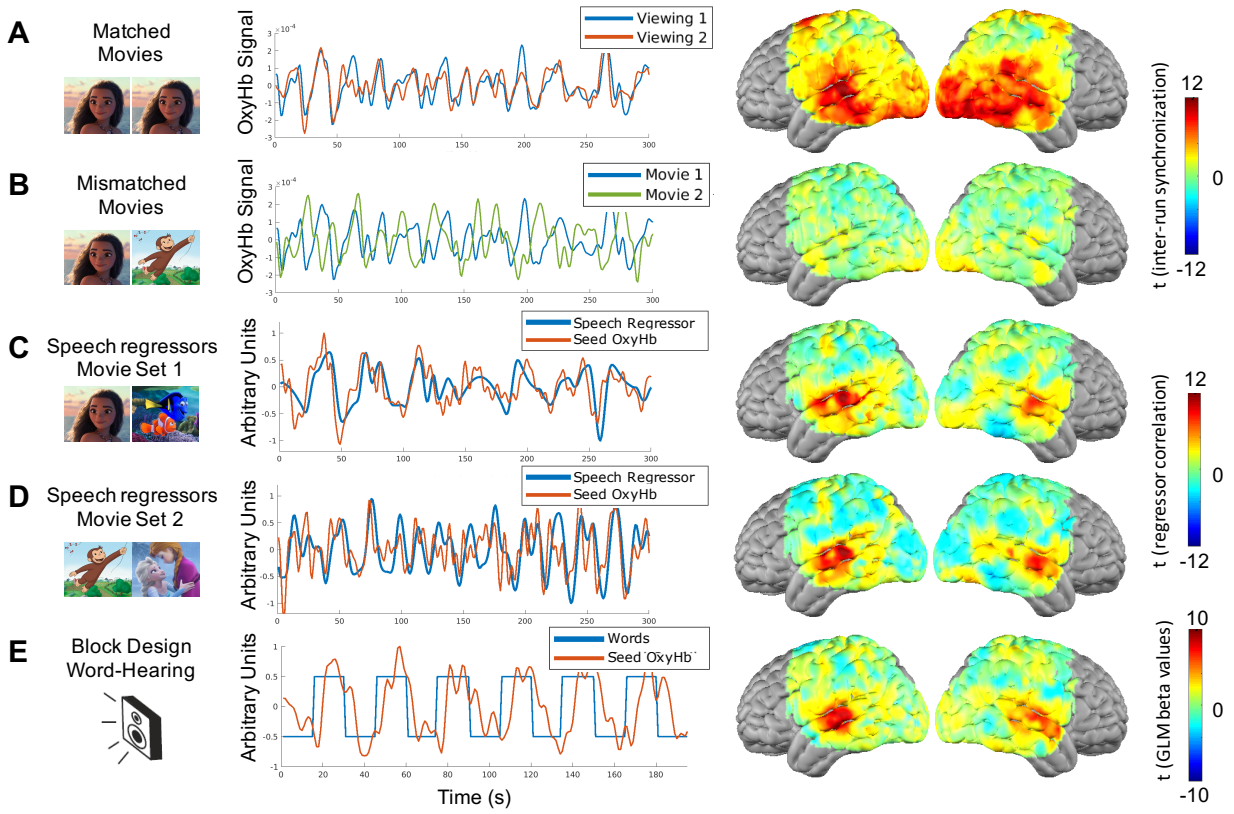

Figure 3.4: Functional brain mapping using movies
Clips from animated children's movies were used as stimuli for mapping brain function, first in a small group of highly sampled adults (80 movie viewing runs in total across 3 adults). (A) Movies evoked reproducible patterns of brain activity, plotted here for an individual voxel during two viewings of the same movie, and mapped across the brain as voxel-wise inter-run signal correlations over time for 40 pairs of movie viewing runs. (B) These responses are movie-specific – comparing responses across 40 pairs of mismatched movie clips does not reveal the same inter-run synchronization seen with matched clips. (C) Nevertheless, a heterogeneous set of movies can be used to map responses to movie features such as speech through regressor correlation analysis. An exemplary speech regressor time course is plotted here along with the oxyhemoglobin signal in a voxel of the brain that appears to be responsive to speech. A speech regressor map highlights regions putatively involved in speech processing based on their high correlations with speech regressors across 32 movie viewing runs using 8 different movie clips. (D) Speech regressors may have different time courses for different movies, but feature regressor analysis produces maps that are comparable nonetheless. A second group speech regressor map is shown here from a separate set of 34 movie viewing runs using 8 completely different movie clips. (E) Activation maps from session-matched, block-design, word-hearing task data are comparable to movie-derived speech regressor maps

Capitalizing on the plurality of movie clips in our stimulus library, we further assessed potential factors influencing the quality of speech regressor maps. To evaluate whether the specific movie that was presented mattered, we constructed two independent group level regressor maps, using two separate sets of movie viewing runs, for each of the movies and shows from which stimulus clips were taken. We then calculated spatial correlations between every possible pairing of the resulting regressor maps, measuring reproducibility, as well as spatial correlations between each movie's speech regressor map and a reference word-hearing task map, measuring construct validity (Figure 3.5A). Across all but one of the movies, both reproducibility and construct validity were consistently high (mean ± standard deviation of spatial correlations = 0.56 ± 0.12 for reproducibility and 0.45 ± 0.06 for construct validity), the exception being the show "Daniel Tiger". As a negative control, maps constructed for the same movies and runs using speech regressors coded for mismatched clips did not exhibit the same consistency (Figure 3.5B; reproducibility = 0.11 ± 0.12 and construct validity = 0.09 ± 0.11). Reproducibility of speech regressor maps between independent viewings of the same clip by the same participant and construct validity with reference to word-hearing task maps were both positively correlated with mean inter-run synchronization across the field of view (Figure 3.5C-D). Poorer speech regressor maps, e.g. as seen for the show Daniel Tiger, also appeared to be associated with clips involving less modulation of speech content over time (Figure 3.5E-F).

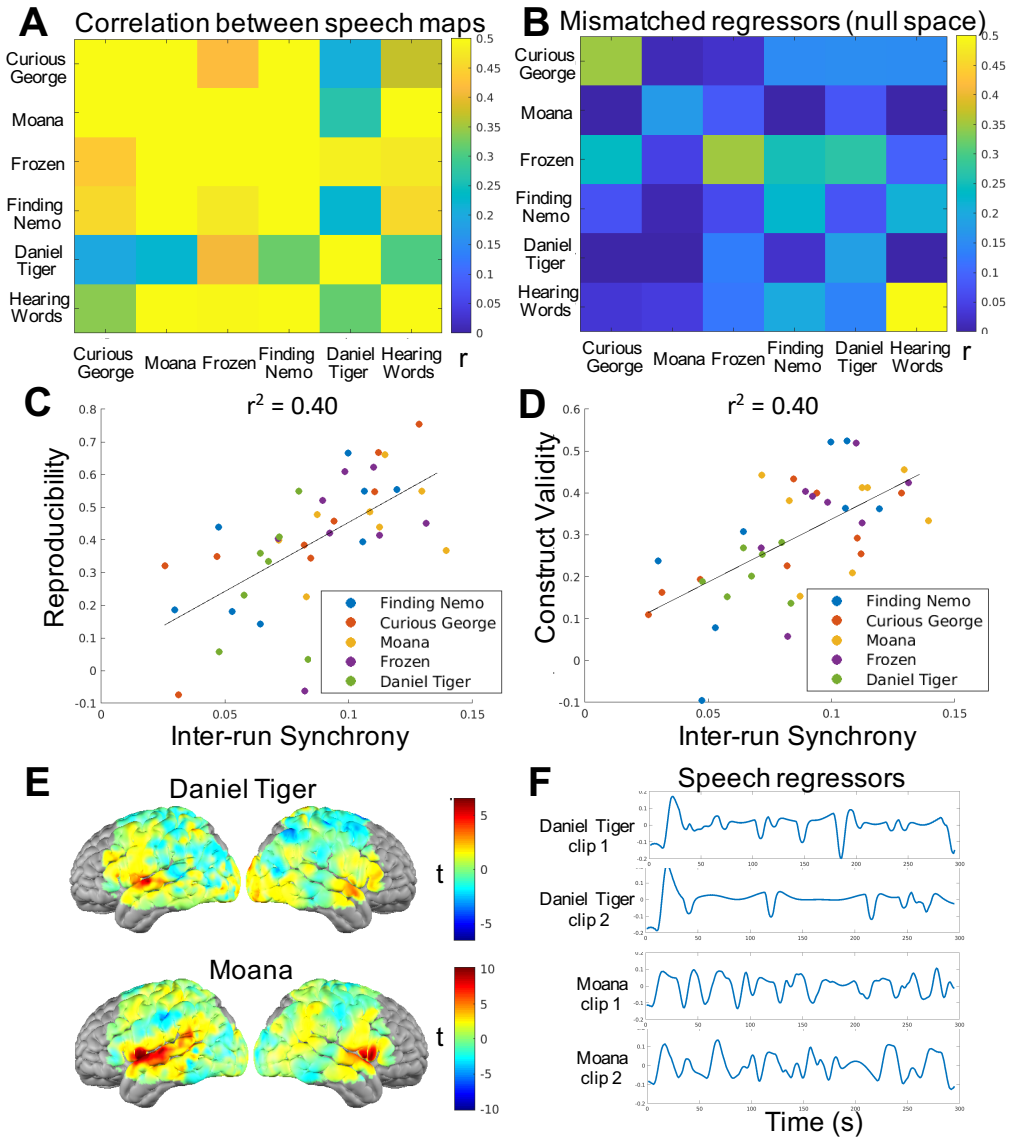

Figure 3.5: Comparison of regressor maps across runs and movies

We utilized our rich movie-viewing data set collected across a variety of movies to study potential factors affecting the quality of regressor maps. (A) Correlation matrix comparing speech regressor maps from participants viewing clips from 5 different children's movies/shows as well as word-hearing task activation maps. High correlations between independent viewings of the same movie (along the main diagonal), across movies (off diagonal; with the exception of Daniel Tiger), and with the word-hearing task (final row and column) illustrate the reproducibility, generalizability, and construct validity of movie speech regressor analysis. (B) As a negative control, maps made with speech regressors from the incorrect movie clip exhibit low correlations with each other and with word-hearing task activation maps. (C)

Reproducibility of speech regressor maps (measured as correlations between speech regressor maps from separate viewings of the same movie clip) is positively correlated with mean inter-run synchrony. (D) Construct validity of speech regressor maps (measured as correlations between speech regressor maps and word-hearing task activation maps) is positively correlated with mean inter-run synchrony. (E) Regressor maps appear worse for some movie clips (e.g. Daniel Tiger) than others (e.g. Moana). (F) Modulation of speech across clips is lower for some clips (e.g. Daniel Tiger) than others (e.g. Moana).

Because movies are such rich, multi-dimensional stimuli, they should also be conducive to mapping multiple types of processing in parallel [36,64]. To assess this point with the new imaging system and stimulus library, we also coded face regressors for all our movies, indicating the salience of faces across the clips. For some clips, these features appeared linearly independent, while for other clips they exhibited considerable covariance over time (Figure 3.6A). Plotting face and speech regressor maps from the same data produced partially overlapping maps (Figure 3.6B). The similarity between face and speech maps from a given run was found to be strongly correlated with the similarity between the face and speech regressors for the movie presented (Figure 3.6C). In contrast, speech regressor maps from a participant presented with movie audio unaccompanied by visuals and face regressor maps from the same participant when presented with silent visual clips stripped of the audio were more distinct, with the face maps notably more right-lateralized (Figure 3.6D). In an attempt to mitigate confounding of speech and face regressor maps from the audiovisual movie viewing data by the covariance between these two movie features, we reran the analysis using a multivariate regression approach. We employed a general linear model (GLM) including both regressors of interest in its design matrix, alongside a series of other nuisance regressors tracking participant motion, stimulus audio envelope, and low-level visual stimulus features such as luminance. This multivariate regression approach yielded more distinct speech and face regressor maps (Figure 3.6E), comparable to those obtained with the purely auditory and purely visual stimuli (Figure 3.6D). Furthermore, the multivariate regression approach also abolished the previously strong positive correlation between regressor map similarity and regressor similarity (Figure 3.6F).

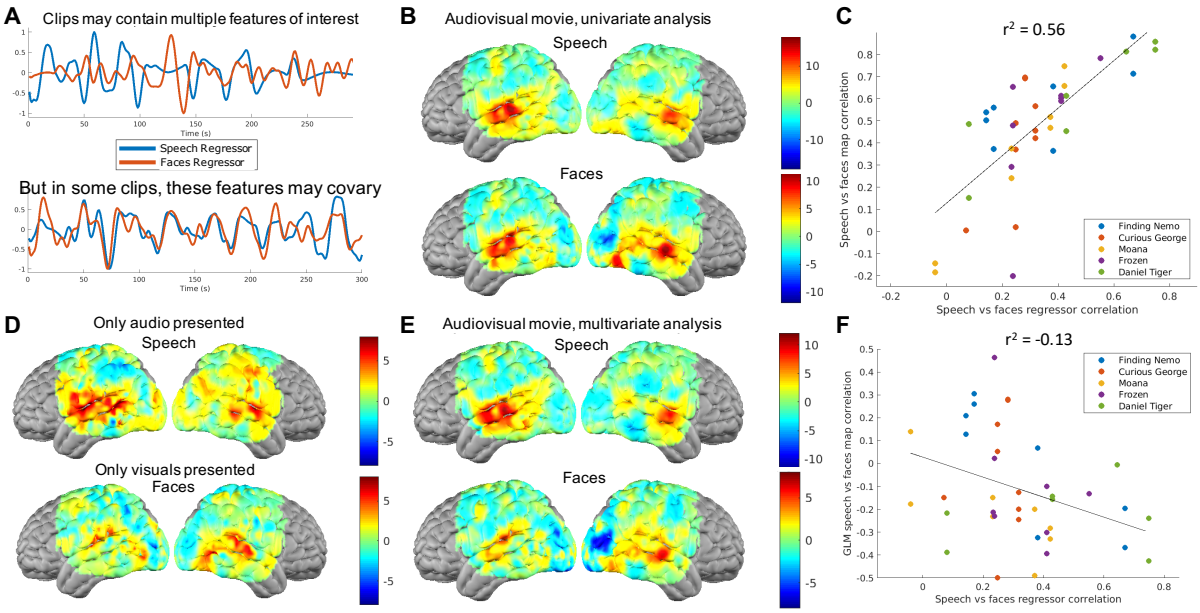

Figure 3.6: Parallel feature mapping
(A) Because movies are such rich stimuli, movie-viewing data can be used to map responses to multiple features of interest present in a movie, e.g. both speech and faces. However, these features of interest may be more independent for some movies and more correlated with one another for other clips. (B) Perhaps as a result, feature correlation maps for speech and faces display some overlap. (C) The overlap between speech and face regressor maps for a movie is strongly correlated with the similarity of the speech and face regressors for that movie. (D) When a participant was presented with just the audio (containing speech and no faces) or just the visuals (containing faces and no speech) of a movie to remove confounds, speech and face regressor maps appeared more distinct. (E) Similarly distinct maps could be obtained from the audiovisual movie viewing data using a GLM-based multivariate regression approach. (F) The multivariate regression abolished the strong positive correlation between regressor map overlap and regressor similarity that was seen with the univariate approach in panel C above.

3.3.5 Functional brain mapping in awake young children

Having developed and validated a new imaging system, stimulus library, and data analysis pipelines, we sought to combine these tools to map responses to speech in young children with HD-DOT. Having previously studied infants while asleep [114,115] and children older than 7 while awake [61], we recruited children between 1 and 7 years of age and attempted to image them while awake (Figure 3.7A). Children appeared to be engaged by presenting them with a movie, based on both experimenter observations as well as eye tracking data. Children kept their gaze on the screen when given a movie to watch more than when asked to maintain central fixation during a conventional word-hearing task (Figure 3.7B). Mapping correlations between the oxyhemoglobin signal evoked by independent viewings of the same movie illustrated high inter-run synchronization across our field of view, especially in temporal cortex (Figure 3.7C). Correlations between responses to mismatched movies were also computed as a negative control (Figure 3.7D), indicating the specificity with which we were able to map movie-evoked responses. Speech regressor analysis yielded speech maps with high correlations centered over the superior temporal gyrus bilaterally but with stronger correlations in the left hemisphere (Figure 3.7E), as was seen in adults. This response pattern was also similar to the activation map obtained from the smaller number of successful block-design word-hearing task runs in children who complied with this task (Figure 3.7F).

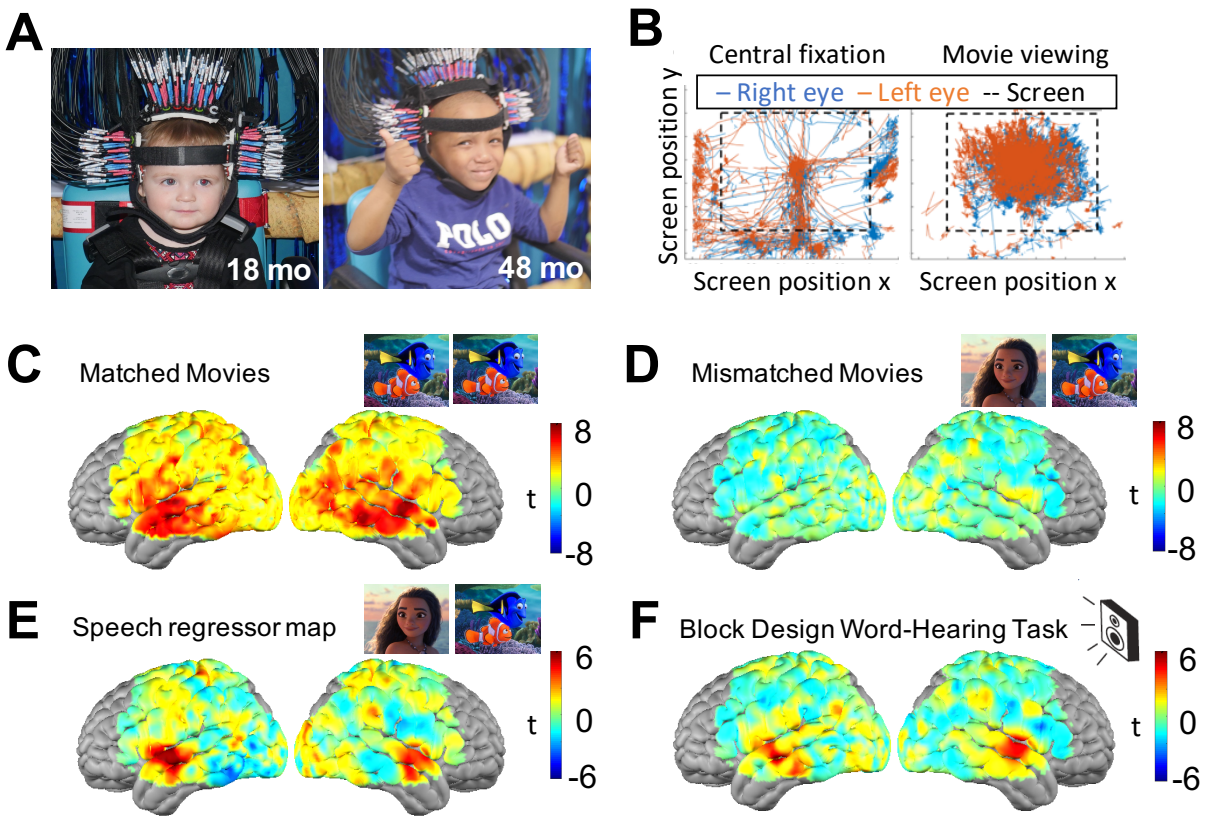

Figure 3.7: Functional brain mapping during movie viewing in 1-6 year-old children
(A) Children ranging from 1 to 6 years of age were successfully imaged using the new system. (B) Eye tracking data shows a child attending to the screen more when presented with movies than when asked to maintain fixation on a central crosshair. (C) High inter-run synchronization is seen in responses to matched movie clips in young children across 38 movie viewing runs from 10 participants. (D) Evoked responses are movie-specific, and inter-run synchronization is not seen across mismatched movie viewing runs. (E) Feature regressor analysis can be used to map receptive language from HD-DOT movie viewing data collected in awake young children. (F) Comparable responses are seen across 17 block-design word-hearing task runs from children who complied with this task.

3.4 Discussion

In summary, we developed a new HD-DOT imaging system targeted towards imaging young children and designed to enhance both image quality and participant comfort. In addition, we compiled a library of animated movies and associated data analysis pipelines for functional brain mapping with child-friendly stimuli. After characterizing and optimizing the imaging system, stimulus library, and analysis methods in adults, we combined these tools to map responses to naturalistic speech in awake young children. We therein establish the utility of HD-DOT imaging combined with a movie viewing paradigm for mapping brain function in early development.

3.4.1 Importance of method development in pediatric functional neuroimaging

While the field of cognitive neuroscience has advanced considerably over the past several decades with the advent and expansion of fMRI research, human developmental studies have been constrained by the challenges of pediatric imaging. Young children are rarely comfortable when asked to lie alone, awake, and still for an extended duration inside the dark, cramped, noisy bore of the MRI scanner [22]. As a result, studies of brain development during the first few years of life have mostly imaged infants while asleep and have relied on resting state functional connectivity analysis to assess the functional architecture of the developing brain [23]. While powerful, this approach does not account for differences in functional networks between awake and sleep states [26,27] and also does not allow for investigation of active processing during diverse task states. Only a few studies have attempted to image infants while awake and

attending to stimuli in an MRI scanner [121,122]. For example, one innovative study had parents lay inside the bore of the scanner with their infants to provide supervision and comfort while they observed visual stimuli, but the experimenters still reported high levels of participant attrition and data exclusion as a result of challenges with compliance and motion artifact [122]. When MRI scans are considered medically necessary for clinical diagnosis, sedation may be used to acquire data and reduce motion in children. However, this approach is avoided when possible due to the health risks of sedation in young children [22,123], is hence unsuitable for widespread use in academic research in the absence of a clinical indication, and is further undesirable for cognitive neuroscience research as anesthesia inherently alters brain function [26,124]. As children grow older, some can learn to comply with the logistics of fMRI through training and practice with mock scanners [125,126]. However, very few studies have reported success with fMRI in awake children younger than 4 years of age [20,22,28].

FNIRS is an alternative imaging modality that has become increasingly popular for developmental research. Using near-infrared light to sample the cortical surface through intact skin and scalp, fNIRS capitalizes on differences in the optical properties of oxygenated and deoxygenated hemoglobin to track blood oxygenation as a surrogate marker of neural activity [29]. In contrast to MRI scanners, fNIRS imaging caps can be used in an open scanning environment that allows pediatric participants to sit comfortably next to caregivers or even in their laps in a child-friendly space [32,33]. FNIRS is therefore conducive to imaging children as well as studying naturalistic behaviors ranging from natural speech to parent-child interactions [127,128]. However, the sparse arrays of light sources and detectors traditionally used by fNIRS

studies result in low spatial resolution and do not allow for accurate anatomical reconstructions of data [30]. Many fNIRS developmental studies hence describe differences between groups and conditions with reference to optode locations on the scalp [32,33], which are not the most biologically relevant unit of measurement. The low channel count is also usually associated with a limited field of view [54]. Furthermore, sparse sampling can lower signal-to-noise ratio [56], cause signal localization errors as artifacts of optode geometry [30,61], and allow contamination of cortical measurements by superficial tissue signals [58]. These artifacts are likely to be especially problematic for mapping distributed changes across the brain associated with complex processes like naturalistic behaviors and brain development. Therefore, the interpretation and applications of fNIRS imaging have been constrained by low image quality.

HD-DOT uses dense arrays of light sources and detectors to support accurate tomographic reconstruction, with improved lateral resolution, increased depth sensitivity, reduced contamination of cortical measurements by superficial signals, and fewer localization artifacts relative to sparse fNIRS grids [30,58,61]. These image quality enhancements have enabled HD-DOT to produce maps of both localized and distributed brain activity patterns comparable to those obtained with fMRI [53,60,64,75]. Furthermore, the imaging approach and its applications continue to develop rapidly, with recent advances including highly portable systems used in low resource settings [61], wearable systems used to image awake infants [56], mapping of responses to naturalistic movies in adults [64], and decoding of stimulus information from HD-DOT data [119]. Despite these promising studies in adults, older children, and infants, HD-DOT had yet to be applied for imaging in the especially challenging age range of 1-6 years. By developing a HD-

DOT system for this age range and applying it to map complex brain responses during naturalistic viewing, the current study establishes methods that could help fill an important void in the developmental cognitive neuroscience literature.

3.4.2 Design and validation of HD-DOT system

In addition to the cap and imaging suite features implemented to increase child comfort, our new HD-DOT system incorporated several attributes to improve image quality relative to previous optical neuroimaging systems.

The channel count – with 5202 source-detector measurements within 5cm of separation – is orders of magnitude higher than traditional fNIRS systems. Much prior pediatric fNIRS research has relied on systems with 1 to 50 channels [54], and even the few more recent studies using high-density systems have only reported using around 300-500 channels [56,61]. Several prior HD-DOT studies in adults described using ~1200 channels [53,64,119], but the current system still represents an expansion. The numerous channels cover a wide field of view and are also packed at a high measurement density. Previous developmental studies have mostly used caps covering more restricted areas of the head surface, e.g. with unilateral or bilateral coverage of only the sides of the head [54,56,61]. While this approach may suffice for mapping changes in targeted brain regions responding to simple stimuli, it is blind to effects outside the a priori region of interest and is inadequate for mapping distributed responses to complex stimuli. Larger HD-DOT caps with posterior and lateral panels have been used to effectively map resting state networks, localized evoked responses to various visual and language tasks, and more dispersed

responses to naturalistic multi-sensory stimuli in adults [53,64]. The current system builds on those wide-field imaging cap designs and adds a panel covering the dorsal surface of the head as well, providing more somatomotor cortex coverage. Moreover, the 11mm separation between adjacent sources and detectors is comparable to the 10-13mm spacing reported for other high-performing HD-DOT systems [53,56], and is significantly tighter than the >20mm separations commonly reported for traditional fNIRS caps [33,54]. This high density was chosen to enhance image quality based on direct, controlled comparisons of high- and low-density measurement arrays by previous studies [30,56,61].

Brighter laser lights were used in place of the LED sources used by many prior systems [53,56,61] to enable more sensitive sampling of brain parenchyma by longer-range measurements, while staying well within recommended safety limits and posing no discomfort to participants. Meanwhile, the 2-pass encoding pattern facilitated a wide dynamic range, important for collecting measurements across a range of source-detector separations to provide depth sensitivity and remove superficial signal contamination [58]. While the higher intensity illumination captured the longer-range source-detector pair measurements with high SNR, the lower intensity illumination ensured that short range measurements were also obtained without being saturated. The effectiveness of using laser sources with 2-pass encoding was evident in the quality of the light-fall off curves obtained with both optical phantoms and human participants, with measurements both unclipped and above the noise floor from 11mm to 50mm of source-detector separation, and with light levels clustering within two orders of magnitude at each separation and spanning 7 orders of magnitude above the noise floor altogether (Figures 3.1I and

3.3A). The clear pulse waveforms visible in measurement timetraces, the 1Hz peak in Fourier spectra, and the pulse band-limited SNR plots all further validated the high SNR of the system for capturing hemodynamic signals in particular. Task maps and hemoglobin time courses across a range of tasks provided a final validation of the SNR and image quality of the system across the various panels of the cap. The motor task maps and time courses exhibited more noise than the other maps (Figure 3.3), likely because the dorsal panel was the newest part of the cap and its design and usage will likely be optimized further over future iterations and applications of the system.

Our data processing pipelines were designed to be stringent in order to maximize the reliability of results. The fMRI literature over the past decade has shown numerous prior findings to be spurious as a product of motion artifact and has characterized best practices for data preprocessing, including global signal regression and censoring of high motion frames, to reduce recurrence of this issue [129–131]. Learning from this literature, we tracked head motion in our HD-DOT data by calculating GVTD, similar to the dvars metric commonly used for fMRI [109]. We then reduced motion artifact in our data through the combination of motion scrubbing (eliminating frames for which GVTD spiked above a statistically determined threshold) and superficial signal regression (as head motion should affect all channels including those sampling the superficial tissues). In addition, we directly compared image quality between generic and subject-specific head models. Previous studies have shown that atlas-derived head models can provide an effective surrogate in the absence of data required for subject-specific head modeling [81]. While this is indeed the case, we here quantified improvements in image quality from

subject-specific head models for our adult data. While our head modeling pipeline depended on having structural MRI data which was not available for the children we scanned, future studies could use photometric methods such as 3-dimensional photographic capture of participants' heads to transform atlas models into subject-specific ones for children as well.

3.4.3 Functional brain mapping with movies

The reproducibility of evoked responses between independent viewings of a movie were previously shown using other imaging modalities and more recently using HD-DOT in adults [35,37,64,132]. Here we first replicated this result using our new HD-DOT system and a library of animated children's movie clips. Consistent with prior studies, we identified high inter-run synchronization across broad regions of cortex spanning visual, auditory, and association areas [35,37]. Our ability to map reproducible and specific responses to complex stimuli across the field of view provides another layer of evidence for the high SNR and image quality of our new system.

Feature regressor analysis has also previously been used with neuroimaging data to map activity correlated with specific features contained in a movie [36,64]. Here, we not only observed similar responses to speech and faces as these prior studies, but also used our library of diverse movie clips to compare regressor maps across 20 different movies. There have only been limited previous studies comparing functional brain maps across different movies. The overall consistency of regressor maps with notable exceptions for just a few movie clips is similar to observations with resting state functional connectivity analysis across different movies in prior

fMRI research [28]. Firstly, the similarity of speech regressor maps from independent viewings of the same movies (along the diagonal in Figure 3.5A) as well as using completely different mixtures of movie clips (Figure 3.4C-D and off-diagonal results in Figure 3.5A) illustrates the overall reproducibility and generalizability of our results. The high similarity of speech regressor maps to maps from a block-design word hearing task (Figure 3.4E and the final row and column of the matrix in Figure 3.5A) provide an additional indicator of construct validity for movie-based speech regressor analysis. Altogether, the fact that we can use data from viewing various different movie clips to obtain similar, meaningful regressor maps highlights the flexibility of the movie-viewing paradigm for functional brain mapping and is of major practical significance. We could, for instance, use different movies geared to study different age groups or cater to individual children's movie preferences and maximize their compliance without compromising feature regressor maps.

A closer look at differences across runs and across clips does, however, point to some interesting sources of variance. For example, both reproducibility and the construct validity of speech regressor maps from a particular clip and participant were positively correlated with inter-run synchronization between the two viewings of that clip by the participant. It may appear intuitive that more consistent responses to a movie in general would be associated with more consistent feature maps as well, as the participant is then more likely attending to the same movie features driving their brain activity across viewings. However, these findings are nevertheless informative, and suggest that overall inter-run synchronization may be a useful criterion for inclusion or exclusion of movie-viewing data for further analysis. Moreover, while regressor

maps were highly consistent for most of the movies used, clips from the show Daniel Tiger appeared to yield poorer speech maps (Figure 3.5E) that were less similar to other speech regressor maps or word-hearing task maps (Figure 3.5A). One possible explanation is that clips of this show were also associated with lower inter-run synchrony in the adult data (Figure 3.5C-D), perhaps linked to subjective reports by adult participants that these clips were less engaging than the other movies. However, another possible source of variance is the audiovisual content and time courses of the clips. The speech regressors for Daniel Tiger exhibited less frequent and less deep modulation over the course of the clips than other movies (Figure 3.5F), which could also help explain the less robust speech response maps. Regressor modulation depth should hence be evaluated during stimulus selection for future studies performing feature regressor mapping.

The richness of movies as a stimulus can be an asset by engaging participants and allowing experimenters to study processing of multiple features in parallel. However, our parallel feature regressor analysis of speech and face responses sheds light on important associated considerations in the selection of stimuli, method of analysis, and interpretation of results. Our speech and face regressor maps initially had notable overlap, both featuring prominent correlations in bilateral superior temporal gyrus (Figure 3.6B), consistent with prior studies [36,64]. There were several possible explanations for this result. It could have represented overlapping neural processing pathways for two features that are commonly concurrently processed in social interactions – we often attend to people's faces when they are speaking [36,133]. Alternatively, the observed regressor maps could be signatures of distinct face and

speech processing areas that are simply located too close to one another to be distinguished given the point-spread function associated with the optics of the current system. For example, primary and secondary auditory cortex lie in close proximity to the superior temporal sulcus that has been implicated in face processing [134]. However, we could also be observing an artifact of the stimuli used – faces and speech are often concurrently emphasized in movies, and this covariation could confound both maps. Indeed, we found that the similarity between face and speech maps generated from a movie clip was strongly predicted by the correlation between the face and speech regressors for that clip (Figure 3.6C). Future studies using movies to investigate processing of specific features of interest such as speech and faces should aim to select clips for which these regressors are more orthogonal, and other potentially salient and covarying features should also be considered and controlled as much as possible. But even in the absence of linearly independent features, taking a multivariate GLM-based approach to regressor analysis can mitigate confounding of results. Incorporating both features of interest and additional confounds of concern as regressors in the GLM design matrix yields more distinct feature maps (Figure 3.6E), with feature map correlations more evenly distributed around zero and no longer strongly driven by regressor correlations (Figure 3.6F). As in our previous analyses, the speech maps from the GLM regressor analysis featured significant correlations in the superior temporal gyrus bilaterally but more prominent on the left, consistent with prior studies [64]. Meanwhile, face maps were right-lateralized with two more posteriorly located temporal hotspots than the speech maps. For comparison, prior fMRI studies have also described right hemispheric specialization of face processing, with face responsive areas including the posterior superior temporal sulcus,

fusiform face area, and lateral occipital face area [36,134,135]. Furthermore, the GLM-derived speech regressor map was comparable to that obtained when a participant was presented with movie audio alone (i.e. with no confounding visual features) and the GLM-derived face regressor map was comparable to that from presenting the visuals stripped of audio (Figure 3.6D). These results provide compelling evidence for the benefits of the GLM-based multivariate regression approach for feature mapping. Future studies will be needed to evaluate what comprises a truly sufficient set of regressors for this multivariate analysis, to address not only features of interest but also more potential confounds. Indeed, covering "all" confounds may be an impossible goal for stimuli as complex as movies with innumerable covarying features, from basic audiovisual attributes like color, optic flow, volume, and frequency content to higher order details such as different categories of objects or different parts of speech. Also, although we used a canonical adult hemodynamic response function (HRF) for all our GLM analyses, studies showing potential differences in neurovascular coupling and the HRF in early development suggest that age-specific HRFs may need to be computed for optimal analysis in future developmental studies [136].

3.4.4 Imaging in awake young children

Finally we used the imaging system, stimuli, and analysis pipelines that we developed and characterized in adults to study a sample of 1-6 year-old children.

Of the 39 children enrolled in the study, 35 cooperated with the cap fit procedure and tolerated at least 5 minutes of imaging. However, as we sought to include in the final analyses only

participants who had provided high quality data (as indicated by light levels, pulse SNR, and the GVTD measure of head motion) for at least two matched and two unmatched movie-viewing runs and/or a word-hearing task run, we finally had usable data from 16 children (23-70mo). Participant attrition is to be expected in pediatric populations, but our retention rates were manageable and could further be offset by increased recruitment in future studies. Furthermore, data retention could likely be further increased through additional improvements in cap comfort (e.g. adding padding to parts of the cap distribute pressure from the tightest points of contact) and through the incorporation of a training and habituation phase prior to imaging sessions. Pediatric fMRI studies have reported success with measures such as mock scan sessions and audio tapes to habituate children to scanner noise [22,126]. We could similarly provide families with mock DOT caps for children to wear and become accustomed to before scans, as well as run training sessions to acclimate children more with the environment and tasks prior to imaging.

It was apparent to the experimenters running imaging sessions that free viewing of movies was a more engaging task for the children that increased their compliance relative to the word-hearing task requiring central fixation, but eye tracking provided further objective behavioral evidence for this claim (Figure 3.7B). There is a wealth of literature using eye tracking during various tasks as a tool to understand typical and atypical brain function and behavior [137,138]. Future studies could further factor gaze position into visual feature regressor analysis, evaluate how brain activity measured with HD-DOT and eye tracking data relate in awake young children during movie viewing or social interactions, and assess differences in populations such as children with autism.

As in adults, we were able to map high inter-run synchronization between independent viewings of matched but not mismatched movie clips, showing that reproducible and specific responses to complex, multi-sensory movie stimuli can already be found early in development (Figure 3.7 C-D). Numerous fMRI studies have shown the reproducibility of movie-evoked brain responses in older children and adults, but very few have evaluated this in children younger than 7 years [28,139,140]. Furthermore, through movie feature regressor analysis in the children, we were able to map responses to speech in the superior temporal gyrus that were present bilaterally but stronger in the left hemisphere, showing that this adult-like response pattern is also established early in development. This movie-derived speech regressor map looked comparable to the word-hearing task activation map from the children who complied with that task. Despite dramatic behavioral changes in language reception and production earlier in childhood, fMRI research on language development has generally studied children older than 4 years of age [20]. Limited prior work in younger children has studied passive speech perception during sleep in 2-3 year-olds [21], focused on infants in the first year of life [56,121], or used fNIRS arrays with low channel counts compromising image quality [127]. Here, we utilized the child-friendly HD-DOT imaging environment and movie paradigm to image active perception of naturalistic speech in awake 1-6 year-old children.

3.4.5 Future directions

Our pilot sample grouped together children across a 5-year-wide age range in order to establish imaging feasibility, but there is much variability in behavior and likely in brain function across

this broad time span. Future studies with more participants at specific ages could further investigate longitudinal changes across early development, with a particular focus on 1-4 year-olds who have been especially challenging to study with fMRI. We coded regressors to study processing of speech and faces, but other regressors could be coded for the same stimuli to study other behaviors of interest (e.g. basic visual processing, emotional responses, etc) using the same data. Future work could also explore other movie-based analyses such as inter-subject synchrony, functional connectivity, and inter-subject functional connectivity to compare other aspects of functional brain organization between children at different ages and adults [28,35,141]. In addition, our effective mapping of reproducible and specific responses to movies and movie features like speech and faces suggests that these sorts of complex, naturalistic stimuli could also be accurately decoded from brain activity sampled by HD-DOT. However, decoding algorithms will need to be designed and their performance will have to be explicitly evaluated in order to build on recent work decoding simpler visual information from HD-DOT data [119]. Furthermore, while movie-viewing may be a powerful, child-friendly paradigm for studying various perceptual processes, other experimental paradigms should also be combined with HD-DOT imaging in order to investigate the developmental neurobiology of speech production, parent-child interactions, and various other fascinating behaviors.

Finally, these approaches could all be applied to not only characterize typical developmental trajectories but also better understand atypical development and plasticity. For instance, the early emergence of atypical speech and face processing could be characterized in children diagnosed with autism spectrum disorder as well as pre-symptomatic at-risk siblings [142]. The plasticity of

speech and motor pathways could be imaged following perinatal brain injury or other causes of childhood movement disorders like cerebral palsy [143]. Children with cerebral palsy could also benefit from augmentative communication based on decoding of speech from brain activity measured with wearable HD-DOT systems. Effects of malnutrition on brain function could be studied during the first 1000 days of life emphasized to be critical for brain development [144], with low-cost, portable imaging systems already shown to be feasible for use in low-resource settings [61]. And Hubel and Wiesel's foundational research on critical periods in the feline visual system could be eloquently translated into clinical studies of visual processing and plasticity in children with developmental strabismus undergoing corrective surgery [14,97]. Further essential optimization of HD-DOT imaging is ongoing – caps are being made wireless and more wearable [107], while algorithms to improve data quality through measures like motion processing and head modeling are being continually refined [109]. However, we believe that our work here, developing imaging tools and establishing the feasibility of mapping brain function in awake young children with HD-DOT, lays an important foundation for future advances in developmental neuroscience with major potential clinical impact.

“Can you read my mind?”

- The Killers

Chapter 4: Decoding naturalistic audiovisual information from high-density diffuse optical tomography neuroimaging data

4.1 Introduction

Decoding of naturalistic information from functional brain imaging data acquired in natural settings could have exciting neuroscientific and clinical implications. Studies using recording methods such as functional magnetic resonance imaging (fMRI) and electrocorticography (ECoG) have accurately reconstructed detailed visual and auditory information from the cortical activity of their research participants. For instance, visual cortex activity sampled with fMRI has been used to reconstruct images and movies, providing insight into how these types of information are encoded in the brain and could potentially be extracted to recreate memories, dreams, and real or imagined scenes [38–41,68]. More recently, intelligible speech has been synthesized from ECoG data [42,43], while accurate and efficient writing has been achieved be decoding intracortical microelectrode recordings [145], with promise for patients who are otherwise unable to communicate due to severe speech and motor disorders. However,

widespread application of these findings is limited by the recording methods used. The logistics of MRI scanners are not conducive to purposes such as daily communication, while ECoG and intracortical microelectrodes require invasive neurosurgery to be implanted. Decoding has also been investigated in a range of populations with less invasive and more portable technologies such as functional near-infrared spectroscopy (fNIRS), but low spatial resolution and image quality have limited the intricacy and accuracy of classifications possible [30,46–51]. Ideally, we could accurately decode detailed naturalistic information using brain imaging methods that can be employed in natural settings.

High-density diffuse optical tomography (HD-DOT) is an emerging method that uses dense optode arrays to combine logistical advantages of optical neuroimaging, such as portability and an open imaging environment, with enhanced image quality closer to that of fMRI [53,60,61,114]. Recent work has shown that simple visual information such as the location of a checkerboard stimulus can be reliably decoded from HD-DOT data [119], capitalizing on the accurate retinotopic mapping previously achieved with the imaging modality [52,60,69]. However, decoding of stimuli more relevant to real-word contexts and decoding of information beyond the visual domain are yet to be investigated with HD-DOT. Studies have shown that HD-DOT can be used to map distributed patterns of brain activity in response to audiovisual movies with high image quality across a wide field of view [64]. Therefore, we hypothesized that these complex, naturalistic, multisensory stimuli could also be decoded from HD-DOT data.

We sought to test this hypothesis using a previously acquired data set from adults viewing a library of movie clips twice each over the course of multiple imaging sessions. We took a

template matching approach to decoding, and examined how parameters such as the number of templates and duration of trials affected performance. We also evaluated decoding of audiovisual stimuli from data confined to separate auditory and visual regions of interest (ROIs), as well as decoding of purely auditory and purely visual versions of the stimuli. Our work establishes the feasibility of decoding complex auditory and visual stimulus information from optical neuroimaging data, and encourages further studies designed for attempting more elaborate decoding with HD-DOT towards goals like reconstructing visual scenes and natural language.

4.2 Methods

4.2.1 Data set

This study utilized data from a previous project that had imaged a small group of highly sampled adults watching a library of animated movie clips. The goal of the original study was to develop and validate a new HD-DOT system, a children's movie viewing paradigm, and associated data analysis pipelines for imaging young children. We repurposed the movie viewing data collected in adults to evaluate the feasibility of decoding naturalistic audiovisual stimuli from HD-DOT data. While further details of the data collection and preprocessing methods can be found in the original report, we summarize here the highlights that we consider most relevant to the current study.

Imaging system: We sought to evaluate decoding performance using optical neuroimaging data with high image quality. Therefore, we utilized data from participants imaged using a HD-DOT

system with 128 laser sources and 125 detectors interleaved in a grid with 11mm spacing, collectively yielding up to 5202 measurements with <50mm source-detector separation, covering posterior, lateral, and dorsal surfaces of the head. The field of view extended ~1cm below the pial surface including regions of occipital, temporal, parietal, and frontal cortex.

Participants: Decoding studies commonly focus on 2-3 participants in order to allow collecting larger amounts of data per participant and reduce challenges associated with inter-individual variability in structural and functional neuroanatomy, compliance, and data quality [38,39]. Therefore, the audiovisual movie viewing data set taken from 3 adults (age 20-31 years) who participated in 5 imaging sessions each was well suited to our study. Additional data was available from 1 of the participants who listened to audio clips without visuals and watched silent movies stripped of audio. Informed consent for participation as well as storage and reuse of data was obtained from all participants in accordance with the IRB protocol approved by the Human Research Protection Office at Washington University School of Medicine.

Task: We aimed to examine factors affecting decoding of complex, naturalistic, auditory and visual stimuli. Our dataset was from adults who had watched 20 5-6-minute-long audiovisual animated movie clips twice each, listened to 4 5-6-minute audio clips twice each, and/or viewed 4 5-6-minute silent visual clips twice each. Incomplete imaging sessions, in which participants did not complete all 8 intended runs, were excluded, as the full set of runs was required for the minimum experiment of 4-way decoding by template matching with 2 independent training and test runs. The remaining data were from 13 audiovisual movie viewing sessions across 3 participants, and 2 audio only sessions and 1 visuals only session in 1 of the participants.

Data preprocessing: Pipelines developed by prior studies were used to preprocess data so as to optimize data quality and prevent artifacts from confounding the decoding analysis. In brief, data cleaning steps included omission of noisy channels with >7.5% variance over the course of a run, high-pass filtering to reduce long-term drift, superficial signal regression to remove superficial tissue and global signals including pulse and motion artifacts, and low-pass filtering to remove residual pulse and high-frequency noise [53,58]. Data were then downsampled to 1Hz. Additional motion artifact removal occurred post-processing, by calculating the global variance of temporal derivatives (GVTD) metric of head motion and censoring all frames for which GVTD rose above a statistically derived threshold [109].

Data reconstruction: Subject-specific head models, optimized using each participant's structural MR images and control task HD-DOT data, were utilized to accurately model light propagation through the head. The resulting sensitivity matrices were inverted using Tikhonov regularization with λ_1=0.05 and spatially variant regularization with λ_2 = 0.1, and then used to reconstruct data in voxel space. Flat field reconstructions of the sensitivity matrices were thresholded at 10% of their maxima to obtain conservative subject-specific estimates of the field of view, which were later applied as spatial masks during template matching calculations for decoding. Changes in oxy- and deoxy- hemoglobin were finally obtained through spectral decomposition [53].

4.2.2 Decoding by template matching

For our first assessment of decoding naturalistic stimuli from HD-DOT data, we wanted to start with a simple classification algorithm and hence tested a template matching approach modified

from a previous visual decoding study [119]. Each imaging session was analyzed separately to avoid challenges associated with cross-registration of HD-DOT data across sessions. The data from each session was split into two sets of runs – training runs and test runs – with one viewing of each movie clip in each half. Spatiotemporal templates were constructed from the training runs and decoding trials were extracted from the test runs as per the time points detailed in Table 4.1. These timings were selected to keep template and trial durations constant across the clip segments within an experiment (except when they were intentionally varied), discard transient stimulus onset responses at the beginning of each run, and allow intervening periods between trials for the sluggish hemodynamic response to the previous trial to fade.

Table 4.1: Timing of templates and trials taken from each of 4 movie clips

Experiment		**Template and trial times for each clip: start - end (note: clips began at 0s and total clip lengths ranged from 299 to 375s)**
Number of templates	**Trial duration**	
4	4 min	30-270
8	2 min	25-145, 175-295
16	45 s	25-70, 100-145, 175-220, 250-295
32	15 s	25-40, 60-75, 95-110, 130-145, 165-180, 200-215, 235-250, 270-285
4	Varied	start at 25; end from 40 to 295 in increments of 15s
Varied	45s	25-70, 100-145, 175-220, 250-295

The cortical oxyhemoglobin signal response (voxels x time) measured by HD-DOT during each trial in the test data was compared to the spatiotemporal (voxels x time) templates for each movie from the training data as follows. Voxel-wise Pearson correlation coefficients $r_{n,m,v}$ were calculated between zero-mean-centered signal time courses for the v^{th} voxel in the n^{th} template $T_{n,v}$ and in the m^{th} trial $S_{m,v}$ as:

$$(1)\ r_{n,m,v} = \frac{(T_{n,v})\cdot(S_{m,v})}{|T_{n,v}||S_{m,v}|}$$

These were averaged across all F voxels within the subject-specific field of view to determine the mean template correlation $r_{n,m}$ for the m^{th} trial with the n^{th} template:

$$(2)\ r_{n,m} = \frac{\sum_{v=1}^{F} r_{n,m,v}}{F}$$

The decoding output D_m for the m^{th} trial was finally determined by the template number n that had the maximum correlation with the trial response:

$$(3)\ D_m = \underset{n}{\operatorname{argmax}}(r_{n,m})$$

Confusion matrices were generated by tracking the number of times that each trial clip was decoded as each of the possible options across all imaging sessions included in the analysis. Mean decoding accuracy was reported as the total percentage of trials across all sessions that were decoded correctly. Statistical significance was evaluated by comparing the distribution of session-wise decoding accuracies to chance, i.e. 1/(total number of templates), using a two-tailed Wilcoxon signed rank test.

4.2.3 Construction of auditory and visual ROIs

For ROI-confined analyses of decoding performance, a combination of pre-acquired control task data and anatomical priors were used to define ROIs.

An auditory ROI was defined using data collected using the same HD-DOT system on 3 adults collectively performing 22 runs of a block-design word-hearing task across 14 imaging sessions. During the task, participants listened passively to spoken word lists presented at 1 word per second for 15 seconds per block, with 15 seconds of silence in between blocks, and a total of 6 blocks per run [53]. Activations from this task were thresholded at 30% of their maximum value, and all remaining contiguous voxels surrounding the peak activation in both temporal lobes were assigned to the auditory ROI. Any voxels lying outside the field of view for the cap on any of the participants (i.e. below the 10% sensitivity threshold for any of the subject-specific head models) were excluded from the final ROI. This ROI was centered around the superior temporal gyrus bilaterally (Figure 4.4A), as anticipated for an auditory ROI.

A visual ROI was defined with reference to data collected using the same HD-DOT system on 4 adults collectively performing 6 runs of a visual stimulation task, in which a rotating checkerboard wedge swept out the participants' visual fields. Specifically, the flickering wedge subtended a polar angle of 60°, a radial angle from 2.5° to 10°, and rotated 10° at a time through 36 positions spanning 360° at 1 second per position, for a total of 10 cycles per run [119]. The voxel with the maximum recorded signal across all the runs was used as the center of a unilateral visual ROI. A 10mm kernel was grown around this center, and then mirrored on the contralateral

hemisphere to create a bilateral visual ROI. Voxels lying outside the field of view for the cap on any of the participants were excluded from the final ROI. This ROI spanned much of the bilateral occipital cortex (Figure 4.4B), as expected for a visual ROI.

4.3 Results

4.3.1 Feasibility of decoding movie viewing HD-DOT data by template matching

Effective decoding of stimulus information from neuroimaging data depends on the measured brain responses being specific to each stimulus condition and reproducible across the instances of each condition. The specificity and reproducibility of movie-evoked responses as captured by HD-DOT was assessed using data from imaging sessions in which participants viewed four different movie clips twice each. The data from each session was split into two halves: the training half contained the first presentation of each movie, and the test half contained the second. Voxel-wise correlations between oxyhemoglobin signal time courses were computed for every possible pairing of one training run and one test run (Figure 4.1A). Strong positive correlations were consistently seen along the main diagonal of the correlation matrix but not in the off-diagonal maps, i.e. for runs in which participants were presented with the same movie clip both times but not for mismatched movie clips. This difference between matched and mismatched movie responses was further evident on averaging correlations across all matched movie runs and all mismatched movie run pairs within the session (Figure 4.1B). The

reproducible, movie-specific responses mapped by HD-DOT implied that the movie presented during each test run could be decoded from the participant's HD-DOT by comparison to template responses from the training data (Figure 4.1C). The performance of this template matching approach to movie decoding from HD-DOT data was evaluated across 13 imaging sessions, involving 3 participants viewing 5 different sets of movies (Figure 4.1D). The bright main diagonal of the confusion matrix implies that the decoded clip typically matched the clip that was actually presented. Decoding accuracy was calculated to be 92.3±4.4% (mean ± standard error of mean), which is significantly greater than chance levels of 25% ($p = 0.0002$, Wilcoxon signed rank test).

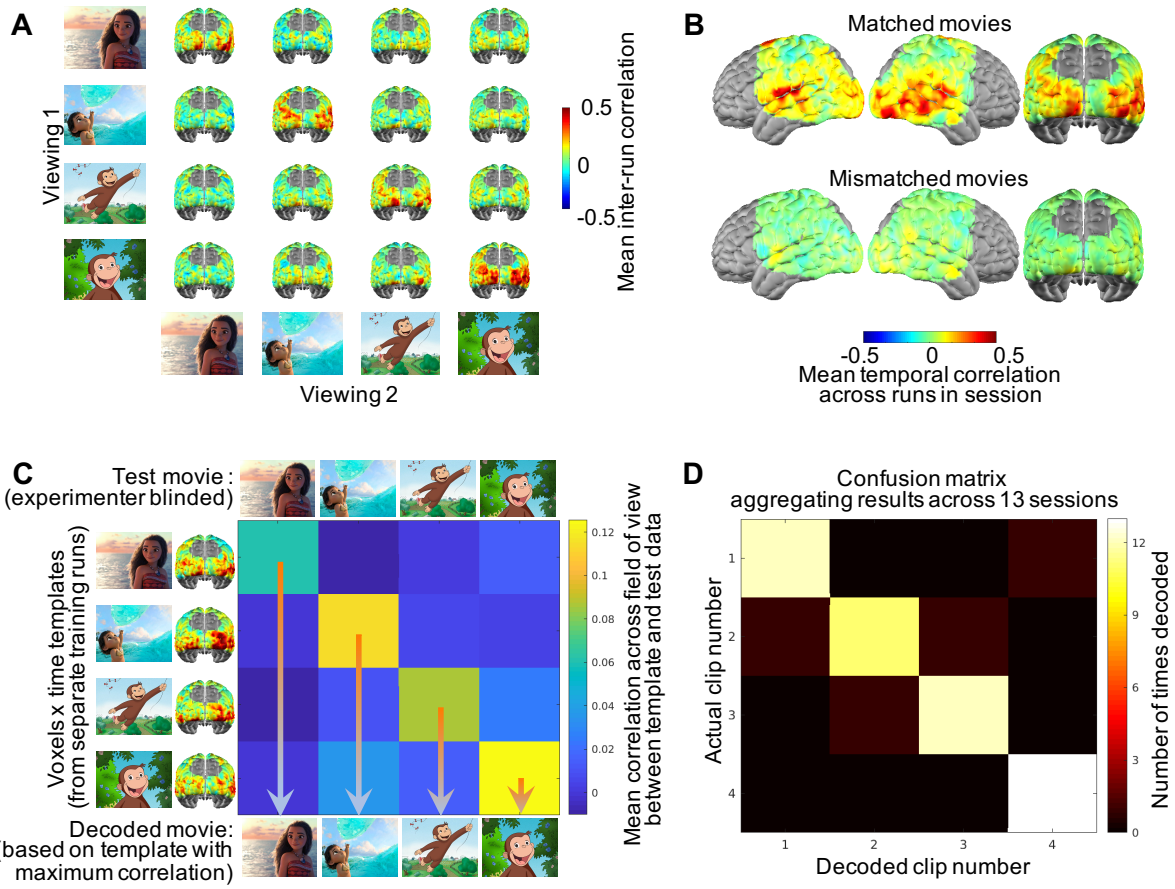

Figure 4.1: Decoding audiovisual movie identity from HD-DOT data by template matching
Participants were shown four different audiovisual movie clips twice each over the course of an imaging session. Comparing cortical responses imaged with HD-DOT between independent viewings of every possible pairing of clips reveals strong correlations between runs in which the participant was presented with the same movie clip. (B) Averaging inter-run correlation maps across all possible pairings of matched and mismatched movies presented during a session illustrates that movies evoke reproducible but clip-specific patterns of brain activity across visual, auditory, and other areas of cortex. (C) A template-matching strategy can hence be taken to decode which of a set of movies a participant was viewing from their HD-DOT data. One presentation of each movie was used as training data to construct templates of the expected brain response. The movies presented during each of the remaining runs were then decoded based on which of the templates had the highest spatiotemporal correlation with the test data. (D) Decoding results were aggregated across 13 sessions involving 5 different sets of movies and 3 participants. The bright main diagonal of the confusion matrix illustrates that the decoded clip matched the presented test clip with an accuracy of 92.3±4.4% (mean ± standard error of mean).

4.3.2 Factors affecting decoding performance

As the initial analysis aggregated observations across multiple sets of movies and participants, decoding performance was also evaluated separately for each set of movies (Figure 4.2 A-D) and for each participant (Figure 4.2 E-G). While the bright main diagonal indicated effective decoding in all cases, more errors (reflected by off-diagonal elements of the confusion matrices) were observed for some sessions and participants than others. Differences in raw data quality were considered as one possible contributor to the observed variance in decoding performance. Since decoding depends on the reproducibility and specificity of movie-evoked responses measured by HD-DOT, these attributes were quantified by a response repeatability score as the difference in mean inter-run correlations between matched and mismatched pairs of movie viewing runs. Meanwhile, as a measure of data quality for each matched run pair, a pulse signal-to-noise score was calculated as the maximum ratio of signal power in the 0.5-2Hz band to bandwidth-scaled median power in the flanking frequency range across the cap. A significant positive correlation was observed between response repeatability and pulse signal-to-noise (Figure 4.2H).

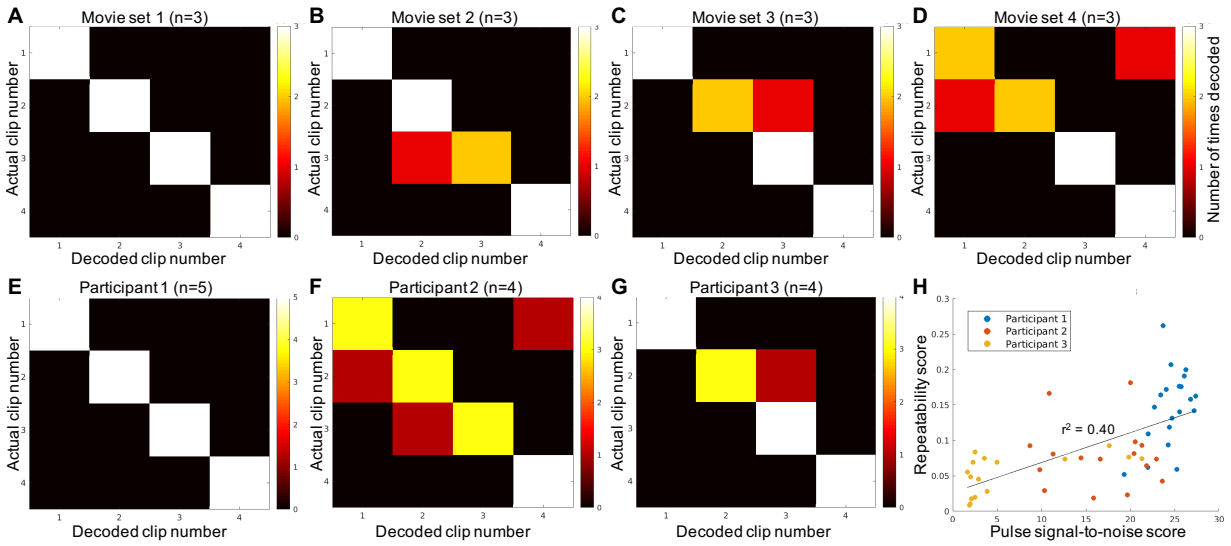

Figure 4.2: Breakdown of decoding results across different sets of movies and participants
(A-D) Confusion matrices showing decoding results for 4 different sets of movie clips that were each presented to 3 participants. (E-G) Confusion matrices showing decoding results for 3 participants who each participated in 4-5 imaging sessions. (H) The repeatability of evoked responses is correlated with measures of data quality. Specifically, the maximum pulse band signal-to-noise ratio during a run is a strong predictor of the difference in mean correlations between responses to matched and mismatched clips, which is the basis of decoding by template matching.

While the initial decoding results were encouraging, they involved only 4 choices, each with a relatively generous 4 minutes of template and test data. In order to challenge the decoding algorithm, the duration of each template and test trial was varied systematically in 15 second decrements from 4.5 minutes down to 15 seconds in length. Decoding accuracy was reevaluated across participants and sessions for each clip duration (Figure 4.3A). Decoding accuracy remained significantly above chance in all cases, although it varied between participants and improved with increasing clip duration. Another way to challenge the decoding algorithm was to present an increasing number of possible classification options. In order to implement this experiment, each movie clip was split into four shorter segments, and decoding accuracy was reevaluated upon varying the total number of templates and hence decoding choices from 4 to 16 in increments of 4, while holding trial duration constant at 45 seconds (Figure 4.3B). As expected, accuracy decreased with an increasing number of options, but performance remained well above chance even for the 16-way classification.

Both the trial duration and the number of templates were then varied in concert, and confusion matrices computed for 8-way decoding with 2-minute-long clips (Figure 4.3C), 16-way decoding with 45-second-long clips (Figure 4.3D), and 32-way decoding with 15-second-long clips (Figure 4.3E). The more challenging decoding was associated with more decoding errors, but accuracy remained significantly above chance ($p<0.05$, Wilcoxon signed rank test) in all cases (Table 4.2).

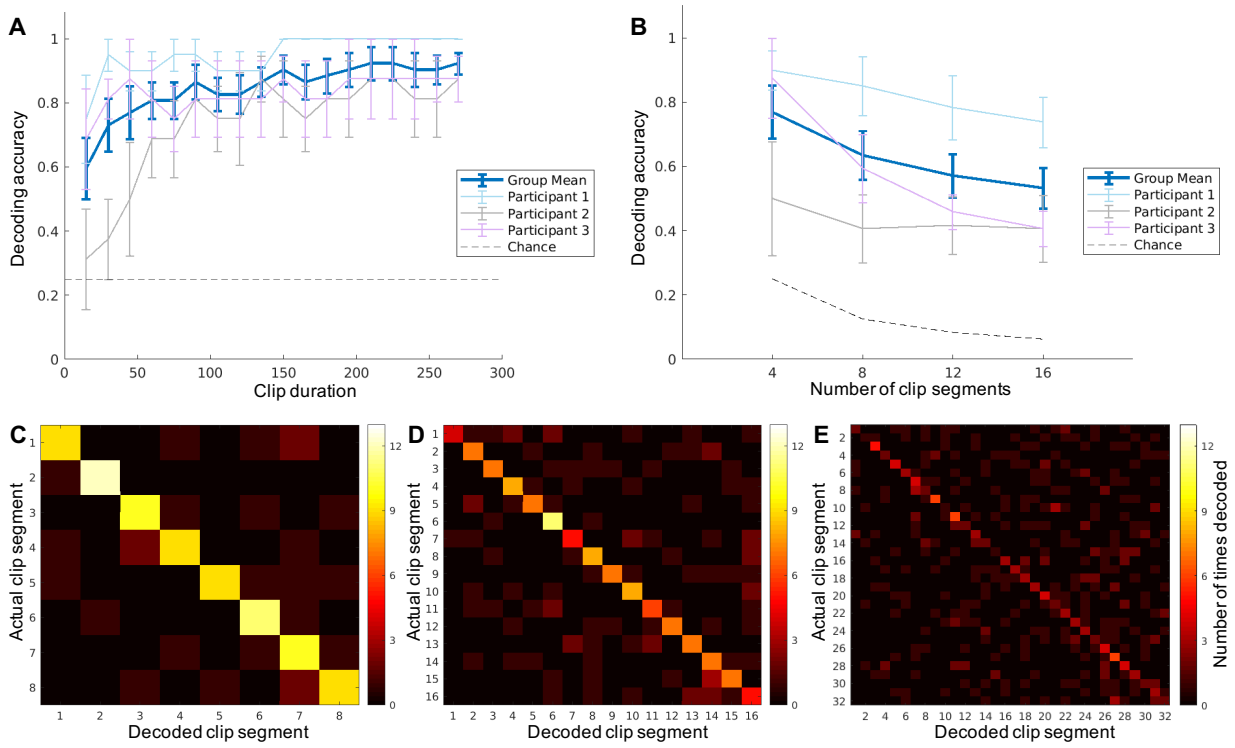

Figure 4.3: Factors affecting decoding performance

(A) The duration of data used for both template construction and testing was varied systematically and decoding accuracy was calculated across all sessions in each case. Decoding improved with increased duration, but was already significantly above chance with as little as 15 seconds worth of data, and plateaued by 150 seconds. (B) Using a fixed clip duration of 45 seconds, the number of templates and decoding choices was varied between 4 and 16, and accuracy was calculated in each case. Decoding accuracy decreased with an increasing number of targets, but remained well above chance even for 16-way decoding. (C) Confusion matrix for decoding 8 clip segments, each 2 minutes long (accuracy = 76.0±5.0%, relative to chance = 12.5%). (D) Confusion matrix for decoding 16 clip segments, each 45 seconds long (accuracy = 53.4±6.3%, relative to chance = 6.25%). (E) Confusion matrix for decoding 64 clip segments, each 15 seconds long (accuracy = 23.3±4.2%, relative to chance = 3.13%). Although accuracy (indicated by the brightness of the main diagonal) decreased as expected by making the decoding more challenging with an increasing number of targets and decreasing amounts of template and test data, mean accuracy remained significantly greater than chance in all these cases ($p<0.05$, Wilcoxon signed rank test).

Table 4.2: Effect of varying both number of choices and trial length on movie decoding accuracy

Number of choices	Trial length	Chance decoding accuracy	Observed decoding accuracy (** significantly above chance)	Figure
4	4 minutes	25%	92.3±4.4% **	1D
8	2 minutes	12.5%	76.0±5.0% **	3C
16	45 seconds	6.3%	53.4±6.3% **	3D
32	15 seconds	3.1%	23.3±4.2% **	3E

4.3.3 Decoding within auditory and visual regions of interest

It was observed that signal correlations for matched movie viewing runs were strong in both temporal cortex and occipital cortex (Figure 4.1A-C), where responses to auditory stimuli and visual stimuli, respectively, are typically localized and have been previously mapped with HD-DOT [53]. To evaluate whether effective decoding of complex naturalistic stimuli might be feasible based on either auditory or visual responses alone, the template matching analysis was repeated using template and test data confined to auditory and visual regions of interest (ROI), defined based on block-design task data and anatomy. The auditory ROI was based on activations from a block-design word-hearing task, while the visual ROI was defined using data from a visual stimulation task in which rotating checkerboard wedges swept around the participants' visual fields. Mapping correlations across matched movie runs and mismatched run pairs from every imaging session revealed strong positive correlations within both of the ROIs (Figure 4.4 A-B). Template matching within each of these ROIs yielded decoding performance comparable to those obtained across the full field of view in both simpler and more complex versions of the experiment, e.g. with >85% mean decoding accuracy for 4-way decoding of 4-minute-long clips (Figure 4.4 C-D) and with >50% mean decoding accuracy for 16-way decoding of 45-second-long clip segments (Figure 4.4 E-F).

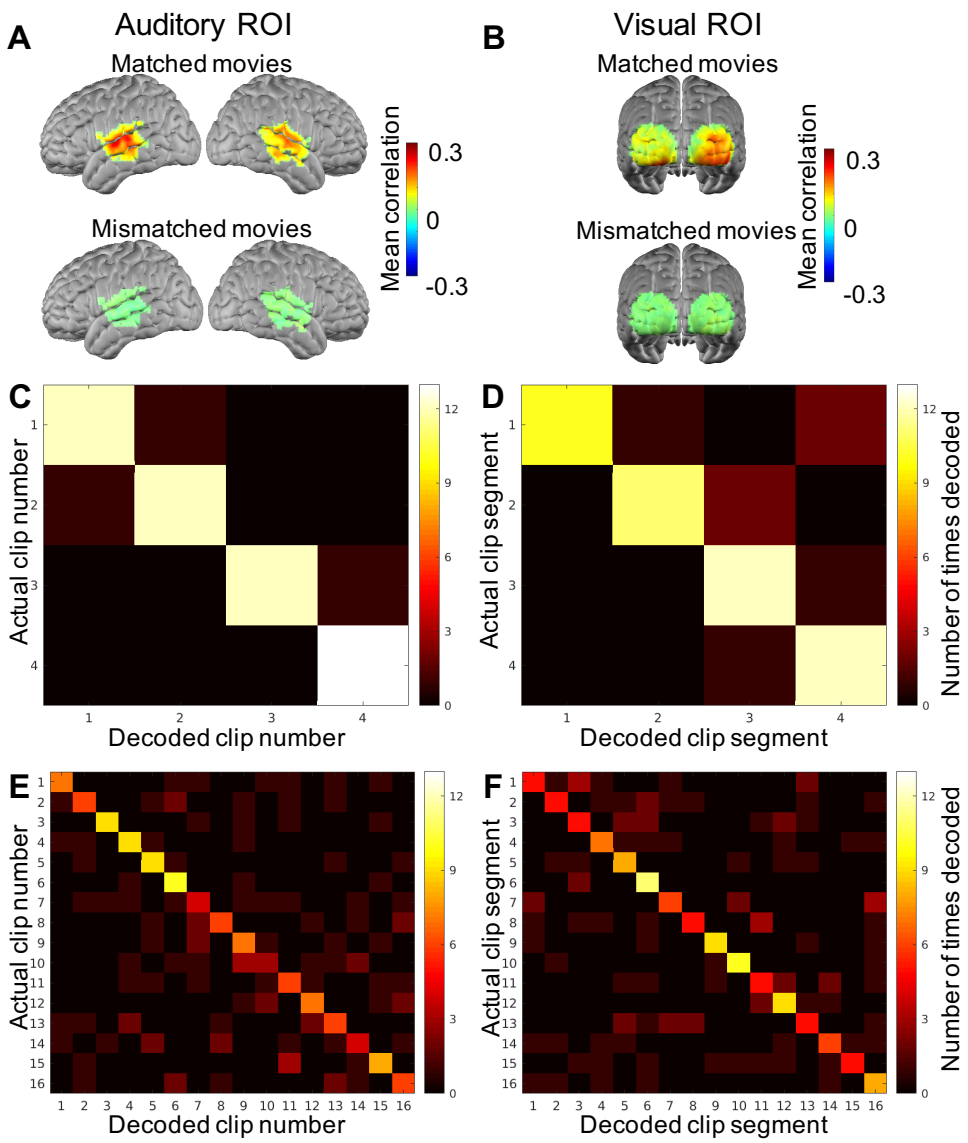

Figure 4.4: Decoding within auditory and visual regions of interest (ROIs)
(A) Mapping inter-run correlations across all sessions shows high correlations between responses to matched and not mismatched clips within a pre-defined, task-based, auditory ROI. (B) Mapping inter-run correlations across all sessions shows high correlations between responses to matched and not mismatched clips within a pre-defined, task- and anatomy- based, visual ROI. (C) Confusion matrix showing the effectiveness of decoding which of four 4-minute-long movie segments was presented to participants based on data from within the auditory ROI alone (accuracy = 94.2±3.0%, relative to chance = 25%). (D) Confusion matrix showing the effectiveness of decoding which of four 4-minute-long movie segments was presented to participants based on data from within the visual ROI alone (accuracy = 86.5±6.7%, relative to

chance = 25%). (E) Confusion matrix showing that decoding using only data from within the auditory ROI is also effective in the more complex case with 16 choices and only 45 seconds of data per template and trial (accuracy = 51.4±5.5%, relative to chance = 6.25%). (F) Confusion matrix showing that decoding using only data from within the visual ROI is also effective in the more complex case with 16 choices and only 45 seconds of data per template and trial (accuracy = 52.4±6.3%, relative to chance = 6.25%). In all cases, decoding accuracy was significantly above chance ($p<0.05$, Wilcoxon signed rank test).

4.3.4 Decoding of purely auditory and visual stimuli

Based on the results of decoding within auditory and visual ROIs, it was hypothesized that responses to complex, naturalistic stimuli that were purely auditory or purely visual could also be decoded from HD-DOT data. A smaller amount of data was available to test this hypothesis. In 2 imaging sessions in 1 participant, four movie audio clips had been presented sans any of the visuals twice each. Results are illustrated for one of these sessions, but the observed decoding accuracy was 100% in both cases. These data sets were each split into training and test halves, and voxel-wise correlations between oxyhemoglobin response time courses were mapped for every possible pairing of training and test runs. Strong positive correlations were observed bilaterally in the superior temporal gyrus for matched but not mismatched movie viewings (Figure 4.5A). A template matching approach was used to decode which audio clip participants were hearing in the test data, using 4-minute-long templates and trials confined to the pre-defined auditory ROI (Figure 4.5B). In 1 imaging session, a full set of four movies had been presented twice each with the visuals stripped of any audio. Here, strong positive correlations were observed in the occipital cortex for matched but not mismatched movie viewings (Figure 4.5C). Although only one complete session's worth of data was available, decoding which silent movie has been presented was 100% accurate using templates and test data confined to the visual ROI (Figure 4.5D).

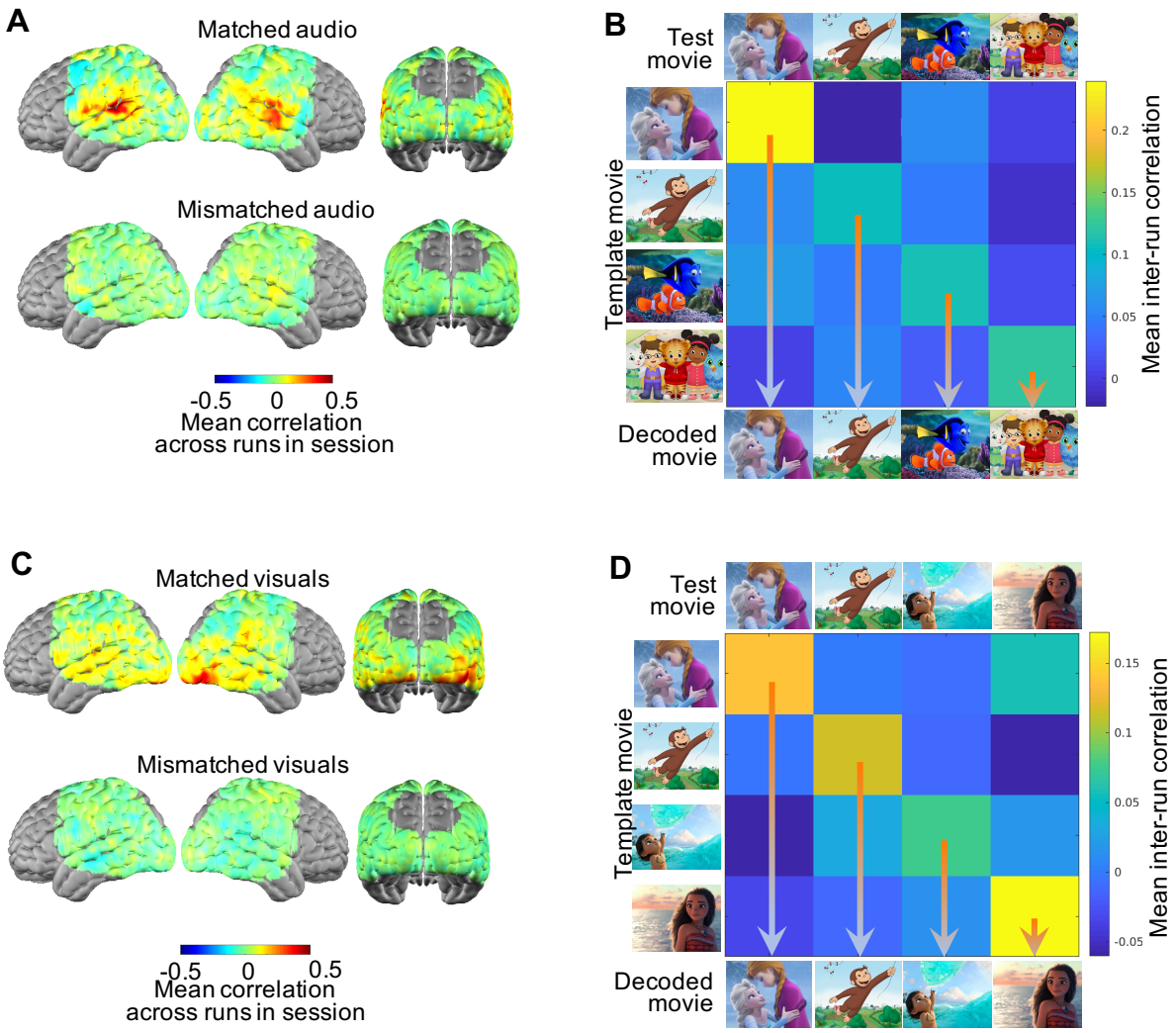

Figure 4.5: Decoding purely auditory and purely visual naturalistic stimuli from HD-DOT data (A) Participants were presented with four different audio clips without any visuals twice each over the course of an imaging session. Comparing cortical responses imaged with HD-DOT between independent presentations of every possible pairing of audio clips reveals strong correlations in the superior temporal gyrus for runs in which the participant was presented with matched clips but not mismatched clips. (B) Decoding of audio clip identity by spatiotemporal template matching within a pre-defined auditory region of interest. (C) Participants were presented with four different silent movie clips twice each over the course of an imaging session. Comparing cortical responses imaged with HD-DOT between independent presentations of every possible pairing of clips reveals strong correlations in occipital cortex between runs in which the participant was presented with the same clip but not mismatched clips. (D) Decoding of silent movies by spatiotemporal template matching within a pre-defined visual region of interest.

4.4 Discussion

In summary, we used a previously acquired set of HD-DOT data from young healthy adults watching audiovisual movie clips to evaluate the feasibility and performance of decoding naturalistic auditory and visual stimuli from optical neuroimaging data. Based on the reproducibility and specificity of movie-evoked brain activity captured by HD-DOT imaging, we used a template matching strategy to decode which of 4 movies participants had viewed from 4-minute-long trials with 92.3±4.4% accuracy. Decoding accuracy remained significantly above chance as trial duration was systematically reduced to 15 seconds and as the number of template and trial types was increased up to 32. Mean decoding accuracy also remained above 85% for 4-way decoding and above 50% for 16-way decoding when using template and test data confined to auditory or visual ROIs. In a smaller number of available data sets, decoding was also effective (100% accurate across 3 sessions altogether) for purely auditory and purely visual stimuli. We therein establish that naturalistic auditory and visual stimulus information can be effectively decoded from HD-DOT data.

4.4.1 Study design in relation to prior research

Previous studies using other imaging modalities have accomplished impressively detailed and accurate decoding of visual, auditory, and other information from recordings of brain activity. Of note, fMRI recordings of visual cortex activity have been used to reconstruct naturalistic images, dynamic silent visual scenes, and semantic content from movies [38–40,74]. In addition, electrocorticographic recordings from brain areas associated with language and motor function

have been used to synthesize intelligible speech in epilepsy patients [42,43]. However, translation of this research into other clinical populations and more widespread applications is currently constrained by the limitations of the recording methods used. For example, neural decoding could provide a means of augmented communication to patients with severe speech and motor disorders, or perhaps even an alternative mode of interfacing for all humans. However, the size, technical requirements, cost, and other logistics of MRI scanners preclude their use for applications such as longitudinal communication in day-to-day life. Meanwhile, ECoG requires invasive neurosurgery, which imposes risks for patients and should certainly be avoided for the masses. Other noninvasive, portable, and wearable imaging technologies, like fNIRS, have also been used in decoding research to access populations that are not as easily imaged with fMRI, such as infants and bed-bound patients [46,48]. However, image quality has generally been a limiting factor for fNIRS, as low spatial resolution and artifacts constrain the granularity and accuracy with which brain activity can be mapped and decoded. Thus, most prior fNIRS decoding studies have been constrained to pairwise or four-way classifications [46–51]. HD-DOT uses high-density optical arrays to combine logistical advantages of fNIRS with improved image quality [30,56,61], and has been used to map localized and distributed patterns of brain activity in various populations with results comparable to fMRI [53,60,64,75]. Furthermore, HD-DOT is being constantly developed to further enhance image quality, progressively streamlined to simultaneously improve wearability [56,107,108], and increasingly used in a range of populations and applications. One recent study characterized the performance of retinotopic decoding from HD-DOT data using checkerboard stimuli with up to 36 possible

stimulus positions [119]. In other research, HD-DOT has been used to reliably map brain responses to movies and movie features during naturalistic viewing [64]. The current study aimed to build on all this prior work by using HD-DOT data collected during movie viewing to evaluate non-binary decoding of complex and naturalistic visual and auditory stimuli from optical neuroimaging data.

4.4.2 Feasibility of decoding movie identity by template matching

It has already been shown that naturalistic viewing of movies evokes reproducible patterns of brain activity that can be captured by several neuroimaging approaches including HD-DOT [35,64,132]. However, previous HD-DOT studies mapping brain activity during movie viewing have reported this effect across a large number of runs collected across multiple imaging sessions and participants. The current study both applies and further builds on those findings by evaluating decoding of individual movie presentations. The >90% classification accuracy observed using single runs to construct templates for decoding single trials indicates both the degree of replicability of responses to a given movie from run to run and the high discernibility of responses between individual viewings of different movies (Figure 4.1). While decoding from a limited menu of 4 choices using generously long 4-minute trials may appear coarse on the surface, this initial experiment provided a valuable starting point to establish the feasibility of decoding complex naturalistic stimuli from HD-DOT data and justified further investigation of factors influencing decoding performance.

4.4.3 Data quality and processing considerations

The observed relationship between response repeatability and pulse SNR (Figure 4.2H) is consistent with previous observations of associations between decoding performance and data quality [119], and underscores the importance of data quality control in decoding and neuroimaging research in general. Even given participant compliance, a smooth cap fit, and high raw light levels based on real-time data quality visualizations during data collection, SNR variation in the data can result from variability in other factors such as optode-to-scalp coupling or participant physiology. If the cardiac pulse is not strongly detected above background noise, it is unlikely that more subtle variations in hemodynamic signals associated with localized or distributed brain activity will be reliably captured. However, pulse SNR must be only one of several factors contributing to variation in decoding performance, given that data from participant 2 usually had higher SNR but also produced more decoding errors than data from participant 3 (Figure 4.2 F-H and Figure 4.3 A-B). Other possible sources of variance include differences between runs and/or individuals in other aspects of data quality, structural and functional neuroanatomy, and cognitive processing of movie content.

Another aspect of data quality that was carefully controlled in the current study was head motion. We wanted to ensure that our decoding was based on neural signals and not on nuisance signals such as recurring patterns of head motion. Indeed, provocative findings of several early fMRI studies of brain function were eventually overturned by methodological research ascribing them to be artifacts of head motion [130,131], and best practices for minimizing motion artifact in fMRI data have since been developed [129,146]. Learning from this literature as well as from

prior studies of motion artifact in HD-DOT data [109], we calculated the global variance of temporal derivatives (GVTD) across all our data, and censored all time points at which the GVTD rose above a statistically derived threshold. In addition, our preprocessing pipeline involved exclusion of channels with >7.5% temporal variance over a run as well as superficial signal regression, in which the average of all first-nearest-neighbor measurements is linearly regressed out of all other source-detector measurements as an estimate of systemic nuisance signals including global head motion [53,58]. This combination of exclusion of noisy channels, superficial signal regression, and motion scrubbing forms a set of stringent measures to prevent head motion from confounding our decoding results. Although these approaches may have reduced the total number of time points ultimately included in the spatiotemporal template matching calculation, this loss of data was preferred over the inclusion of motion artifacts that could either impair decoding or spuriously enhance the measured accuracy of neural decoding if classification was driven by any replicable patterns of motion across runs.

Inter-individual differences in structural and functional neuroanatomy have become an area of increasing focus in fMRI research [89], and also influence the precise neural responses that are sampled by optical imaging methods. The data processing pipeline utilized subject-specific head models during tomographic reconstruction and associated subject-specific field-of-view masks during data processing in an attempt to mitigate the effects of this anatomical variation. Indeed, other studies by our group, yet to be published, have shown that decoding performance is higher with neuroanatomically reconstructed data than with data in measurement space. Our interpretation of this result is that head models transform data from source-detector measurement

space into an alternative space where the features have more information content. In this regard, tomographic reconstruction has similar effects to commonly utilized dimensionality reduction techniques that remap data into more informative feature spaces, except here the remapping algorithm is based in neuroanatomy, biophysical models of light propagation, and sensitivity matrix inversion involving Tikhonov regularization to address an ill-posed inverse problem. In light of this observation, it is reasonable to hypothesize that more accurate reconstructions would amount to a more effective feature space for more accurate decoding. Previous studies have shown that atlas-derived head models can be effective in the absence of subject-specific structural data [81], but subject-specific head models improve image quality. Therefore, subject-specific head models were used for data reconstruction in the current study. Future studies could compare decoding performance with generic and subject-specific head models to systematically assess the effects of this data processing step on decoding.

Aside from all these raw data quality and analysis considerations, some variation in decoding performance may result from differences between runs and individuals in cognitive processing of movies. For example, the perceptual phenomenon of repetition suppression [147] may have diminished responses over repeated movie viewings, or the repeated exposure may have otherwise influenced measured responses [148]. Eye tracking studies have shown the reproducibility of gaze patterns over repeated movie viewings both within and across participants [149,150], but their observed variance in gaze position is still non-zero and likely to influence low level sensory responses as well as attentional focus. Since attentional modulation can greatly influence sensory processing [151–153], changes in both attentional biases and general arousal

level between runs could have caused sporadic variation in neural responses to movie stimuli. Other factors such as changes in emotional state, either driven by or unrelated to the movies, could have also contributed to variability in how participants engaged and responded to stimuli across viewings [154,155].

4.4.4 Further characterization of decoding performance

The consistently high decoding accuracy observed with our initial parameters established a performance baseline that supported challenging the template matching algorithm and the HD-DOT data through manipulations such as decreasing trial duration and increasing the number of trial types. Mean decoding accuracy remained above 80% on reducing template and trial duration from 4 minutes down to 1 minute. While further reducing trial duration had a larger impact on accuracy, decoding remained significantly better than chance even with just 15 seconds of data. However, more marked variation between participants was noted for the shorter trial durations. For instance, decoding accuracy for participant 3 declined precipitously for trials shorter than 1 minute, but remained above 90% for participant 1 even with 30-second trials. With 15-second trials, accuracy remained above 65% for 2 participants, but fell near chance levels for the third. As expected, errors also became more common with an increasing number of template and trial types to be classified, but remained significantly better than chance in all cases tested.

The trial durations used here are longer than those used by a previous HD-DOT decoding study that classified checkerboard stimulus location based on data from individual time points [119], but the stimuli used in the current study are also significantly more dynamic, multi-dimensional,

and variable. One study decoding silent movies from fMRI data used clips that were 10-20 seconds long but compiled into 1-10 minute chunks [39]. Another study used training clips averaging 23 seconds in duration and test clips averaging 4 seconds, but both combined into many 8-minute-long blocks that were presented multiple times each [40]. However, comparisons to these fMRI studies come with major caveats, as they used substantially larger stimulus libraries and more elaborate decoding algorithms to reconstruct novel stimuli in 1-second bins or framewise. These more ambitious decoding endeavors capitalized on a more well-established imaging modality that is routinely used to pool many hours' worth of data across multiple imaging sessions, as was required to amass sufficient training and test data. Replicating these full-scale fMRI decoding studies with HD-DOT imaging will be more feasible once ongoing efforts address challenges with cross-registration and pooling of HD-DOT data across imaging sessions. However, the current work still makes an important first step in this direction, and still represents an important advance for the optical neuroimaging field. The >90% accuracy observed with 4-way decoding is on par with the most accurate of decoding outcomes previously achieved with optical neuroimaging in humans [119]. Furthermore, the majority of prior fNIRS decoding studies have been limited to decoding only 2-4 different states. A previous HD-DOT decoding study reported on 18-way and 36-way classifications of visual stimulus position using checkerboard stimuli [119]. This prior work is emulated by the 8-way, 16-way, and 32-way classifications reported here using significantly more complex, multi-sensory, naturalistic stimuli.

4.4.5 Decoding isolated auditory and visual responses

With the strong inter-run correlations observed in both temporal areas associated with auditory processing and occipital areas associated with visual processing, the ROI-confined analysis attempted to disentangle decoding of auditory and visual responses while still utilizing the wealth of audiovisual movie viewing data available. Decoding performance remained high with template matching confined to either ROI, suggesting the feasibility of decoding naturalistic auditory or visual information independent of one another. However, the ROI-based approach was still an imperfect test of this hypothesis, as neither ROI definitively excludes responses to the other sensory modality. For example, responses to visual features such as faces have been reported in right temporal cortical regions like the superior temporal sulcus [134], which lies within our task-data-derived auditory ROI due to its close proximity to the superior temporal gyrus where auditory responses are centered and due to the spatial blurring associated with the point-spread function of optical imaging systems. The purely auditory and purely visual stimulus data available from this data set allowed for a more controlled evaluation of auditory and visual decoding in isolation, even though only a smaller amount of this data was available. Decoding remained accurate across these sessions as well, for both audio clips presented without visuals and silent movies stripped of audio. The promising results of this pilot analysis support further studies focused on decoding of both sensory modalities with HD-DOT, with various interesting questions remaining to be addressed. Although the audio clips included much naturalistic speech, it is unclear from the current results how much auditory decoding was driven by low level features of speech, its semantic content, or non-speech sound such as background music. Future

experiments could evaluate these possibilities by comparing decoding of normal speech, reversed speech, songs, instrumental music, and other non-speech sounds from HD-DOT data. Similarly, it remains to be addressed what components of the movies drove the observed visual decoding – low level characteristics such as luminance, color, and optic flow, higher level visual features such as faces and objects, or other factors such as social interactions depicted and emotional reactions evoked by the movies. Greater understanding of these questions could inform or be informed by the development of more elaborate feature encoding models, which would also enable extrapolative decoding of novel stimuli outside the training set of stimuli [3,38,39].

4.4.6 Limitations and future directions

As already alluded to, the greatest limitations of the current study stem from its opportunistic use of a preexisting dataset, designed and acquired for other purposes, to begin exploring the simplest version of a complicated problem. The stimulus set, number of presentations, and number of participants were limited: 20 different movie clips were presented twice each to 3 participants. Movie clips were several minutes long, and their order of presentation was not randomized. And while training and test HD-DOT data were separate, classification of test data was only among a limited set of 4-32 choices contained within the training stimuli. However, many influential imaging and decoding studies have focused on small numbers of highly sampled participants [38,39]. Future data collection efforts targeted for decoding could address the other study design limitations. And the current study utilizes existing data to address a question that is within its scope. While the most powerful decoding algorithms strive to remain

versatile and classify diverse stimuli, the straightforward template matching strategy used herein circumvents that concern by intentionally seeking to address a simpler but still crucial precursory question: Are responses to complex, naturalistic, multisensory stimuli as measured by HD-DOT rich, repeatable, and specific enough to support reliable classification of test data among a limited set of options represented in the independent training data? Our analyses are affirmative, and show that HD-DOT recordings of brain activity have high information content that could be mined by future studies developing more elaborate encoding models and decoding algorithms towards broader applications. Among the immediate next steps, dimensionality reduction methods such as principal components analysis could be used to identify a smaller set of specific features accounting for the majority of the variance in our data. We could then test whether transforming data into this feature space prior to decoding enhances accuracy under the stress tests of shorter trial durations and more template options. Furthermore, it has been argued that accurate feature encoding models facilitate effective feature decoding [3]. We have previously developed movie feature regressors to map responses to content such as speech, faces, and other high and low level visual and auditory features of movies onto the brain [64]. Inverting a matrix of such feature regressors could be one approach to decoding specific movie features from HD-DOT data and potentially decoding movies that were excluded from the training set.

Previous fMRI studies of visual decoding have used receptive field models to decode novel static images [38], motion energy encoding models to reconstruct dynamic visual stimuli [39], and convolutional neural networks to encode and decode hierarchical layers of visual stimulus features and neural responses [40,41]. Other fMRI studies have used a variety of methods,

including multivariate pattern analysis, ridge regression, and generative models, to encode and decode acoustic, articulatory, and semantic content of speech [71,72,156–159] as well as other auditory stimuli like music [73]. In addition, ECoG studies have taken various approaches, from linear classification methods to a variety of feature encoding models, to decode articulatory or spectral features during speech production and thereafter synthesize intelligible speech from electrical recordings of brain activity [42,43,160,161]. Even more recently, highly accurate and efficient brain-to-text communication has also been achieved in a quadriplegic patient by decoding motor cortex activity sampled with surgically implanted intracortical microelectrodes during imagined handwriting [145]. While certainly impressive and inspiring, these studies are currently constrained in their end-applications by the logistical limitations of fMRI and the invasiveness of ECoG. Nevertheless, this literature provides invaluable leads for decoding studies using alternative imaging modalities to replicate and build on with fewer logistical constraints limiting their translation into real-world contexts.

In fact, reliable decoding of the sorts of complex signals classified here even just among a limited menu of options could have powerful applications. For instance, if comparable classification performance can be achieved using covert signals, HD-DOT could be used for algorithmic approaches to augmented communication, e.g. using different internally generated mental states to represent key words [46] or represent different characters to spell out any word [162]. Though these approaches may have a low bit rate, they could still be immensely empowering for individuals such as patients with locked in syndrome who have few or no other means to communicate. Furthermore, the complex decoding achieved with HD-DOT could still

increase bit rate relative to the binary decoding previously pursued in these patients with traditional fNIRS [46]. Future studies could explore these modes of communication with HD-DOT first in typical populations and then in clinical cohorts.

In addition to investigating more elaborate decoding and developing clinical applications, future studies could also push towards decoding during increasingly natural paradigms. While free viewing of audiovisual movies provide a segue from block-design tasks, imaging studies have begun to explore even more naturalistic paradigms such as immersive three-dimensional movies [163], interactive virtual reality [164], and face-to-face human interactions [128]. While harder to control, these interactive paradigms attempt to enhance the ecological validity of laboratory research and move closer to real world applications. With its combination of high image quality and an open scanning environment, HD-DOT is well suited as a tool for both mapping and decoding information during such paradigms.

In all, the template matching decoding performance observed here establishes the fidelity of HD-DOT signals in response to naturalistic audiovisual stimuli in both visual and auditory regions of the brain, and further establishes the feasibility of decoding these complex signals. These findings comprise a major step forward for decoding with optical neuroimaging, and encourage future HD-DOT studies geared towards more elaborate decoding, more natural paradigms, and clinical applications.

"Was it everything you wanted to find?"

- Train

Chapter 5: Conclusions and future directions

The final chapter aims to synthesize the highlights of the other chapters and discuss the implications and future directions for the fields of optical neuroimaging, developmental cognitive neuroscience, and neural decoding research.

5.1 Summary of previous chapters

5.1.1 Decoding visual stimulus position from HD-DOT data (Chapter 2)

- Applications of neural decoding have previously been constrained by mainstream neuroimaging modalities, specifically by the tradeoffs that they require between image quality and logistics. Elaborate decoding has been achieved with data from invasive electrical recording methods or MRI scanners confined to large clinical and research facilities. More portable, noninvasive methods like fNIRS and EEG can be utilized in a wider range of populations and settings, but their traditionally poor image quality has limited the level of detail and accuracy of decoding.

- HD-DOT combines the logistical advantages of optical imaging with improved image quality, but the feasibility and performance of decoding had not previously been evaluated using brain activity measured with HD-DOT.

- We were able to decode the binary position of a flickering checkerboard stimulus from HD-DOT data collected in typical adults at a single-trial level with high sensitivity, specificity, and reproducibility. For instance, we recorded ROCs with areas under the curve greater than 0.97 across 10 imaging sessions in a highly sampled participant.

- ROC analyses of decoding across 5 participants established both replicability of decoding in multiple individuals and the feasibility of inter-individual decoding, although performance varied between individuals.

- More complex, non-binary, pseudo-real-time decoding was evaluated through 18-way and 36-way classifications of the positions of moving checkerboard stimulus patterns at each time point in their trajectory. For instance, across 3 highly sampled participants, the phase of a 60° wide checkerboard wedge rotating 10° per second through 360° was decoded with a within-participant error of 25.8±24.7°. Decoding between participants was also feasible here according to permutation-based significance testing.

- Thus, we concluded that visual stimulus information can be decoded accurately, reproducibly, and across a range of detail (for both binary and non-binary outcomes) at the single-trial level (without needing to block-average test data) using HD-DOT data.

5.1.2 Mapping brain function during naturalistic viewing in awake young children with HD-DOT (Chapter 3)

- Typical and atypical trajectories of functional brain development during early critical periods remain understudied because of the challenges of mapping brain function in young children. It is difficult for young children to lie alone, calm, and still inside the dark, cramped, loud bore of an MRI scanner for an extended duration. Alternative neuroimaging modalities like fNIRS have traditionally lacked image quality due to sparse sampling. In addition, conventional tasks for brain mapping in compliant adults elicit low task engagement, high head motion, and considerable participant attrition in pediatric populations.

- We designed and built a new HD-DOT system geared towards imaging young children with a combination of features to ensure both high image quality and child comfort.

- Features to maximize image quality included laser optics and a high channel count of densely packed measurements across a wide field of view. An array with 11mm grid spacing of 128 laser sources and 125 detectors using a 2-pass source encoding pattern yielded up to 5202 usable measurement channels with <50mm source-detector separation across the posterior, lateral, and dorsal surfaces of the head, spanning 7 orders of magnitude of optical power.

- Features to increase cap wearability and child comfort included an adjustable multi-panel cap design, soft plastic fiber tips, lightweight fiber optic cables evenly supported by a pair of suspended halos and a padded frame, a cushioned and adjustable pediatric imaging chair with adjacent seating space for caregivers, and a child-friendly under-the-sea themed décor across the exterior of the imaging system and the imaging suite.

- Healthy young adults were imaged to validate system performance, using light levels and pulse signals as indicators of raw data quality, and responses during auditory, visual, and motor tasks to evaluate image quality.

- Subject-specific head models were developed using structural imaging data and shown to produce more accurate reconstructions than a generic head model

- A library of children's movie clips was characterized as a stimulus set for brain mapping. We found that HD-DOT could be used to map reproducible responses to movies and contained features such as speech and faces across a variety of animated clips. Variation in the reproducibility, construct validity, and orthogonality of regressor maps could be partially explained by properties of both the movie stimuli (such as feature modulation and orthogonality) as well as the imaging data (such as inter-run synchronization).

- A multivariate approach to feature regression was found to mitigate confounding of regressor maps by feature covariance during parallel feature mapping.

- Finally, the optimized imaging system, selected movies, and refined speech regressor analysis pipeline were used to map receptive language function during movie viewing with HD-DOT in a group of awake, 1-6 year-old children. We mapped responses to speech during movies bilaterally in the superior temporal gyrus, comparable to activations observed in both children and adults during a traditional block-design word-hearing task.

- Therefore, we concluded that HD-DOT imaging during movie viewing can be used to map reproducible cortical responses to movie features such as naturalistic speech in awake young children.

5.1.3 Decoding naturalistic auditory and visual information from HD-DOT data (Chapter 4)

- Effective decoding of naturalistic information from brain activity measured in natural settings would have exciting neuroscientific and clinical implications. Thus, having established the feasibility and performance of both decoding simple visual stimulus information from HD-DOT data (Chapter 2) and mapping responses to audiovisual movies with HD-DOT (Chapter 3), we sought to assess whether complex, naturalistic, auditory and visual movie stimulus information could be decoded from HD-DOT data.

- The previously collected data set from adults viewing our children's movie stimulus library was repurposed to study decoding.

- A template matching strategy was used to decode which of 4 audiovisual clips a participant had viewed during any given 4-minute trial from their HD-DOT data with 92.3±4.4% accuracy across 13 imaging sessions in 3 participants.

- Systematically reducing trial duration and increasing the number of template choices gradually reduced decoding accuracy, but performance remained significantly above chance even with 15-second trials and 32 choices.

- Decoding accuracy was similarly high when template matching was confined to task-derived visual and auditory regions of interest or applied to data in which participants were presented with purely auditory and purely visual clips.

- As a result, we concluded that complex, naturalistic auditory and visual stimulus information can be accurately decoded from HD-DOT data.

5.2 Implications and Applications

The methods and findings established in this dissertation could be further built upon and applied through a variety of avenues of future investigation. This section focuses on discussing some of the: (1) next steps in developing optical neuroimaging methods, (2) potential developmental cognitive neuroscience research applications to studying models of neuroplasticity, and (3) possible applications of decoding beyond the laboratory.

5.2.1 Further optical neuroimaging method development

From a methodological perspective, this body of work focused on improving instrumentation and data analysis pipelines while exploring novel applications for HD-DOT. Specific efforts to improve instrumentation included further increasing channel count, measurement density, and field of view relative to previous systems, as well as incorporating laser optics and a new, adjustable cap design. Work on analysis pipelines included optimizing data synchronization, developing a children's stimulus library with associated feature regressors, showing the benefits of multivariate feature regression, and evaluating the impact of subject-specific head models. The main applications pioneered were imaging in awake young children during naturalistic movie viewing and decoding of sensory stimulus information from brain activity as measured by HD-DOT. The experimenters' experiences while implementing these efforts and the results finally obtained encourage several avenues of further investigation pertaining to method development for optical neuroimaging in humans.

Previous studies have directly shown tangible improvements in image quality resulting from dense sampling in other age groups [30,56,61]. However, the benefits for decoding and preschool-age imaging of using high-density arrays relative to sparser ones, while suggested by our studies, have yet to be explicitly quantified. Decoding performance could be reevaluated for both checkerboard stimulus position classification and naturalistic movie decoding using sparser subsets of our measurement arrays and the results compared to those obtained with our full set of measurements. Similarly, maps of inter-run synchronization and speech responses during movie viewing could be computed with sparser subsets of the collected channels to directly assess

improvements in pediatric image quality resulting from our high channel count and high density of measurements. Furthermore, new systems could be built with even higher densities of sources and detectors to evaluate how far image quality can be pushed in practice by extending this strategy.

While we were satisfied with the image quality obtained across the posterior and lateral panels of our preschooler imaging cap, pulse SNR was found to be lower for the dorsal pad, and maps and time traces were correspondingly noisier for motor tasks. This shortcoming will be crucial to address if we aspire to effectively decode overt or covert motor signals associated with limb movements or articulation. However, this was our first attempt at building a wide-field HD-DOT system with coverage of the dorsal surface of the head, and cap building is notoriously an iterative process in industrial design. Future iterations of the imaging system can focus on optimizing the design and use of the dorsal pad to improve SNR across the somatomotor cortex. Furthermore, future systems can attempt to further expand the field of view, adding further frontal cortex coverage. Indeed, prefrontal cortex is a popular target of fNIRS studies due in part to the relative ease of managing hair over the forehead [165], and it should similarly be easy to use HD-DOT for mapping prefrontal cortex activity during cognitive control tasks. The main challenge is likely to in molding a cap across the entire head surface in spite of all the variable points of curvature. However, this design challenge is likely to prove solvable through technical innovation in the materials and methods used to construct imaging caps.

When building the preschooler imaging system, we incorporated several features intended to make the cap and imaging environment comfortable for children. However, additional comfort

features could help further improve child compliance and data retention. Future builds could attempt to disperse pressure points by adding some padding where the cap tends to most tightly contact the head, using even higher optode densities (to not only further increase resolution but also magnify the bed of nails effect), and incorporating even softer materials to cover fiber tips and Velcro buckles. It may also be worthwhile to evaluate the impact of gradually habituating children to the sensation of the cap on the head by providing mock caps to wear prior to imaging sessions or by conducting training sessions similar to mock MRI scans [126]. Perhaps the most exciting prospect for increasing wearability is the development of wireless, fiber-less HD-DOT systems [107,108], which would also present other advantages. By using compact optics and electronics to incorporate sources and detectors in the body of the cap itself, fiber-less systems remove the weight, tension, torque, and logistical constraints of the massive bundles of fiber optic cables currently used by fiber-based high-density caps. As a result, fiber-less systems are also likely to reduce motion artifact and facilitate building caps with increased areal coverage. And of course, wearable systems will enable studies of more natural behaviors in more natural settings, unshackling neuroimaging research from the constraints of the laboratory.

In addition to advances in hardware, the pediatric imaging project also drove the development of several data analysis tools. Considerable work occurred behind the scenes on troubleshooting initial issues with synchronizing the neuroimaging data with the presented movie stimuli and associated feature regressors, culminating in the modification of data processing pipelines to ensure the accurate alignment that is required for time-locked stimulus response analyses. The final solution involved computing cross-correlations between audio recorded concurrently with

the imaging data and the audio attached to the stimuli and regressors, and shifting time courses by the lag that maximized the correlation between the two audio tracks. This measure corrected for artifactual jittering of data relative to stimuli by previous versions of the preprocessing code, and drastically improved the reliability and validity of all subsequent analysis. Equally integral to meaningful movie data analysis was the generation and quality control of a set of feature regressors for the movie stimulus library. These stimuli and regressors could be reutilized for a host of future studies, but additional feature regressors could also be developed for further analysis of data sets that have already been collected. This option to probe new hypotheses with the existing data is a product of the richness of the movie stimulus paradigm and the flexibility of the feature regressor analysis approach. One ongoing endeavor involves creating regressors for additional features of interest besides speech and faces, ranging from low level visual stimulus characteristics to emotional valence and social content. In addition, we are trying to utilize machine learning approaches for speech and face recognition to efficiently generate feature regressors for any new movie clip. This innovation would allow us to analyze data collected with an infinitely diverse stimulus library that could cater to any child's movie viewing preferences and maximize their compliance, capitalizing on our finding that reproducible feature regressor maps can be obtained across heterogeneous sets of movies. A richer set of regressors would also further bolster the multivariate regression approach that we showed minimizes confounding of feature response maps by covarying movie features.

Besides refining these tools specific to the analysis of movie viewing data, we also showed that subject-specific head modeling more generally improves the accuracy with which HD-DOT data

can be reconstructed. However, our current approach to generating these models is highly labor intensive, involving iterative adjustments to cap positioning based on subjective assessments of control task activation maps. Though it may prove to be computationally intensive, automating this process would allow for less subjective, more accurate, and more efficient generation of not only subject-specific but also session-specific head models. This would facilitate more accurate reconstructions and cross-registration of data across multiple imaging sessions, allowing for more robust analyses. For instance, it has been shown that larger quantities of resting state data support more reliable functional connectivity analysis [99], while large quantities of movie viewing data can be pooled to fuel more elaborate decoding algorithms [38–40]. In addition, extending these advances to children requires developing a subject-specific head modeling pipeline that does not depend on having the participant's structural MRI data, which may be challenging to acquire in children. Ongoing efforts in the lab aim to collect three-dimensional photographs of participants both with and without the imaging cap use these data to morph atlas-derived head models. If successful, this approach would be applicable not only to pediatric studies but also HD-DOT research in populations with contraindications to MRI and in settings where an MRI scanner is not accessible or convenient to use.

Finally, our work focused on high-density continuous wave measurements of hemoglobin signal contrasts, but there are several other optical imaging techniques being developed by other groups that could be used in parallel or potentially even combined with our approach to further improve image quality or capture other contrasts. A non-exhaustive list of these other methods includes frequency domain DOT [166], speckle contrast DOT [167], optical imaging of cytochrome

oxidase [168,169], and holography to increase the depth and resolution of image reconstructions [170]. Furthermore, as a result of the compatibility of optical imaging instrumentation with most other technology, HD-DOT could be combined with electroencephalography, electromyography, eye tracking, transcranial magnetic stimulation, and various other recording and stimulation methods in multimodal studies probing diverse neuroscience questions. With so many avenues for further development, it is fair to say the future of human optical neuroimaging is bright.

5.2.2 Models of plasticity

Early childhood is packed with myriad dramatic behavioral changes reflecting dynamic underlying changes in brain function. Both typical and atypical trajectories of early brain development can hence serve as models of neuroplasticity. These models have not been easily accessed by extant neuroimaging modalities due to the challenges of imaging children with some (e.g. fMRI) and the limited image quality of others (e.g. fNIRS). However, the HD-DOT imaging methods developed in this dissertation, combining a child-friendly imaging environment with enhanced image quality, are well suited to addressing this void in the developmental cognitive neuroscience literature. While a complete listing of currently understudied models of plasticity that could be investigated with HD-DOT is beyond the scope of this chapter, several compelling examples are discussed here. Through the proposed studies, HD-DOT could shed new light on not only typical brain development and developmental neurological disorders, but also more generally on how neural circuits might reorganize across the lifespan in health, amid disease, and during recovery.

Strabismus

The importance of normal binocular integration during development was first brought to light in the 1960s by the Nobel Prize-winning work of Hubel and Wiesel on feline models of strabismus (abnormal deviation of one or both eyes) and monocular deprivation [13–15,97]. Their seminal experiments found that binocular decorrelation early in life causes lasting abnormalities of visual cortex structure and function. This finding inspired decades of work on how sensory experience can shape brain development. To honor that rich neuroscience research history, this subsection takes a deep dive into strabismus as a model for studying neural mechanisms underlying atypical development and cortical plasticity with HD-DOT.

Since Hubel and Wiesel's original experiments, studies have identified increased interocular suppression in visual cortex in animal models of strabismus, wherein input from one eye appears to inhibit the response of visual cortex to the other eye [171–175]. This excessive interocular suppression has been variably interpreted as an adaptive response to strabismus [176], as a pathologic mediator of its detrimental effects on vision [177–179], and, intriguingly, even as a potential cause of strabismus itself [180]. However, our ability to resolve these theories about the mechanistic importance of strabismic suppression in humans is limited by a lack of knowledge about when this process develops. Due to the challenges of pediatric neuroimaging, strabismic suppression has thus far mainly been studied using behavioral approaches and in older children or adults who have already had strabismus for years or decades, even though strabismus itself most often emerges in children between birth and 4 years of age [175,180–182] As a result, it is currently debated whether cortical suppression precedes strabismus, develops immediately after

the onset of strabismus, or emerges gradually afterwards. Thus, over 50 years after the original studies by Hubel and Wiesel, we still lack a clear understanding of the neurobiology of human strabismus. Strabismus and its resulting visual deficits continue to affect up to 5% of the population, and optimal disease management remains a matter of debate [183].

The increased interocular suppression seen in strabismus has long been argued to serve an adaptive role [176]. The ocular misalignment of strabismus effectively creates a state of constant dichoptic viewing. It has been posited that a high degree of long-term suppression of visual input from the misaligned eye serves to prevent strabismic individuals from experiencing constant double vision [176].

However, it has also conversely been argued that strabismic suppression may play a pathological role [177–179]. Chronic suppression during development could prevent activity-dependent maintenance of the visual cortical circuitry downstream of the misaligned eye, serving as a potential mechanism underlying the visual acuity loss that is identified clinically as amblyopia. In addition to this developmental effect, interocular suppression may also actively interfere with visual processing by the deviated eye and preclude binocular function in strabismic individuals [177–179]. Experiments show that at least some architecture capable of supporting binocular integration survives despite the occurrence of strabismus through development, but remains masked by interocular suppression [173,174,184–186]. Overcoming this suppression – either pharmacologically in electrophysiological experiments on animal models of strabismus [173,185,186], or through stimulus manipulations in psychophysical studies with amblyopic adults [174] – enables a restoration of some degree of binocular integration. This concept is the

basis of a growing field of research attempting to develop visual exercises and video games to overcome interocular suppression and enhance binocular function in patients with amblyopia [187,188]. The increased interocular suppression in strabismus could hence be maladaptive both during development and in adulthood, highlighting the importance of understanding how this disease mechanism emerges and operates.

Importantly, it has also been suggested that interocular suppression may not only play an adaptive or pathological role downstream of strabismus, but may also precede strabismus and cause it to develop in the first place [180]. Clinical evidence and animal model studies suggest that strabismus may itself be triggered by various other initial causes of cortical binocular decorrelation [189,190]. For instance, research has shown that several manipulations imposing binocular decorrelation early in life cause strabismus to emerge in monkeys [189,190]. It has also been argued that various perinatal insults (e.g. periventricular hemorrhage, cerebral visual pathway white matter injury) may serve as natural causes of binocular decorrelation in humans, explaining their strong association with infantile strabismus [189]. It is thought that such initial triggers of binocular decorrelation may initiate processes such as a loss of horizontal connections between ocular dominance columns and increased interocular suppression in visual cortex, similar to the effects of strabismus itself [171,189,191]. These processes may then prevent the binocular integration that is required to guide the control of fusional vergence by the medial superior temporal area, resulting in the observed ocular misalignment and associated symptoms of strabismus [180,189,192]. This hypothesis hence ascribes interocular suppression a causative

role in the etiology of strabismus and implies that it precedes the occurrence of strabismus itself, but it is yet to be shown whether this is the case in humans.

Interocular interactions can be studied in the context of binocular rivalry, a phenomenon wherein individuals are presented with different non-fusible images to each eye and spontaneously alternate between perceiving one image and then the other over time [193]. FMRI studies of binocular rivalry in adults have shown alternating patterns of visual cortex activity aligning with the alternating perception reported by participants [194–197]. Furthermore, behavioral studies have shown that some properties of binocular rivalry change over typical development; for instance, the rate of perceptual alternation has been found to slow with age [198]. However, the neural correlates of binocular rivalry have not yet been compared across different age groups, and there is a lack of studies of binocular rivalry in children who are 3-4 years old, an age at which abnormalities in cortical interocular interactions may drive strabismus [180].

Although it is challenging for children to continuously report their subject experience during a dynamic paradigm like binocular rivalry, it has been shown in adults that a participant's visual experience during binocular rivalry can be decoded from their visual cortex activity recorded using fMRI [196]. Building on the visual decoding results detailed in Chapters 2 and 4, future studies could attempt to decode visual experience during binocular rivalry from HD-DOT recordings of visual cortex activity in typical adults. If successful, this method could be extended to studying adults with strabismus, typically developing children, and children with a new onset of strabismus, so as to investigate whether abnormal interocular suppression exists at the onset of strabismus or emerges later. If abnormal cortical suppression is already found at the time of

strabismus diagnosis, longitudinal studies of children predisposed to strabismus by a family history or other risk factors could evaluate whether the suppression in fact predates strabismus. This work could identify neuroimaging biomarkers of abnormal cortical suppression and also pave the way for longitudinal studies of visual cortex plasticity and critical periods associated with clinical interventions for strabismus. For instance, cortical interocular interactions could be mapped before and after the surgical correction of eye alignment in children incidentally receiving treatment at different ages. Such a series of studies would be the ultimate translation of Hubel and Wiesel's foundational work into humans, elucidating developmental trajectories and critical periods for binocular integration, mechanisms of disease and resilience, diagnostic and prognostic biomarkers, and optimal clinical management for strabismus and amblyopia.

Malnutrition

Many research efforts, humanitarian service projects, and public policy initiatives have focused on "the first 1000 days of life" as a period when sufficient nutrition is paramount for fueling the high energy and nutrient needs of the developing brain [144]. Epidemiological and behavioral studies show that nutritional deficiencies during this epoch can permanently impair cognitive development. However, the specific effects of malnutrition on brain function remain poorly understood for multiple reasons. For one, the general challenges of pediatric neuroimaging that have been discussed at length in previous chapters of course apply to imaging children with malnutrition during the first few years of life. In addition, given that a local lack of resources is a root of the problem of malnutrition, imaging equipment like MRI scanners have been a distant

luxury and functional brain imaging studies have been a low priority for affected communities. However, malnutrition affects hundreds of millions of children worldwide, accounts for 45% of child deaths, and limits cognitive potential in survivors [199]. Just as standardized anthropometric indices facilitate grading the severity of malnutrition and tracking responses to clinical interventions based on somatic measurements [200], biomarkers of brain function are needed to predict neurological repercussions of malnutrition and monitor responses to treatment in both research and clinical care.

Several groups have piloted the use of fNIRS, as a cost-effective method conducive to imaging in low-resource settings, for studying how malnutrition affects the brain [32,201]. However, the low channel count of the imaging systems used in most of these studies have yielded coarse indicators of neural function, such as overall cerebral bloodflow and metabolism [201] or oxyhemoglobin signals that are poorly localized to approximate positions on the scalp as opposed to specific regions of the brain [32]. A recent study established the feasibility of HD-DOT imaging in 7-10 year-old children in low-resources settings in Cali, Colombia [61]. The next phase of that research is underway, using the same imaging system to map neural correlates of malnutrition in this age range. Future studies could use an imaging system like the one developed in Chapter 3 to extend these studies into younger children. Moreover, there have been exciting movements to develop and administer locally targeted therapeutic interventions for children with severe malnutrition in communities around the world. These solutions include ready-to-use therapeutic foods [202], locally sourced and produced therapeutic foods [203,204], and microbiome-directed food interventions [205,206]. HD-DOT imaging could be utilized in

low resource settings to study how brain function responds to these nutritional therapies whenever they are administered, potentially helping to identify which interventions could have the best cognitive outcomes as well as mapping neural changes and critical periods associated with cognitive recovery.

Language

While the physician's mind leaps at the opportunity to investigate clinical problems, one of the most remarkable and certainly most widespread models of plasticity is the typically developing brain. Children exhibit a wide range of behavioral changes as they hit their developmental milestones, indicating underlying neuroplasticity across functional systems, spanning sensory, motor, cognitive, emotional, and social domains. Among the most striking and celebrated of developmental changes is the emergence of language. Children typically progress from being able to distinguish speech from analogous sounds by 3 months postpartum [18,207] to recognizing familiar words by 8 months [208] to exhibiting a 200-fold expansion in their receptive vocabulary by 2 years of age [21]. Simultaneously, they move from recognizing prosodic patterns and other coarse features of speech at birth through prenatal intrauterine exposure to learning complex syntactic rules and other elements of grammar that complement their growing vocabulary [18]. On the heels of these rapidly expanding comprehension skills, children's speech production abilities typically grow from incomprehensible babbling at 6 months to first words around 12 months to sentences and eventually fluent stories over the next 1-4 years [19]. Despite all these exciting events, fMRI studies of language development have

almost skipped over the period from birth to 4 years due to the challenges of scanning across this epoch [20]. Exceptions include a study of speech perception in 3-month-old infants [121], who have barely begun to identify speech, and a study of passive listening in sleeping 2-3 year-olds [21]. Neuroimaging studies of active listening and speech generation could provide much additional insight into the neural changes underlying the developmental burst of language comprehension and production capabilities. And as language is of course largely a uniquely human capability, our understanding of its emergence is heavily dependent on developing the tools to conduct these studies in humans.

Our effective mapping of receptive language function during movie viewing with HD-DOT in a pilot sample of awake 1-6 year-old children suggests that HD-DOT could be applied to bridge this gap in the literature. We have already begun recruiting more children within discrete age brackets spanning 0-4 years to conduct a larger study of early functional brain development with HD-DOT. Readouts pertaining to language development including movie speech regressor maps, functional connectivity of language areas, and behavioral tests measuring receptive and productive language function. These measures could be collectively used to evaluate changes in functional brain organization concurrent with language acquisition and brain-behavior correlations associated with language performance. In addition to illuminating the typical developmental trajectories of language networks in the brain, this work will provide a reference for better understanding differences in atypical populations, such as children with autism or specific language impairment. Given its feasibility in young children, HD-DOT imaging can be applied to study these atypical populations as soon as they have been diagnosed or even

beforehand by preemptively studying participants predisposed to these conditions by a family history or other risk factors. Finally, studies in atypical populations receiving clinical interventions would offer an additional opportunity for investigating cortical remapping in language networks. For example, tracking changes in brain responses to heard speech after hearing impaired children receive cochlear implants at different ages could provide a window into critical periods for the development of speech processing. Such studies would further capitalize on the compatibility of HD-DOT with metallic medical devices such as cochlear implants, which are a contraindication for MRI scans even in adults. HD-DOT studies of older children with autism spectrum disorder and adults with cochlear implants are already underway in the lab, and the new instrumentation and movie viewing data analysis methods developed by this dissertation will facilitate extending that research into younger participants.

Motor system

Recent research has highlighted the remarkable plasticity of motor networks even in the mature adult brain. For instance, one study attempted to induce plasticity in typical adults by casting the dominant arm for 2 weeks and evaluated the effects of this intervention on resting state functional connectivity measured with fMRI [209]. A marked reduction of functional connectivity was observed between brain regions responsible for controlling the casted arm and the rest of the somatomotor network, and this change reversed after removal of the cast. Large, spontaneous pulses of brain activity were also observed propagating through the motor control regions being underutilized due to the limb constraint as well as cingulo-opercular executive

control regions, potentially helping to maintain neural circuits that are otherwise temporarily disconnected and thereby possibly facilitating the rapid recovery later observed [209,210]. The authors attributed their success in capturing such dramatic signatures of neural plasticity in part to their high level of sampling, as they conducted over 30-minute-long fMRI scans on a daily basis for 6-9 weeks surrounding the casting intervention [209]. This study design epitomizes the burgeoning field of precision functional neuroimaging, wherein studies focus on a small number of highly sampled individuals [89,99]. This approach aims to maximize the reliability of findings by reducing temporal sampling bias [99], to minimize the blurring of results as a product of inter-individual variability by conducting analyses within individuals instead of averaging across them [89], and to generate rich datasets that can be mined to investigate numerous hypotheses with high granularity.

HD-DOT has the potential to extend this approach even further than fMRI has been able to thus far. Portable HD-DOT systems are already being utilized to collect hours' worth of longitudinal data at the bedside to investigate network disruptions and cortical remapping in populations such as recovering stroke patients [62,63]. Some neuroscience research studies using other modalities have sent participants home with wearable imaging devices to collect longitudinal data in ecologically valid settings, for instance using sparse EEG arrays to track brain activity during natural sleep over multiple nights [211,212]. As wireless, fiber-less HD-DOT technology [107,108] become increasingly wearable and available, participants could conceivably be imaged for hours every day while engaging in their daily activities in their home environment. Of course, methodological rigor will be required in the analysis of such data, from addressing artifacts

caused by head motion to correcting for multiple comparisons on an immense scale. However, these experiments could provide insight into potential contributions to inter-individual differences identified in the casting study, such as diurnal variation in spontaneous activity patterns or the effects of the recent history of motor activity immediately preceding data acquisition. Moreover, given the compatibility of optical imaging hardware with other equipment, multimodal studies could be implemented to address intriguing physiological questions, e.g. HD-DOT could be combined with electromyography and accelerometry to explore just how far the spontaneous pulses of brain activity detected during disuse propagate along the motor system. And while these particular ideas aim to further investigate an adult model of neuroplasticity, the concept could also be expanded to pediatric studies of motor system development, disorder, and resilience. For instance, HD-DOT could be used to track somatomotor cortex physiology as typically developing children acquire numerous gross and fine motor skills over their first few years of life, or as the brain develops an atypical functional architecture working around lesions caused by perinatal brain injury. All these proposals would invoke and further build on methodological advances pioneered in this dissertation, such as pediatric HD-DOT and wide-field HD-DOT including coverage of the dorsal head surface to capture somatomotor cortex physiology and distributed patterns of brain activity. In the long run, this work could greatly augment our understanding of when, how, and why the motor system or the nervous system in general does or does not display its remarkable plasticity.

5.2.3 Decoding applications beyond the laboratory

Previous chapters have already discussed at length how our decoding results relate to the breadth of neural decoding research conducted using other imaging modalities, and how several notable fMRI and electrocorticography decoding studies lay out a roadmap for potential future studies of decoding with HD-DOT [39,43,159]. Rather than restate all of the same points, this final section summarizes them briefly and ends by describing a specific example of how this work could be translated to have a significant clinical impact.

In essence, we have established the feasibility of reliably decoding stimulus information from HD-DOT measurements of brain activity. Ongoing and future work will attempt to utilize more elaborate and versatile decoding algorithms, based on parametric feature encoding models and machine learning approaches, to decode increasingly detailed information, novel stimuli, and covert signals such as imagined scenes and motor intention. Furthermore, decoding algorithms could be combined with real-time data processing and closed loop feedback to enable non-invasive brain-computer interfaces. These advances would greatly broaden the scope for applications like augmentative communication and prosthetic control in patients with severe motor disabilities, such as locked in syndrome following a stroke or amyotrophic lateral sclerosis. Furthermore, decoding of data acquired with non-invasive optical imaging methods could be applied more broadly beyond the clinic to the general population as well. Perhaps one day we will all interface to some extent with technology, social media, and each other using wearable optical neuroimaging caps to capture our brain activity and utilizing decoding algorithms to decipher it.

But for now, consider the compelling case of children with cerebral palsy (CP). CP is the most common motor disability in children [143]. It can be caused by genetic factors or various perinatal neurological insults, and results in permanent deficits in speech and movement with no known cure. The degree of functional motor impairment and level of associated cognitive disability can both vary greatly depending on the precise etiology and territory of the lesion. As a result, children with CP can have largely unimpaired cognition and emotional processing and yet be unable to communicate due to severe spasticity and paralysis. Effectively "locked in" and underappreciated by others who misunderstand their condition, these children suffer further mental health and social development consequences of their impaired communication [143]. The augmentative and alternative communication technologies currently available to improve quality of life mostly exploit residual motor function [213], which in severe CP can be subtle, slow, and unreliable. It is hard for most of us to imagine being a schoolchild, thinking and feeling the same sorts of things as all the other children, but having to rely entirely on something like eye movements in order to learn and play and communicate with others. Yet it is easy to appreciate the impact in this situation of an alternative means of communication with a higher bit rate. This is a vision that has inspired much of the work in this dissertation. We have developed HD-DOT for pediatric functional neuroimaging and evaluated decoding performance with HD-DOT data. Building on this research, ongoing developments like parametric model-based decoding and wireless DOT systems hold promise for high impact applications of decoding beyond the laboratory, such as conveying mental images or generating speech from brain activity sampled

with wearable imaging systems. These advances could one day give a new voice to children with cerebral palsy.

"There's a feeling you give me, an everglow."

- Coldplay

References

[1] M.E. Raichle, A Brief History of Human Functional Brain Mapping, Brain Mapp. Syst. (2000) 33–75. https://doi.org/10.1016/b978-012692545-6/50004-0.

[2] J.-D. Haynes, G. Rees, Decoding mental states from brain activity in humans, Nat. Rev. Neurosci. 7 (2006) 523–534. https://doi.org/10.1038/nrn1931.

[3] T. Naselaris, K.N. Kay, S. Nishimoto, J.L. Gallant, Encoding and decoding in fMRI, Neuroimage. 56 (2011) 400–410. https://doi.org/10.1016/j.neuroimage.2010.07.073.Encoding.

[4] M.N. Hebart, C.I. Baker, Deconstructing multivariate decoding for the study of brain function, Neuroimage. 180 (2018) 4–18. https://doi.org/10.1016/j.neuroimage.2017.08.005.

[5] M.S. Gazzaniga, R.B. Ivry, G.R. Mangun, Methods of Cognitive Neuroscience, in: Cogn. Neurosci. Biol. Mind, 2019: pp. 72–123.

[6] M.D. Wheelock, J.P. Culver, A.T. Eggebrecht, High-density diffuse optical tomography for imaging human brain function, Rev. Sci. Instrum. 90 (2019). https://doi.org/10.1063/1.5086809.

[7] T.K. Hensch, Critical period regulation, Annu. Rev. Neurosci. 27 (2004) 549–579. https://doi.org/10.1146/annurev.neuro.27.070203.144327.

[8] C.M. Alberini, A. Travaglia, Infantile amnesia: A critical period of learning to learn and remember, J. Neurosci. 37 (2017) 5783–5795. https://doi.org/10.1523/JNEUROSCI.0324-17.2017.

[9] N. Friedmann, D. Rusou, Critical period for first language: The crucial role of language input during the first year of life, Curr. Opin. Neurobiol. 35 (2015) 27–34. https://doi.org/10.1016/j.conb.2015.06.003.

[10] M. Fagiolini, J.J. Leblanc, Autism: A critical period disorder?, Neural Plast. 2011 (2011). https://doi.org/10.1155/2011/921680.

[11] C.A.I. Nelson, C.H. Zeanah, N. a Fox, P.J. Marshall, A.T. Smyke, D. Guthrie, Cognitive Recovery in Socially Deprived Young Children: The Bucharest Early Intervention Project, Science (80-.). 318 (2007) 1937–1940.

[12] M.S. Banks, R.N. Aslin, R.D. Letson, Sensitive period for the development of human binocular vision, Science (80-.). 190 (1975) 675 LP – 677.

https://doi.org/10.1126/science.1188363.

[13] T.N. Wiesel, D.H. Hubel, Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens, J Neurophysiol. 28 (1965) 1029–1040. https://doi.org/10.1152/jn.1965.28.6.1029.

[14] T.N. Wiesel, D.H. Hubel, Extent of recovery from the effects of visual deprivation in kittens, J Neurophysiol. 28 (1965) 1060–1072. https://doi.org/10.1152/jn.1965.28.6.1060.

[15] T.N. Wiesel, D.H. Hubel, Single-Cell Responses in Striate Cortex of Kittens Deprived of Vision in One Eye., J Neurophysiol. 26 (1963) 1003–1017. https://doi.org/10.1152/jn.1963.26.6.1003.

[16] R.K. Reh, B.G. Dias, C.A. Nelson, D. Kaufer, J.F. Werker, B. Kolbh, J.D. Levine, T.K. Hensch, Critical period regulation acrossmultiple timescales, Proc. Natl. Acad. Sci. U. S. A. 117 (2020) 23242–23251. https://doi.org/10.1073/pnas.1820836117.

[17] M. Pagel, Q&A: What is human language, when did it evolve and why should we care?, BMC Biol. 15 (2017) 1–6. https://doi.org/10.1186/s12915-017-0405-3.

[18] J. Gervain, Typical language development, 1st ed., Elsevier B.V., 2020. https://doi.org/10.1016/B978-0-444-64150-2.00016-2.

[19] M.I. Visser-Bochane, S.A. Reijneveld, W.P. Krijnen, C.P. van der Schans, M.R. Luinge, Identifying Milestones in Language Development for Young Children Ages 1 to 6 Years, Acad. Pediatr. 20 (2020) 421–429. https://doi.org/10.1016/j.acap.2019.07.003.

[20] L.J. Weiss-Croft, T. Baldeweg, Maturation of language networks in children: A systematic review of 22years of functional MRI, Neuroimage. 123 (2015) 269–281. https://doi.org/10.1016/j.neuroimage.2015.07.046.

[21] E. Redcay, F. Haist, E. Courchesne, Functional neuroimaging of speech perception during a pivotal period in language acquisition, Dev. Sci. 11 (2008) 237–252. https://doi.org/10.1111/j.1467-7687.2008.00674.x.

[22] N. Raschle, J. Zuk, S. Ortiz-Mantilla, D.D. Sliva, A. Franceschi, P.E. Grant, A.A. Benasich, N. Gaab, Pediatric neuroimaging in early childhood and infancy: challenges and practical guidelines, Ann N Y Acad Sci. 1252 (2012) 43–50.

[23] C.D. Smyser, A.Z. Snyder, J.J. Neil, Functional connectivity MRI in infants: Exploration of the functional organization of the developing brain, Neuroimage. 56 (2011) 1437–1452. https://doi.org/10.1016/j.neuroimage.2011.02.073.

[24] B.R. Howell, M.A. Styner, W. Gao, P.T. Yap, L. Wang, K. Baluyot, E. Yacoub, G. Chen, T. Potts, A. Salzwedel, G. Li, J.H. Gilmore, J. Piven, J.K. Smith, D. Shen, K. Ugurbil, H. Zhu, W. Lin, J.T. Elison, The UNC/UMN Baby Connectome Project (BCP): An overview

of the study design and protocol development, Neuroimage. 185 (2019) 891–905. https://doi.org/10.1016/j.neuroimage.2018.03.049.

[25] H. Zhang, D. Shen, W. Lin, Resting-state functional MRI studies on infant brains: A decade of gap-filling efforts, Neuroimage. 185 (2019) 664–684. https://doi.org/10.1016/j.neuroimage.2018.07.004.

[26] L.M. Brier, E.C. Landsness, A.Z. Snyder, P.W. Wright, G.A. Baxter, A.Q. Bauer, J.-M. Lee, J.P. Culver, Separability of calcium slow waves and functional connectivity during wake, sleep, and anesthesia, Neurophotonics. 6 (2019) 1. https://doi.org/10.1117/1.nph.6.3.035002.

[27] A. Mitra, A.Z. Snyder, E. Tagliazucchi, H. Laufs, J. Elison, R.W. Emerson, M.D. Shen, J.J. Wolff, K.N. Botteron, S. Dager, A.M. Estes, A.C. Evans, G. Gerig, H.C. Hazlett, S.J. Paterson, R.T. Schultz, M.A. Styner, L. Zwaigenbaum, C. Chappell, A. Estes, D. Shaw, K. Botteron, R. McKinstry, J. Constantino, J. Pruett, R. Schultz, S. Paterson, D.L. Collins, G.B. Pike, V. Fonov, P. Kostopoulos, S. Das, M. Styner, H. Gu, B.L. Schlaggar, J. Piven, J.R. Pruett, M. Raichle, Resting-state fMRI in sleeping infants more closely resembles adult sleep than adult wakefulness, PLoS One. 12 (2017) 1–19. https://doi.org/10.1371/journal.pone.0188122.

[28] T. Vanderwal, J. Eilbott, F.X. Castellanos, Movies in the magnet: Naturalistic paradigms in developmental functional neuroimaging, Dev. Cogn. Neurosci. 36 (2019) 100600. https://doi.org/10.1016/j.dcn.2018.10.004.

[29] P. Pinti, I. Tachtsidis, A. Hamilton, J. Hirsch, C. Aichelburg, S. Gilbert, P.W. Burgess, The present and future use of functional near-infrared spectroscopy (fNIRS) for cognitive neuroscience, Ann N Y Acad Sci. 1464 (2020) 5–29. https://doi.org/10.1111/nyas.13948.

[30] B.R. White, J.P. Culver, Quantitative evaluation of high-density diffuse optical tomography: in vivo resolution and mapping performance, J. Biomed. Opt. 15 (2010) 026006. https://doi.org/10.1117/1.3368999.

[31] R. Saager, A. Berger, Measurement of layer-like hemodynamic trends in scalp and cortex: implications for physiological baseline suppression in functional near-infrared spectroscopy, J. Biomed. Opt. 13 (2008) 034017. https://doi.org/10.1117/1.2940587.

[32] S. Lloyd-Fox, K. Begus, D. Halliday, L. Pirazzoli, A. Blasi, M. Papademetriou, M.K. Darboe, A.M. Prentice, M.H. Johnson, S.E. Moore, C.E. Elwell, Cortical specialisation to social stimuli from the first days to the second year of life: A rural Gambian cohort, Dev. Cogn. Neurosci. 25 (2017) 92–104. https://doi.org/10.1016/j.dcn.2016.11.005.

[33] S. Lloyd-Fox, A. Blasi, C.E. Elwell, T. Charman, D. Murphy, M.H. Johnson, Reduced neural sensitivity to social stimuli in infants at risk for autism, Proc. R. Soc. B Biol. Sci. 283 (2013). https://doi.org/10.1098/rspb.2016.0260.

[34] D.J. Greene, J.M. Koller, J.M. Hampton, V. Wesevich, A.N. Van, A.L. Nguyen, C.R. Hoyt, L. Mcintyre, E.A. Earl, R.L. Klein, J.S. Shimony, S.E. Petersen, B.L. Schlaggar, D.A. Fair, N.U.F. Dosenbach, Behavioral interventions for reducing head motion during MRI scans in children, Neuroimage. 171 (2018) 234–245. https://doi.org/10.1016/j.neuroimage.2018.01.023.

[35] U. Hasson, Y. Nir, I. Levy, G. Fuhrmann, R. Malach, Intersubject Synchronization of Cortical Activity During Natural Vision, 303 (2004) 1634–1640.

[36] A. Bartels, S. Zeki, Functional Brain Mapping During Free Viewing of Natural Scenes, 5 (2004) 75–85. https://doi.org/10.1002/hbm.10153.

[37] Y. Golland, S. Bentin, H. Gelbard, Y. Benjamini, R. Heller, Y. Nir, U. Hasson, R. Malach, Extrinsic and intrinsic systems in the posterior cortex of the human brain revealed during natural sensory stimulation, Cereb. Cortex. 17 (2007) 766–777. https://doi.org/10.1093/cercor/bhk030.

[38] K.N. Kay, T. Naselaris, R.J. Prenger, J.L. Gallant, Identifying natural images from human brain activity, Nature. 452 (2008) 352–5. https://doi.org/10.1038/nature06713.

[39] S. Nishimoto, A.T. Vu, T. Naselaris, Y. Benjamini, B. Yu, J.L. Gallant, Reconstructing Visual Experiences from Brain Activity Evoked by Natural Movies, Curr. Biol. 21 (2011) 1641–1646. https://doi.org/10.1016/j.cub.2011.08.031.

[40] H. Wen, J. Shi, Y. Zhang, K.-H. Lu, J. Cao, Z. Liu, Neural Encoding and Decoding with Deep Learning for Dynamic Natural Vision, Cereb. Cortex. 28 (2018) 4136–4160. https://doi.org/10.1093/cercor/bhx268.

[41] T. Horikawa, Y. Kamitani, Generic decoding of seen and imagined objects using hierarchical visual features, Nat. Commun. 8 (2017) 15037. https://doi.org/10.1038/ncomms15037.

[42] C. Herff, L. Diener, M. Angrick, E. Mugler, M.C. Tate, M.A. Goldrick, D.J. Krusienski, M.W. Slutzky, T. Schultz, Generating Natural, Intelligible Speech From Brain Activity in Motor, Premotor, and Inferior Frontal Cortices, Front. Neurosci. 13 (2019) 1–11. https://doi.org/10.3389/fnins.2019.01267.

[43] G.K. Anumanchipalli, J. Chartier, E.F. Chang, Speech synthesis from neural decoding of spoken sentences, Nature. 568 (2019) 493–498. https://doi.org/10.1038/s41586-019-1119-1.

[44] A. Stecco, A. Saponaro, A. Carriero, Patient safety issues in magnetic resonance imaging: state of the art, Radiol. Medica. 112 (2007) 491–508. https://doi.org/10.1007/s11547-007-0154-4.

[45] C. Vogt, Design Concepts and Validation of In-field MRI Electronics, (2017) 24767.

[46] A. Abdalmalak, D. Milej, L. Norton, D.B. Debicki, T. Gofton, M. Diop, A.M. Owen, K. St. Lawrence, Single-session communication with a locked-in patient by functional near-infrared spectroscopy, Neurophotonics. 4 (2017) 040501. https://doi.org/10.1117/1.nph.4.4.040501.

[47] S.M.H. Hosseini, Y. Mano, M. Rostami, M. Takahashi, M. Sugiara, R. Kawashima, Decoding what one likes or dislikes from single-trial fNIRS measurements, Neuroreport. 22 (2011) 269–273. https://doi.org/10.1097/WNR.0b013e3283451f8f.

[48] L.L. Emberson, B.D. Zinszer, R.D.S. Raizada, R.N. Aslin, Decoding the infant mind: Multivariate pattern analysis (MVPA) using fNIRS, PLoS One. 12 (2017) e0172500. https://doi.org/10.1371/journal.pone.0172500.

[49] A. Abdalmalak, D. Milej, L.C.M. Yip, A.R. Khan, M. Diop, A.M. Owen, K. St. Lawrence, Assessing Time-Resolved fNIRS for Brain-Computer Interface Applications of Mental Communication, Front. Neurosci. 14 (2020) 105. https://doi.org/10.3389/fnins.2020.00105.

[50] S. Luu, T. Chau, Decoding subjective preference from single-trial near-infrared spectroscopy signals, J Neural Eng. 6 (2009) 016003. https://doi.org/10.1088/1741-2560/6/1/016003.

[51] R. Sitaram, H. Zhang, C. Guan, M. Thulasidas, Y. Hoshi, A. Ishikawa, K. Shimizu, N. Birbaumer, Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain–computer interface, Neuroimage. 34 (2007) 1416–1427. https://doi.org/10.1016/j.neuroimage.2006.11.005.

[52] B.W. Zeff, B.R. White, H. Dehghani, B.L. Schlaggar, J.P. Culver, Retinotopic mapping of adult human visual cortex with high-density diffuse optical tomography, Proc Natl Acad Sci U S A. 104 (2007) 12169–12174. https://doi.org/10.1073/pnas.0611266104.

[53] A.T. Eggebrecht, S.L. Ferradal, A. Robichaux-Viehoever, S.S. Mahlega, H. Dehghani, A.Z. Snyder, T. Hershey, J.P. Culver, Mapping distributed brain function and networks with diffuse optical tomography, Nat. Photonics. 8 (2014) 448–454. https://doi.org/10.1038/nphoton.2014.107.

[54] S. Lloyd-Fox, A. Blasi, C.E. Elwell, Illuminating the developing brain: The past, present and future of functional near infrared spectroscopy, Neurosci. Biobehav. Rev. 34 (2010) 269–284. https://doi.org/10.1016/j.neubiorev.2009.07.008.

[55] F. Tian, G. Alexandrakis, H. Liu, Optimization of probe geometry for diffuse optical brain imaging based on measurement density and distribution, Appl. Opt. 48 (2009) 2496–2504. https://doi.org/10.1364/AO.48.002496.

[56] E.M. Frijia, A. Billing, S. Lloyd-Fox, E. Vidal Rosas, L. Collins-Jones, M.M. Crespo-Llado, M.P. Amadó, T. Austin, A. Edwards, L. Dunne, G. Smith, R. Nixon-Hill, S. Powell, N.L. Everdell, R.J. Cooper, Functional imaging of the developing brain with wearable high-density diffuse optical tomography: A new benchmark for infant neuroimaging outside the scanner environment, Neuroimage. 225 (2021) 117490. https://doi.org/10.1016/j.neuroimage.2020.117490.

[57] H. Dehghani, B.R. White, B.W. Zeff, J.P. Culver, Depth sensitivity analysis of high-density imaging arrays for mapping brain function with diffuse optical tomography, Biomed. Opt. BIOMED 2008. 48 (2009) 137–143. https://doi.org/10.1364/biomed.2008.bmd27.

[58] N.M. Gregg, B.R. White, B.W. Zeff, A.J. Berger, J.P. Culver, Brain specificity of diffuse optical imaging: improvements from superficial signal regression and tomography, Front. Neuroenergetics. 2 (2010) 14. https://doi.org/10.3389/fnene.2010.00014.

[59] H. Dehghani, B.W. Pogue, S.P. Poplack, K.D. Paulsen, Multiwavelength three-dimensional near-infrared tomography of the breast: initial simulation, phantom, and clinical results, Appl. Opt. 42 (2003) 135–45. https://doi.org/10.1364/ao.42.000135.

[60] A.T. Eggebrecht, B.R. White, S.L. Ferradal, C. Chen, Y. Zhan, A.Z. Snyder, H. Dehghani, J.P. Culver, A quantitative spatial comparison of high-density diffuse optical tomography and fMRI cortical mapping, Neuroimage. 61 (2012) 1120–1128. https://doi.org/10.1016/j.neuroimage.2012.01.124.

[61] A.K. Fishell, A.M. Arbeláez, C.P. Valdés, T.M. Burns-Yocum, A. Sherafati, E.J. Richter, M. Torres, A.T. Eggebrecht, C.D. Smyser, J.P. Culver, Portable, field-based neuroimaging using high-density diffuse optical tomography, Neuroimage. 215 (2020) 116541. https://doi.org/10.1016/j.neuroimage.2020.116541.

[62] J.P. Culver, K.M. Bergonzi, A.E. Eggebrecht, A.K. Fishell, J. Lee, Optical imaging of functional connectivity at the bedside, Conf Proc IEEE Eng Med Biol Soc. (2016) 65–67. https://doi.org/10.1109/EMBC.2016.7590641.

[63] B. Burke, K. Bergonzi, A. Sherafati, A. Fishell, T. Burns-yocum, S. Ferradal, B. Palanca, R. Dhar, G. Kumar, J. Lee, J. Culver, A. Eggebrecht, Bedside Diffuse Optical Tomography of Disrupted Brain Connectivity During Acute Stroke, Prepr. (Version 1) Available Res. Sq. (2020) 1–20. https://doi.org/10.21203/rs.3.rs-107369/v1.

[64] A.K. Fishell, T.M. Burns-Yocum, K.M. Bergonzi, A.T. Eggebrecht, J.P. Culver, Mapping brain function during naturalistic viewing using high-density diffuse optical tomography, Sci. Rep. 9 (2019) 11115. https://doi.org/10.1038/s41598-019-45555-8.

[65] J.L. Collinger, B. Wodlinger, J.E. Downey, W. Wang, E.C. Tyler-kabara, D.J. Weber, A.J.C. Mcmorland, M. Velliste, M.L. Boninger, A.B. Schwartz, High-performance

neuroprosthetic control by an individual with tetraplegia, Lancet. 381 (2013) 557–564. https://doi.org/10.1016/S0140-6736(12)61816-9.

[66] A.M. Albers, P. Kok, I. Toni, H.C. Dijkerman, F.P. de Lange, Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex, Curr. Biol. 23 (2013) 1427–1431. https://doi.org/10.1016/j.cub.2013.05.065.

[67] S.A. Harrison, F. Tong, Decoding reveals the contents of visual working memory in early visual areas, Nature. 458 (2009) 632–635. https://doi.org/10.1038/nature07832.

[68] T. Horikawa, M. Tamaki, Y. Miyawaki, Y. Kamitani, Neural Decoding of Visual Imagery During Sleep, Science (80-.). 340 (2013) 639–643. https://doi.org/10.1126/science.1234330.

[69] B.R. White, J.P. Culver, Phase-encoded retinotopy as an evaluation of diffuse optical neuroimaging, Neuroimage. 49 (2010) 568–577. https://doi.org/10.1016/j.neuroimage.2009.07.023.

[70] D.M. Brandman, S.S. Cash, L.R. Hochberg, Review: Human intracortical recording and neural decoding for brain computer interfaces, IEEE Trans Neural Syst Rehabil Eng. 25 (2018) 1687–1696. https://doi.org/10.1109/TNSRE.

[71] J. Correia, E. Formisano, G. Valente, L. Hausfeld, B. Jansma, M. Bonte, Brain-Based Translation: fMRI Decoding of Spoken Words in Bilinguals Reveals Language-Independent Semantic Representations in Anterior Temporal Lobe, J Neurosci. 34 (2014) 332–338. https://doi.org/10.1523/JNEUROSCI.1302-13.2014.

[72] J.M. Correia, B.M.B. Jansma, M. Bonte, Decoding Articulatory Features from fMRI Responses in Dorsal Speech Regions, J Neurosci. 35 (2015) 15015–15025. https://doi.org/10.1523/JNEUROSCI.0977-15.2015.

[73] S. Hoefle, A. Engel, R. Basilio, V. Alluri, P. Toiviainen, M. Cagy, J. Moll, Identifying musical pieces from fMRI data using encoding and decoding models, Sci. Rep. 8 (2018) 2266. https://doi.org/10.1038/s41598-018-20732-3.

[74] A.G. Huth, T. Lee, S. Nishimoto, N.Y. Bilenko, A.T. Vu, J.L. Gallant, Decoding the Semantic Content of Natural Movies from Human Brain Activity, Front. Syst. Neurosci. 10 (2016) 81. https://doi.org/10.3389/fnsys.2016.00081.

[75] M.S. Hassanpour, A.T. Eggebrecht, J.P. Culver, J.E. Peelle, Mapping cortical responses to speech using high-density diffuse optical tomography, Neuroimage. 117 (2015) 319–326. https://doi.org/10.1016/j.neuroimage.2015.05.058.

[76] D.H. Hubel, T.N. Wiesel, Receptive Fields, Binocular Interaction and Functional Architecture in the Cat’s Visual Cortex, J Physiol. 160 (1962) 106–154.

https://doi.org/10.1113/jphysiol.1962.sp006837.

[77] B. Thirion, E. Duchesnay, E. Hubbard, J. Dubois, J. Poline, D. Lebihan, S. Dehaene, Inverse retinotopy : Inferring the visual content of images from brain activation patterns, Neuroimage. 33 (2006) 1104–1116. https://doi.org/10.1016/j.neuroimage.2006.06.062.

[78] Y. Kamitani, F. Tong, Decoding the visual and subjective contents of the human brain, Nat. Neurosci. 8 (2005) 679–685. https://doi.org/10.1038/nn1444.

[79] V. Fonov, A.C. Evans, K. Botteron, C.R. Almli, R.C. McKinstry, D.L. Collins, Unbiased average age-appropriate atlases for pediatric studies, Neuroimage. 54 (2011) 313–327. https://doi.org/10.1016/j.neuroimage.2010.07.033.

[80] J. Mazziotta, A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, R. Kawashima, B. Mazoyer, A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM), Philos. Trans. R. Soc. B Biol. Sci. 356 (2001) 1293–1322. https://doi.org/10.1098/rstb.2001.0915.

[81] S.L. Ferradal, A.T. Eggebrecht, M. Hassanpour, A.Z. Snyder, J.P. Culver, Atlas-based head modeling and spatial normalization for high-density diffuse optical tomography: In vivo validation against fMRI, Neuroimage. 85 (2014) 117–126. https://doi.org/10.1016/j.neuroimage.2013.03.069.

[82] A.M. Dale, B. Fischl, M.I. Sereno, Cortical Surface-Based Analysis, Neuroimage. 9 (1999) 179–194. https://doi.org/10.1006/nimg.1998.0395.

[83] B. Fischl, A. Liu, A.M. Dale, Automated Manifold Surgery: Constructing Geometrically Accurate and Topologically Correct Models of the Human Cerebral Cortex, IEEE Trans. Med. Imaging. 20 (2001) 70–80. https://doi.org/10.1109/42.906426.

[84] F. Ségonne, A.M. Dale, E. Busa, M. Glessner, D. Salat, H.K. Hahn, B. Fischl, A hybrid approach to the skull stripping problem in MRI, Neuroimage. 22 (2004) 1060–1075. https://doi.org/10.1016/j.neuroimage.2004.03.032.

[85] H. Dehghani, M.E. Eames, P.K. Yalavarthy, S.C. Davis, S. Srinivasan, C.M. Carpenter, B.W. Pogue, K.D. Paulsen, Near infrared optical tomography using NIRFAST: Algorithm for numerical model and image Reconstruction, Commun. Numer. Methods Eng. 25 (2008) 711–732. https://doi.org/10.1002/cnm.1162.

[86] F. Bevilacqua, D. Piguet, P. Marquet, J.D. Gross, B.J. Tromberg, C. Depeursinge, In vivo local determination of tissue optical properties: applications to human brain, Appl. Opt. 38 (1999) 4939. https://doi.org/10.1364/ao.38.004939.

[87] G. Strangman, J.P. Culver, J.H. Thompson, D.A. Boas, A Quantitative Comparison of Simultaneous BOLD fMRI and NIRS Recordings during Functional Brain Activation, Neuroimage. 17 (2002) 719–731. https://doi.org/10.1006/nimg.2002.1227.

[88] A. Custo, W.M. Wells, A.H. Barnett, E.M.C. Hillman, D.A. Boas, Effective scattering coefficient of the cerebral spinal fluid in adult head models for diffuse optical imaging, Appl. Opt. 45 (2006) 4747–4755. https://doi.org/10.1364/AO.45.004747.

[89] E.M. Gordon, T.O. Laumann, A.W. Gilmore, D.J. Newbold, D.J. Greene, J.J. Berg, M. Ortega, C. Hoyt-Draze, C. Gratton, H. Sun, J.M. Hampton, R.S. Coalson, A.L. Nguyen, K.B. McDermott, J.S. Shimony, A.Z. Snyder, B.L. Schlaggar, S.E. Petersen, S.M. Nelson, N.U.F. Dosenbach, Precision functional mapping of individual human brains, Neuron. 95 (2017) 791–807. https://doi.org/10.1016/j.neuron.2017.07.011.

[90] S.O. Dumoulin, B.A. Wandell, Population receptive field estimates in human visual cortex, Neuroimage. 39 (2008) 647–660. https://doi.org/10.1016/j.neuroimage.2007.09.034.

[91] D.H. Brainard, The Psychophysics Toolbox, Spat. Vis. 10 (1997) 433–436.

[92] G.M. Boynton, S.A. Engel, G.H. Glover, D.J. Heeger, Linear systems analysis of functional magnetic resonance imaging in human V1, J. Neurosci. 16 (1996) 4207–4221. https://doi.org/10.1523/jneurosci.16-13-04207.1996.

[93] M.A. Lindquist, J. Meng Loh, L.Y. Atlas, T.D. Wager, Modeling the hemodynamic response function in fMRI: efficiency, bias and mis-modeling., Neuroimage. 45 (2009) S187–S198. https://doi.org/10.1016/j.neuroimage.2008.10.065.

[94] C.E. Metz, Basic Principles of ROC Analysis, Semin. Nucl. Med. 8 (1978) 283–98. https://doi.org/10.1016/s0001-2998(78)80014-2.

[95] C.E. Metz, Receiver Operating Characteristic Analysis: A Tool for the Quantitative Evaluation of Observer Performance and Imaging Systems, J Am Coll Radiol. 3 (2006) 413–422. https://doi.org/10.1016/j.jacr.2006.02.021.

[96] W.J. Youden, Index for rating diagnostic tests, Cancer. 3 (1950) 32–35. https://doi.org/10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3.

[97] D.H. Hubel, T.N. Wiesel, Binocular interaction reared in striate artificial cortex squint, J. Neurophysiol. 28 (1965) 1041–1059. https://doi.org/10.1152/jn.1965.28.6.1041.

[98] Y. Kamitani, F. Tong, Decoding Seen and Attended Motion Directions from Activity in the Human Visual Cortex, Curr. Biol. 16 (2006) 1096–1102. https://doi.org/10.1016/j.cub.2006.04.003.

[99] T.O. Laumann, E.M. Gordon, B. Adeyemo, A.Z. Snyder, S.J. Joo, M.-Y. Chen, A.W.

Gilmore, K.B. Mcdermott, S.M. Nelson, N.U.F. Dosenbach, B.L. Schlaggar, J.A. Mumford, R.A. Poldrack, S.E. Petersen, Functional system and areal organization of a highly sampled individual human brain, Neuron. 87 (2015) 657–670. https://doi.org/10.1016/j.neuron.2015.06.037.

[100] E.A. DeYoe, G.J. Carman, P. Bandettini, S. Glickman, J. Wieser, R. Cox, D. Miller, J. Neitz, Mapping striate and extrastriate visual areas in human cerebral cortex, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 2382–2386. https://doi.org/10.1073/pnas.93.6.2382.

[101] B.A. Wandell, S.O. Dumoulin, A.A. Brewer, Visual Field Maps in Human Cortex, Neuron. 56 (2007) 366–383. https://doi.org/10.1016/j.neuron.2007.10.012.

[102] P. Kochunov, M. Hasnain, J. Lancaster, T. Grabowski, P. Fox, Improvement in Variability of the Horizontal Meridian of the Primary Visual Area Following High-Resolution Spatial Normalization, Hum. Brain Mapp. 18 (2003) 123–134. https://doi.org/10.1002/hbm.10080.

[103] P.M. Thompson, C. Schwartz, R.T. Lin, A.A. Khan, A.W. Toga, Three-dimensional Statistical Analysis of Sulcal Variability in the Human Brain, J Neurosci. 16 (1996) 4261–4274. https://doi.org/10.1523/JNEUROSCI.16-13-04261.1996.

[104] G. Iaria, M. Petrides, Occipital Sulci of the Human Brain: Variability and Probability Maps, J Comp Neurol. 501 (2007) 243–259. https://doi.org/10.1002/cne.

[105] R.F. Dougherty, V.M. Koch, A.A. Brewer, B. Fischer, J. Modersitzki, B.A. Wandell, Visual field representations and locations of visual areas V1/2/3 in human visual cortex, J. Vis. 3 (2003) 586–598. https://doi.org/10.1167/3.10.1.

[106] K. Amunts, A. Malikovic, H. Mohlberg, T. Schormann, K. Zilles, Brodmann's Areas 17 and 18 Brought into Stereotaxic Space- Where and How Variable?, Neuroimage. 11 (2000) 66–84. https://doi.org/10.1006/nimg.1999.0516.

[107] H. Zhao, R.J. Cooper, Review of recent progress toward a fiberless, whole-scalp diffuse optical tomography system, Neurophotonics. 5 (2017) 011012. https://doi.org/10.1117/1.NPh.5.1.011012.

[108] D. Chitnis, R.J. Cooper, L. Dempsey, S. Powell, S. Quaggia, D. Highton, C. Elwell, J.C. Hebden, N.L. Everdell, Functional imaging of the human brain using a modular, fibre-less, high-density diffuse optical tomography system, Biomed. Opt. Express. 7 (2016) 4275. https://doi.org/10.1364/boe.7.004275.

[109] A. Sherafati, A.Z. Snyder, A.T. Eggebrecht, K.M. Bergonzi, T.M. Burns-Yocum, H.M. Lugar, S.L. Ferradal, A. Robichaux-Viehoever, C.D. Smyser, B.J. Palanca, T. Hershey, J.P. Culver, Global motion detection and censoring in high-density diffuse optical tomography, Hum. Brain Mapp. (2020) 4093–4112. https://doi.org/10.1002/hbm.25111.

[110] E. Guzman-Martinez, P. Leung, S. Franconeri, M. Grabowecky, S. Suzuki, Rapid eye-fixation training without eye tracking, Psychon Bull Rev. 16 (2009) 491–496. https://doi.org/10.3758/PBR.16.3.491.

[111] S. Yang, T.J. Ross, Y. Zhang, E.A. Stein, Y. Yang, Head motion suppression using real-time feedback of motion information and its effects on task performance in fMRI, Neuroimage. 27 (2005) 153–162. https://doi.org/10.1016/j.neuroimage.2005.02.050.

[112] W. Penfield, E. Boldrey, Somatic Motor And Sensory Representation In The Cerebral Cortex Of Man As Studied By Electrical Stimulation, Brain. 60 (1937) 389–443. https://doi.org/10.1093/brain/60.4.389.

[113] J. Besle, O. Mougin, R.-M. Sánchez-Panchuelo, C. Lanting, P. Gowland, R. Bowtell, S. Francis, K. Krumbholz, Is Human Auditory Cortex Organization Compatible With the Monkey Model ? Contrary Evidence From Ultra-High-Field Functional and Structural MRI, Cereb. Cortex. 29 (2019) 410–428. https://doi.org/10.1093/cercor/bhy267.

[114] S.L. Ferradal, S.M. Liao, A.T. Eggebrecht, J.S. Shimony, T.E. Inder, J.P. Culver, C.D. Smyser, Functional Imaging of the Developing Brain at the Bedside Using Diffuse Optical Tomography, Cereb. Cortex. 26 (2016) 1558–1568. https://doi.org/10.1093/cercor/bhu320.

[115] S.M. Liao, S.L. Ferradal, B.R. White, N. Gregg, T.E. Inder, J.P. Culver, High-density diffuse optical tomography of term infant visual cortex in the nursery, J. Biomed. Opt. 17 (2012) 081414. https://doi.org/10.1117/1.JBO.17.8.081414.

[116] K.M. Bergonzi, T.M. Burns-Yocum, J.R. Bumstead, E.M. Buckley, P.C. Mannion, C.H. Tracy, E. Mennerick, S.L. Ferradal, H. Dehghani, A.T. Eggebrecht, J.P. Culver, Lightweight sCMOS-based high-density diffuse optical tomography, Neurophotonics. 5 (2018) 035006. https://doi.org/10.1117/1.NPh.5.3.035006.

[117] J.J. Vidal, Realtime Detection of Brain Events in EEG, Proc. IEEE. 65 (1977) 633–641. https://doi.org/10.1109/PROC.1977.10542.

[118] R.M. Rahn, L.M. Brier, A.R. Bice, M.D. Reisman, J.D. Dougherty, J.P. Culver, Functional connectivity of the developing mouse cortex, Biorxiv. 2 (2021).

[119] K. Tripathy, Z.E. Markow, A.K. Fishell, A. Sherafati, T.M. Burns-Yocum, M.L. Schroeder, A.M. Svoboda, A.T. Eggebrecht, M.A. Anastasio, B.L. Schlaggar, J.P. Culver, Decoding visual information from high-density diffuse optical tomography neuroimaging data, Neuroimage. 226 (2021) 117516. https://doi.org/10.1016/j.neuroimage.2020.117516.

[120] M.S. Hassanpour, B.R. White, A.T. Eggebrecht, S.L. Ferradal, A.Z. Snyder, J.P. Culver, Statistical analysis of high density diffuse optical tomography, Neuroimage. 85 (2014) 104–116. https://doi.org/10.1016/j.neuroimage.2013.05.105.Statistical.

[121] G. Dehaene-lambertz, S. Dehaene, L. Hertz-pannier, Functional Neuroimaging of Speech Perception in Infants, 298 (2002) 2013–2016.

[122] B. Deen, H. Richardson, D.D. Dilks, A. Takahashi, B. Keil, L.L. Wald, N. Kanwisher, R. Saxe, Organization of high-level visual cortex in human infants, Nat. Commun. 8 (2017) 1–10. https://doi.org/10.1038/ncomms13995.

[123] S.Z. Dong, M. Zhu, D. Bulas, Techniques for minimizing sedation in pediatric MRI, J. Magn. Reson. Imaging. 50 (2019) 1047–1054. https://doi.org/10.1002/jmri.26703.

[124] W.L. Gross, K.K. Lauer, X. Liu, S. Liu, S. Gollapudy, J.R. Binder, S. Li, A.G. Hudetz, Propofol sedation alters perceptual and cognitive functions in healthy volunteers as revealed by functional magnetic resonance imaging, 131 (2019) 254–265. https://doi.org/10.1097/ALN.0000000000002669.Propofol.

[125] S. Rothman, A. Gonen, A. Vodonos, V. Novack, I. Shelef, Does preparation of children before MRI reduce the need for anesthesia? Prospective randomized control trial, Pediatr. Radiol. 46 (2016) 1599–1605. https://doi.org/10.1007/s00247-016-3651-6.

[126] H.M.A. De Bie, M. Boersma, M.P. Wattjes, S. Adriaanse, R.J. Vermeulen, K.J. Oostrom, J. Huisman, D.J. Veltman, H.A. Delemarre-Van De Waal, Preparing children with a mock scanner training protocol results in high quality structural and functional MRI scans, Eur. J. Pediatr. 169 (2010) 1079–1085. https://doi.org/10.1007/s00431-010-1181-z.

[127] L.K. Butler, S. Kiran, H. Tager-Flusberg, Functional Near-Infrared Spectroscopy in the Study of Speech and Language Impairment Across the Life Span: A Systematic Review, Am J Speech Lang Pathol. 29 (2020) 1674–1701.

[128] A. Czeszumski, S. Eustergerling, A. Lang, D. Menrath, M. Gerstenberger, S. Schuberth, F. Schreiber, Z.Z. Rendon, P. König, Hyperscanning: A Valid Method to Study Neural Inter-brain Underpinnings of Social Interaction, Front. Hum. Neurosci. 14 (2020) 1–17. https://doi.org/10.3389/fnhum.2020.00039.

[129] R. Ciric, D.H. Wolf, J.D. Power, D.R. Roalf, G.L. Baum, K. Ruparel, R.T. Shinohara, M.A. Elliott, S.B. Eickho, C. Davatzikos, R.C. Gur, R.E. Gur, D.S. Bassett, Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity, 154 (2017) 174–187. https://doi.org/10.1016/j.neuroimage.2017.03.020.

[130] J.D. Power, K.A. Barnes, A.Z. Snyder, B.L. Schlaggar, S.E. Petersen, Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion, Neuroimage. 59 (2012) 2142–2154. https://doi.org/10.1016/j.neuroimage.2011.10.018.

[131] J.D. Power, B.L. Schlaggar, S.E. Petersen, Recent progress and outstanding issues in

motion correction in resting state fMRI, Neuroimage. 105 (2015) 536–551. https://doi.org/10.1016/j.neuroimage.2014.10.044.

[132] U. Hasson, R. Malach, D.J. Heeger, Reliability of cortical activity during natural stimulation, Trends Cogn. Sci. 14 (2010) 40–48. https://doi.org/10.1016/j.tics.2009.10.011.

[133] G.A. Calvert, E.T. Bullmore, M.J. Brammer, R. Campbell, S.C.R. Williams, P.K. McGuire, P.W.R. Woodruff, S.D. Iversen, A.S. David, Activation of auditory cortex during silent lipreading, Science (80-.). 276 (1997) 593–596. https://doi.org/10.1126/science.276.5312.593.

[134] D.Y. Tsao, M.S. Livingstone, Mechanisms of face perception, Annu. Rev. Neurosci. 31 (2008) 411–437. https://doi.org/10.1146/annurev.neuro.30.051606.094238.

[135] R. Le Grand, C.J. Mondloch, D. Maurer, H.P. Brent, Expert face processing requires visual input to the right hemisphere during infancy, Nat. Neurosci. 6 (2003) 1108–1112. https://doi.org/10.1038/nn1121.

[136] T. Arichi, G. Fagiolo, M. Varela, A. Melendez-Calderon, A. Allievi, N. Merchant, N. Tusor, S.J. Counsell, E. Burdet, C.F. Beckmann, A.D. Edwards, Development of BOLD signal hemodynamic responses in the human brain, Neuroimage. 63 (2012) 663–673. https://doi.org/10.1016/j.neuroimage.2012.06.054.

[137] J.N. Constantino, S. Kennon-McGill, C. Weichselbaum, N. Marrus, A. Haider, A.L. Glowinski, S. Gillespie, C. Klaiman, Infant viewing of social scenes is under genetic control and atypical in autism, 547 (2018) 340–344. https://doi.org/10.1038/nature22999.Infant.

[138] M.H. Black, N.T.M. Chen, K.K. Iyer, O. V. Lipp, S. Bölte, M. Falkmer, T. Tan, S. Girdler, Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography, Neurosci. Biobehav. Rev. 80 (2017) 488–515. https://doi.org/10.1016/j.neubiorev.2017.06.016.

[139] H. Richardson, G. Lisandrelli, A. Riobueno-Naylor, R. Saxe, Development of the social brain from age three to twelve years, Nat. Commun. 9 (2018) 1–12. https://doi.org/10.1038/s41467-018-03399-2.

[140] D. Moraczewski, G. Chen, E. Redcay, Inter-subject synchrony as an index of functional specialization in early childhood, Sci. Rep. 8 (2018) 1–12. https://doi.org/10.1038/s41598-018-20600-0.

[141] Y. Ren, V.T. Nguyen, L. Guo, C.C. Guo, Inter-subject Functional Correlation Reveal a Hierarchical Organization of Extrinsic and Intrinsic Systems in the Brain, Sci. Rep. 7 (2017) 1–12. https://doi.org/10.1038/s41598-017-11324-8.

[142] V. Costanzo, N. Chericoni, F.A. Amendola, L. Casula, F. Muratori, M.L. Scattoni, F. Apicella, Early detection of autism spectrum disorders: From retrospective home video studies to prospective “high risk” sibling studies, Neurosci. Biobehav. Rev. 55 (2015) 627–635. https://doi.org/10.1016/j.neubiorev.2015.06.006.

[143] A. Colver, C. Fairhurst, P.O.D. Pharoah, Cerebral palsy, Lancet. 383 (2014) 1240–1249. https://doi.org/10.1016/S0140-6736(13)61835-8.

[144] S.E. Cusick, M.K. Georgieff, The Role of Nutrition in Brain Development: The Golden Opportunity of the “First 1000 Days,” J. Pediatr. 175 (2016) 16–21. https://doi.org/10.1016/j.jpeds.2016.05.013.

[145] F.R. Willett, D.T. Avansino, L.R. Hochberg, J.M. Henderson, K. V Shenoy, High-performance brain-to-text communication via handwriting, Nature. 593 (2021). https://doi.org/10.1038/s41586-021-03506-2.

[146] J.S. Siegel, J.D. Power, J.W. Dubis, A.C. Vogel, J.A. Church, B.L. Schlaggar, S.E. Petersen, Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points, Hum. Brain Mapp. 35 (2014) 1981–1996. https://doi.org/10.1002/hbm.22307.

[147] C.L. Wiggs, A. Martin, Properties and mechanisms of perceptual priming, Curr. Opin. Neurobiol. 8 (1998) 227–233. https://doi.org/10.1016/S0959-4388(98)80144-X.

[148] M. Andric, S. Goldin-Meadow, S.L. Small, U. Hasson, Repeated movie viewings produce similar local activity patterns but different network configurations, Neuroimage. 142 (2016) 613–627. https://doi.org/10.1016/j.neuroimage.2016.07.061.

[149] U. Hasson, E. Yang, I. Vallines, D.J. Heeger, N. Rubin, A hierarchy of temporal receptive windows in human cortex, J. Neurosci. 28 (2008) 2539–2550. https://doi.org/10.1523/JNEUROSCI.5487-07.2008.

[150] R.B. Goldstein, R.L. Woods, E. Peli, Where people look when watching movies: do all vie, Comput. Biol. Med. 37 (2007) 957–964.

[151] S. Kastner, L.G. Ungerleider, Mechanisms of visual attention in the human cortex, Annu. Rev. Neurosci. 23 (2000) 315–341.

[152] I. Choi, J.Y. Lee, S.H. Lee, Bottom-up and top-down modulation of multisensory integration, Curr. Opin. Neurobiol. 52 (2018) 115–122. https://doi.org/10.1016/j.conb.2018.05.002.

[153] S.A. Hillyard, R.F. Hink, V.L. Schwent, T.W. Picton, Electrical signs of selective attention in the human brain, Science (80-.). 182 (1973) 177–180. https://doi.org/10.1126/science.182.4108.177.

[154] R. Subramanian, D. Shankar, N. Sebe, D. Melcher, Emotion modulates eye movement patterns and subsequent memory for the gist and details of movie scenes, J. Vis. 14 (2014) 1–18. https://doi.org/10.1167/14.3.31.

[155] A. Zlomuzica, F. Preusser, C. Totzeck, E. Dere, J. Margraf, The impact of different emotional states on the memory for what, where and when features of specific events, Behav. Brain Res. 298 (2016) 181–187. https://doi.org/10.1016/j.bbr.2015.09.037.

[156] E. Formisano, F. De Martino, M. Bonte, R. Goebel, "Who" Is Saying "What"? Brain-Based Decoding of Human Voice and Speech, Science (80-.). 322 (2008) 970–973. https://doi.org/10.1126/science.1164318.

[157] M. Dehghani, R. Boghrati, K. Man, J. Hoover, S.I. Gimbel, A. Vaswani, J.D. Zevin, M.H. Immordino-Yang, A.S. Gordon, A. Damasio, J.T. Kaplan, Decoding the neural representation of story meanings across languages, Hum. Brain Mapp. 38 (2017) 6096–6106. https://doi.org/10.1002/hbm.23814.

[158] W.A. de Heer, A.G. Huth, T.L. Griffiths, J.L. Gallant, F.E. Theunissen, The hierarchical cortical organization of human speech processing, J. Neurosci. 37 (2017) 6539–6557. https://doi.org/10.1523/JNEUROSCI.3267-16.2017.

[159] A.G. Huth, W.A. De Heer, T.L. Griffiths, F.E. Theunissen, J.L. Gallant, Natural speech reveals the semantic maps that tile human cerebral cortex, Nature. 532 (2016) 453–458. https://doi.org/10.1038/nature17637.

[160] E.M. Mugler, J.L. Patton, R.D. Flint, Z.A. Wright, S.U. Schuele, J. Rosenow, J.J. Shih, D.J. Krusienski, M.W. Slutzky, Direct classification of all American English phonemes using signals from functional speech motor cortex, 11 (2015) 1–16. https://doi.org/10.1088/1741-2560/11/3/035015.Direct.

[161] D.A. Moses, M.K. Leonard, J.G. Makin, E.F. Chang, Real-time decoding of question-and-answer speech dialogue using human cortical activity, Nat. Commun. 10 (2019). https://doi.org/10.1038/s41467-019-10994-4.

[162] B. Sorger, J. Reithler, B. Dahmen, R. Goebel, A Real-Time fMRI-Based Spelling Device Immediately Enabling Robust Motor-Independent Communication, Curr. Biol. 22 (2012) 1333–1338. https://doi.org/10.1016/j.cub.2012.05.022.

[163] H.C. Kim, S. Jin, S. Jo, J.H. Lee, A naturalistic viewing paradigm using 360° panoramic video clips and real-time field-of-view changes with eye-gaze tracking: Naturalistic viewing paradigm based on 360° panoramic video and real-time eye gaze, Neuroimage. 216 (2020) 116617. https://doi.org/10.1016/j.neuroimage.2020.116617.

[164] N. Reggente, J.K.Y. Essoe, Z.M. Aghajan, A. V. Tavakoli, J.F. McGuire, N.A. Suthana, J. Rissman, Enhancing the ecological validity of fMRI memory research using virtual

reality, Front. Neurosci. 12 (2018) 1–9. https://doi.org/10.3389/fnins.2018.00408.

[165] N. Naseer, K.S. Hong, fNIRS-based brain-computer interfaces: A review, Front. Hum. Neurosci. 9 (2015) 1–15. https://doi.org/10.3389/fnhum.2015.00003.

[166] M. Doulgerakis, A.T. Eggebrecht, H. Dehghani, High-density functional diffuse optical tomography based on frequency-domain measurements improves image quality and spatial resolution, Neurophotonics. 6 (2019) 1. https://doi.org/10.1117/1.nph.6.3.035007.

[167] H.M. Varma, C.P. Valdes, A.K. Kristoffersen, J.P. Culver, T. Durduran, Speckle contrast optical tomography: A new method for deep tissue three-dimensional tomography of blood flow, Biomed. Opt. Express. 5 (2014) 1275. https://doi.org/10.1364/boe.5.001275.

[168] C. Kolyva, A. Ghosh, I. Tachtsidis, D. Highton, C.E. Cooper, M. Smith, C.E. Elwell, Cytochrome c oxidase response to changes in cerebral oxygen delivery in the adult brain shows higher brain-specificity than haemoglobin, Neuroimage. 85 (2014) 234–244. https://doi.org/10.1016/j.neuroimage.2013.05.070.

[169] I. de Roever, G. Bale, R.J. Cooper, I. Tachtsidis, Functional NIRS measurement of cytochrome-c-oxidase demonstrates a more brain-specific marker of frontal lobe activation compared to the haemoglobins, Adv. Exp. Med. Biol. 977 (2017) 141–147. https://doi.org/10.1007/978-3-319-55231-6_19.

[170] N.S. Balbekin, M.S. Kulya, A. V. Belashov, A. Gorodetsky, N. V. Petrov, Increasing the resolution of the reconstructed image in terahertz pulse time-domain holography, Sci. Rep. 9 (2019) 1–9. https://doi.org/10.1038/s41598-018-36642-3.

[171] L. Tychsen, A. Burkhalter, Neuroanatomic abnormalities of primary visual cortex in macaque monkeys with infantile esotropia: preliminary results, J Pediatr Ophthalmol Strabismus. 32 (1995) 323–328.

[172] F. Sengpiel, C. Blakemore, P.C. Kind, R. Harrad, Interocular suppression in the visual cortex of strabismic cats, J Neurosci. 14 (1994) 6855–6871. https://doi.org/DOI: 10.1016/j.visres.2006.08.017.

[173] F. Sengpiel, K.U. Jirmann, V. Vorobyov, U.T. Eysel, Strabismic suppression is mediated by inhibitory interactions in the primary visual cortex, Cereb. Cortex. 16 (2006) 1750–1758. https://doi.org/10.1093/cercor/bhj110.

[174] B. Mansouri, B. Thompson, R.F. Hess, Measurement of suprathreshold binocular interactions in amblyopia, Vision Res. 48 (2008) 2775–2784. https://doi.org/10.1016/j.visres.2008.09.002.

[175] R. Farivar, B. Thompson, B. Mansouri, R.F. Hess, Interocular suppression in strabismic amblyopia results in an attenuated and delayed hemodynamic response function in early

visual cortex, J. Vis. 11 (2011) 16–16. https://doi.org/10.1167/11.14.16.

[176] R. Harrad, Psychophysics of suppression, Eye. 10 (1996) 270–273. https://doi.org/10.1038/eye.1996.57.

[177] J. Li, B. Thompson, C.S.Y. Lam, D. Deng, L.Y.L. Chan, G. Maehara, G.C. Woo, M. Yu, R.F. Hess, The role of suppression in amblyopia., Invest. Ophthalmol. Vis. Sci. 52 (2011) 4169–4176. https://doi.org/10.1167/iovs.11-7233.

[178] F. Sengpiel, C. Blakemore, The neural basis of suppression and amblyopia in strabismus, Eye. 10 (1996) 250–258. https://doi.org/10.1038/eye.1996.54.

[179] A.M.F. Wong, Implications of interocular suppression in the treatment of amblyopia, J. AAPOS. 15 (2011) 417–418. https://doi.org/10.1016/j.jaapos.2011.08.001.

[180] L. Tychsen, Why do humans develop strabismus?, in: C.S. Hoyt, D. Taylor (Eds.), Pediatr. Ophthalmol. Strabismus, 4th Editio, Elsevier, 2012: pp. 756–763.

[181] I.P. Conner, J.V. Odom, T.L. Schwartz, J.D. Mendola, Retinotopic maps and foveal suppression in the visual cortex of amblyopic adults., J. Physiol. 583 (2007) 159–73. https://doi.org/10.1113/jphysiol.2007.136242.

[182] V.J. Chen, K. Tarczy-Hornoch, Functional Magnetic Resonance Imaging of Binocular Interactions in Visual Cortex in Strabismus., J. Pediatr. Ophthalmol. Strabismus. (2010) 1–9. https://doi.org/10.3928/01913913-20101118-01.

[183] M.X. Repka, Amblyopia : basics , questions , and practical management, in: Taylor Hoyt's Pediatr. Ophthalmol. Strabismus, Fifth Edit, Elsevier Ltd, 2016: pp. 754-761.e2. https://doi.org/10.1016/B978-0-7020-6616-0.00073-6.

[184] D.H. Baker, T.S. Meese, B. Mansouri, R.F. Hess, Binocular summation of contrast remains intact in strabismic amblyopia, Investig. Ophthalmol. Vis. Sci. 48 (2007) 5332–5338. https://doi.org/10.1167/iovs.07-0194.

[185] G.D. Mower, W.G. Christen, J.L. Burchfiel, F.H. Duffy, Microiontophoretic bicuculline restores binocular responses to visual cortical neurons in strabismic cats, 309 (1984) 168–172.

[186] G.D. Mower, W.G. Christen, Evidence for an enhanced role of G A B A inhibition in visual cortical ocular dominance of cats reared with abnormal monocular experience, 45 (1989) 211–218.

[187] C.X. Guo, R.J. Babu, J.M. Black, W.R. Bobier, C.S.Y. Lam, S. Dai, T.Y. Gao, R.F. Hess, M. Jenkins, Y. Jiang, L. Kowal, V. Parag, J. South, S.E. Staffieri, N. Walker, A. Wadham, B. Thompson, Binocular treatment of amblyopia using videogames (BRAVO): study protocol for a randomised controlled trial, Trials. 17 (2016) 504.

https://doi.org/10.1186/s13063-016-1635-3.

[188] L. To, B. Thompson, J.R. Blum, G. Maehara, R.F. Hess, J.R. Cooperstock, A game platform for treatment of amblyopia, IEEE Trans. Neural Syst. Rehabil. Eng. 19 (2011) 280–289. https://doi.org/10.1109/TNSRE.2011.2115255.

[189] L. Tychsen, Causing and curing infantile esotropia in primates: the role of decorrelated binocular input (an American Ophthalmological Society thesis)., Trans. Am. Ophthalmol. Soc. 105 (2007) 564–93. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2258131&tool=pmcentrez&rendertype=abstract.

[190] M.W. Quick, M. Tigges, J.A. Gammon, R.G. Boothe, Early abnormal visual experience induces strabismus in infant monkeys, Investig. Ophthalmol. Vis. Sci. 30 (1989) 1012–1017.

[191] B. Zhang, H. Bi, E. Sakai, I. Maruko, J. Zheng, E.L. Smith, Y.M. Chino, Rapid plasticity of binocular connections in developing monkey visual cortex (V1), Proc. Natl. Acad. Sci. 102 (2005) 9026–9031. https://doi.org/10.1073/pnas.0500280102.

[192] L. Tychsen, M. Richards, A. Wong, P. Foeller, D. Bradley, A. Burkhalter, The Neural Mechanism for Latent (Fusion Maldevelopment) Nystagmus, J. Neuro-Ophthalmology. 30 (2010) 276–283. https://doi.org/10.1097/WNO.0b013e3181dfa9ca.

[193] F. Tong, M. Meng, R. Blake, Neural bases of binocular rivalry, Trends Cogn. Sci. 10 (2006) 502–511. https://doi.org/10.1016/j.tics.2006.09.003.

[194] F. Tong, K. Nakayama, J.T. Vaughan, N. Kanwisher, Binocular Rivalry and Visual Awareness in Human Extrastriate Cortex, 21 (1998) 753–759.

[195] F. Tong, S.A. Engel, Interocular rivalry revealed in the human cortical blind-spot representation, Nature. 1 (2001) 130–130. https://doi.org/10.1167/1.3.130.

[196] J. Haynes, G. Rees, Predicting the Stream of Consciousness from Activity in Human Visual Cortex, Curr. Biol. 15 (2005) 1301–1307. https://doi.org/10.1016/j.cub.2005.06.026.

[197] J.-D. Haynes, G. Rees, Predicting the orientation of invisible stimuli from activity in human primary visual cortex, J. Vis. 5 (2010) 221–221. https://doi.org/10.1167/5.8.221.

[198] M. Hudak, Increased readiness for adaptation and faster alternation rates under binocular rivalry in children, Front. Hum. Neurosci. 5 (2011) 1–7. https://doi.org/10.3389/fnhum.2011.00128.

[199] R.E. Black, C.G. Victora, S.P. Walker, Z.A. Bhutta, P. Christian, M. De Onis, M. Ezzati, S. Grantham-Mcgregor, J. Katz, R. Martorell, R. Uauy, Maternal and child undernutrition

and overweight in low-income and middle-income countries, Lancet. 382 (2013) 427–451. https://doi.org/10.1016/S0140-6736(13)60937-X.

[200] M. Duggan, Anthropometry as a tool for measuring malnutrition: impact of the new WHO growth standards and reference., Ann Trop Paediatr. 30 (2010) 1–17.

[201] S.B. Roberts, M.A. Franceschini, R.E. Silver, S.F. Taylor, A.B. De Sa, R. Có, A. Sonco, A. Krauss, A. Taetzsch, P. Webb, S.K. Das, C.Y. Chen, B.L. Rogers, E. Saltzman, P.Y. Lin, N. Schlossman, W. Pruzensky, C. Balé, K.K.H. Chui, P.C.T. Muentener, Effects of food supplementation on cognitive function, cerebral blood flow, and nutritional status in young children at risk of undernutrition: Randomized controlled trial, BMJ. 370 (2020). https://doi.org/10.1136/bmj.m2397.

[202] M.A. Ciliberto, H. Sandige, M.J. Ndekha, P. Ashorn, A. Briend, H.M. Ciliberto, M.J. Manary, Comparison of home-based therapy with ready-to-use therapeutic food with standard therapy in the treatment of malnourished Malawian children: A controlled, clinical effectiveness trial, Am. J. Clin. Nutr. 81 (2005) 864–870. https://doi.org/10.1093/ajcn/81.4.864.

[203] M.J. Manary, Local production and provision of ready-to-use therapeutic food (RUTF) spread for the treatment of severe childhood malnutrition, Food Nutr. Bull. 27 (2006) 83–89. https://doi.org/10.1177/15648265060273s305.

[204] S. Subramanian, S. Huq, T. Yatsunenko, R. Haque, M. Mahfuz, M.A. Alam, A. Benezra, J. Destefano, M.F. Meier, B.D. Muegge, M.J. Barratt, L.G. VanArendonk, Q. Zhang, M.A. Province, W.A. Petri, T. Ahmed, J.I. Gordon, Persistent gut microbiota immaturity in malnourished Bangladeshi children, Nature. 510 (2014) 417–421. https://doi.org/10.1038/nature13421.

[205] L. V. Blanton, M.J. Barratt, M.R. Charbonneau, T. Ahmed, J.I. Gordon, Childhood undernutrition, the gut microbiota, and microbiota-directed therapeutics, Science (80-.). 352 (2016). https://doi.org/10.1126/science.aad9359.

[206] R.Y. Chen, I. Mostafa, M.C. Hibberd, S. Das, M. Mahfuz, N.N. Naila, M.M. Islam, S. Huq, M.A. Alam, M.U. Zaman, A.S. Raman, D. Webber, C. Zhou, V. Sundaresan, K. Ahsan, M.F. Meier, M.J. Barratt, T. Ahmed, J.I. Gordon, A Microbiota-Directed Food Intervention for Undernourished Children, N. Engl. J. Med. (2021). https://doi.org/10.1056/nejmoa2023294.

[207] A. Vouloumanos, M.D. Hauser, J.F. Werker, A. Martin, The tuning of human neonates' preference for speech, Child Dev. 81 (2010) 517–527. https://doi.org/10.1111/j.1467-8624.2009.01412.x.

[208] P.W. Jusczyk, R.N. Aslin, Infants' Detection of the Sound Patterns of Words in Fluent Speech, Cogn. Psychol. 29 (1995) 1–23. https://doi.org/10.1006/cogp.1995.1010.

[209] D.J. Newbold, T.O. Laumann, C.R. Hoyt, J.M. Hampton, D.F. Montez, R. V. Raut, M. Ortega, A. Mitra, A.N. Nielsen, D.B. Miller, B. Adeyemo, A.L. Nguyen, K.M. Scheidter, A.B. Tanenbaum, A.N. Van, S. Marek, B.L. Schlaggar, A.R. Carter, D.J. Greene, E.M. Gordon, M.E. Raichle, S.E. Petersen, A.Z. Snyder, N.U.F. Dosenbach, Plasticity and Spontaneous Activity Pulses in Disused Human Brain Circuits, Neuron. 107 (2020) 580-589.e6. https://doi.org/10.1016/j.neuron.2020.05.007.

[210] D.J. Newbold, E.M. Gordon, T.O. Laumann, N.A. Seider, D.F. Montez, S.J. Gross, A. Zheng, A.N. Nielsen, C.R. Hoyt, J.M. Hampton, M. Ortega, B. Adeyemo, D.B. Miller, A.N. Van, S. Marek, B.L. Schlaggar, A.R. Carter, B.P. Kay, D.J. Greene, M.E. Raichle, S.E. Petersen, A.Z. Snyder, N.U.F. Dosenbach, Cingulo-opercular control network and disused motor circuits joined in standby mode, Proc. Natl. Acad. Sci. U. S. A. 118 (2021) 1–10. https://doi.org/10.1073/pnas.2019128118.

[211] B.P. Lucey, J.S. Mcleland, C.D. Toedebusch, J. Boyd, J.C. Morris, E.C. Landsness, K. Yamada, D.M. Holtzman, Comparison of a single-channel EEG sleep study to polysomnography, J. Sleep Res. 25 (2016) 625–635. https://doi.org/10.1111/jsr.12417.

[212] B.P. Lucey, A. McCullough, E.C. Landsness, C.D. Toedebusch, J.S. McLeland, A.M. Zaza, A.M. Fagan, L. McCue, C. Xiong, J.C. Morris, T.L.S. Benzinger, D.M. Holtzman, Reduced non-rapid eye movement sleep is associated with tau pathology in early Alzheimer's disease, Sci. Transl. Med. 12 (2020). https://doi.org/10.1126/scitranslmed.aba5187.

[213] A. Myrden, L. Schudlo, S. Weyand, T. Zeyl, T. Chau, Trends in communicative access solutions for children with cerebral palsy, J. Child Neurol. 29 (2014) 1108–1118. https://doi.org/10.1177/0883073814534320.

"This life is more than just a read through"

- Red Hot Chili Peppers

Curriculum Vitae

Kalyan Tripathy

Education

Washington University in St. Louis, School of Medicine, St. Louis, MO
M.D.-Ph.D. candidate, August 2014 – present (anticipated graduation: May 2023)
Graduate program: Neurosciences Graduate school GPA: 4.00/4.00

University of Pennsylvania, College of Arts and Sciences, Philadelphia, PA
BA, *summa cum laude*, with Distinction in Biology, May 2012
Major: Biology Concentration: Neurobiology Minor: Psychology GPA: 3.92/4.00

Awards and Honors

Tom Thach Award for Systems Neuroscience Research
McDonnell Center for Systems Neuroscience, Washington University in St. Louis, 2020

Best Research Poster Award
Medical Scientist Training Program Research Retreat, Washington University School of Medicine, 2019

David F. Silbert Outstanding Teaching Assistant Award
Washington University School of Medicine, 2018

McDonnell Fellowship in Cognitive, Computational, and Systems Neuroscience
McDonnell Center for Systems Neuroscience, Washington University in St. Louis, 2017-2018

Nathan Edward Hellman, MD, PhD, Memorial Award
Washington University School of Medicine, 2016

Sol Feinstone Award for creating constructive social and educational change
University of Pennsylvania, 2012

Dean's List
University of Pennsylvania, 2008-09, 2009-10, 2010-11, 2011-12

Publications

Tripathy K, Markow ZE, Fishell AK, Sherafati A, Burns-Yocum TM, Schroeder ML, Svoboda AM, Eggebrecht AT, Anastasio MA, Schlaggar BL, Culver JP. Decoding visual information from high-density diffuse optical tomography neuroimaging data. Neuroimage. 2021 Feb; 1;226:117516. doi: 10.1016/j.neuroimage.2020.117516

Walker AK*, **Tripathy K***, Restrepo CR, Ge G, Xu Y, Kwong LK, Trojanowski JQ, Lee VM. An insoluble frontotemporal lobar degeneration-associated TDP-43 C-terminal fragment causes neurodegeneration and hippocampus pathology in transgenic mice. Hum Mol Genet. 2015 December; 24(25):7241-54. doi: 10.1093/hmg/ddv424 ***Co-first-author publication**

Walker AK, Spiller KJ, Ge G, Zheng A, Xu Y, Zhou M, **Tripathy K**, Kwong LK, Trojanowski JQ, Lee VM. Functional recovery in new mouse models of ALS/FTLD after clearance of pathological cytoplasmic TDP-43. Acta Neuropathol. 2015 November; 130(5):643-60. doi: 10.1007/s00401-015-1460-x

Pesiridis GS, **Tripathy K**, Tanik S, Trojanowski JQ, Lee VM-Y. A "two-hit" hypothesis for inclusion formation by carboxy-terminal fragments of TDP-43 protein linked to RNA depletion and impaired microtubule-dependent transport. J Biol Chem. 2011 May; 286(21):18845-55. doi: 10.1074/jbc.M111.231118

Conference Talks

Tripathy K, Svoboda AM, Schroeder ML, Fishell AK, Richter EJ, Rafferty S, Tracy C, Markow ZE, Wheelock MD, Burke B, Eggebrecht AT, Culver JP. Functional brain mapping in preschool-age children with high density diffuse optical tomography. Oral presentation at SPIE Photonics West. San Francisco, CA, USA. February 2020.

Tripathy K, Fishell AK, Markow ZE, Burns-Yocum TM, Eggebrecht AT, Schlaggar BL, Culver JP. Decoding visual information from high density diffuse optical tomography neuroimaging data. Oral Presentation at fNIRS. Tokyo, Japan. 2018 October.

Tripathy K, Fishell AK, Markow ZE, Burns-Yocum TM, Newbold DJ, Tripathy P, Schlaggar BL, Culver JP. Decoding visual information from high density diffuse optical tomography neuroimaging data. Oral Presentation at Optical Society of America Biophotonics Congress. Hollywood, FL, USA. 2018 April.

Poster Presentations

Tripathy K, Svoboda AM, Schroeder ML, Fishell AK, Richter EJ, Rafferty S, Tracy C, Markow ZE, Wheelock MD, Burke B, Eggebrecht AT, Culver JP. A new high density diffuse optical tomography system for functional brain mapping in preschool age children. Poster presentation at Society for Neuroscience Annual Meeting. Chicago, IL, USA. October 2019**.**

Tripathy K, Svoboda AM, Schroeder ML, Fishell AK, Richter EJ, Rafferty S, Tracy C, Markow ZE, Wheelock MD, Burke B, Eggebrecht AT, Culver JP. Functional brain mapping in preschool-age children with high density diffuse optical tomography. Poster presentation at Flux Congress: The Society for Developmental Cognitive Neuroscience. New York, NY, USA. August 2019**.**

Tripathy K, Fishell AK, Markow ZE, Eggebrecht AT, Schlaggar BL, Culver JP. Decoding visual stimulus position from high-density diffuse optical tomography neuroimaging data. Poster presentation at Organization for Human Brain Mapping conference. Rome, Italy. June 2019**.**

Tripathy K, Richter EJ, Agato A, Markow ZE, Fishell AK, Sherafati A, Burke BA, Eggebrecht AT, Culver JP. Wireless High-Density Diffuse Optical Tomography for Mapping Brain Activity. Poster presentation at BRAIN Initiative® Investigators Meeting. Washington DC, USA. 2019 April.

Tripathy K, Markow ZE, Fishell AK, Burns-Yocum TM, Eggebrecht AT, Anastasio MA, Schlaggar BL, Culver JP. Decoding visual stimulus identity from high-density diffuse optical tomography measurements of human brain hemodynamics. Poster presentation at Neural Imaging and Sensing sub-conference, Photonics West conference, SPIE. San Francisco, CA, USA. February 2019.

Tripathy K, Fishell AK, Markow ZE, Burns-Yocum TM, Schlaggar BL, Culver JP. Visual decoding and preschool age functional neuroimaging with high density diffuse optical tomography neuroimaging data. Poster Presentation at Flux Congress: The Society for Developmental Cognitive Neuroscience. Berlin, Germany. 2018 September.

Research Experiences

Culver Lab, Washington University in St. Louis: M.D.-Ph.D. candidate (September 2016 – present)

- PI: Joseph P. Culver
- Conducting independent research on neuroimaging method development, neural decoding, and early brain development:
    - designed studies with my PI
    - helped fund research by obtaining a student fellowship and by assisting my PI with grant applications
    - built new high-density diffuse optical tomography (HD-DOT) caps for imaging children and adults
    - created audiovisual stimuli, stimulus presentation code, and data processing code
    - imaged adult and child study participants using HD-DOT and fMRI

    - wrote, optimized, and ran MATLAB code for brain mapping, decoding, and statistical hypothesis testing
    - presented periodic progress updates to funding agencies (e.g. Gates Foundation) and thesis committee
    - presented posters and talks at international conferences and university retreats and seminars
    - writing papers for publication
- Thesis topic: Mapping brain development and decoding brain activity with diffuse optical tomography:
  (1) Evaluated feasibility and performance of visual decoding with HD-DOT
  (2) Developed HD-DOT imaging tools for high-performance functional brain mapping in awake young children
  (3) Evaluated feasibility and performance of decoding naturalistic audiovisual information from HD-DOT data

Bateman Lab, Washington University in St. Louis: M.D.-Ph.D. rotation student (July 2016 – August 2016)
- PI: Randall J. Bateman
- Learned to measure amyloid beta levels in CSF by protein purification and mass spectrometry; performed a pilot study on purifying intact amyloid beta and analyzing its full range of proteoforms by discovery-oriented, top-down mass spectrometry, which provided preliminary data for a subsequent grant application

Neuroimaging Labs, Washington University in St. Louis: M.D.-Ph.D. rotation student (June 2015 – July 2015)
- PIs: Bradley L. Schlaggar, Steven E. Petersen
- Learned to use basic neuroimaging and computational research tools such as MATLAB, Freesurfer, and Workbench, and applied them to study functional connectivity networks and cortical parcellation in children's brains using resting state fMRI data

Center for Neurodegenerative Disease Research (CNDR), University of Pennsylvania: Undergraduate student (May 2009 – May 2012), Research Specialist (August 2012 – May 2014)
- PIs: Virginia M.-Y. Lee, John Q. Trojanowski
- Studied how the protein TDP-43 contributes to the pathogenesis of ALS and FTD using *in vitro* and mouse models of disease, through diverse approaches including protein purification, biochemistry, electron microscopy, molecular biology, cell culture, immunocytochemistry, mouse behavioral testing, mouse perfusion and dissection, and immunohistochemistry
- Honors thesis studying purification and prionoid properties of TDP-43 awarded grade of A+ by PI and graduation with Distinction by Department of Biology
- Gave annual research talks at CNDR; presented posters at several institutional retreats; contributed to multiple publications (below)
- Designed and implemented a weekly symposium for summer student researchers to promote sharing of knowledge and experiences; mentored younger students through this platform during my later years in the lab

Teaching, Leadership, and Extracurricular Experiences

Washington University School of Medicine: Teaching Assistant, Neural Sciences course (March 2018 – May 2018)
- Helped teach neuroscience to 125 first-year medical school students together with a team of professors and other TAs
- Assisted professors and conducted small group discussions with students during brain dissection and histology labs
- Prepared slides, talks, and interactive quizzes for weekly 2-hour-long review sessions
- Prepared exam review materials such as summary slide decks and practice practical tests for the midterm and final
- Provided feedback to professors on lab presentation slides and exam questions
- Responded directly to students who emailed with questions about course content, and liaised between students and professors to respond to logistical requests
- Provided additional focused mentorship to students who requested extra assistance, which helped boost their exam performance

Teaching school children: Volunteer with numerous service projects (September 2008 – present)
- Mentored underprivileged school children from West Philadelphia and tried to foster their interest in science through programs such as the Cardiology Pipeline Program, the West Philadelphia Tutoring Project, and the Upward Bound Math and Science Summer Scholars Program at the University of Pennsylvania
- Ran English language and health literacy workshops for children in slums in North Kolkata through Pratit International
- Led cultural exchange and college preparation sessions for high school students in Navajo Nation during a spring break service trip organized by the Washington University School of Medicine
- Lectured and conducted experiential learning activities to introduce children to neuroscience at events such as the Kids Judge Neuroscience Fair at the University of Pennsylvania, the Camp Neuro summer program at the Washington University School of Medicine, and the Amazing Brain Carnival at the St. Louis Science Center
- Contributed microscope images of neuronal cultures now on permanent display for the new "Your Brain" exhibit at The Franklin Institute, the children's science education museum in Philadelphia
- Coached high school students in neuroanatomy and histology for the National Brain Bee Championship; one of them placed first in the country and then 6th in the world at the International Brain Bee Championship

Pratit International: Charter Member (September 2008), Board Member of Pratit at Penn (January 2010 – August 2010), Co-President of Pratit at Penn (August 2010 – December 2011), Board Member of Pratit International (January 2012 – present)
- Co-founded and ran the Penn chapter of Pratit International, a humanitarian organization providing healthcare to slums in Kolkata, India: established organizational structure, recruited members, led meetings, represented the organization, liaised with team in India
- Helped run year-round humanitarian effort in Kolkata, India: helped direct expansion and improvement of healthcare services, led operations on the ground, contributed to patient care at clinics, contributed to patient care remotely through telemedicine initiatives
- Helped organize responses to health crises affecting our communities including natural disasters and the COVID-19 pandemic: including fundraising, distribution of food relief and sanitation supplies, coordination of situation-specific health care response
- Continuing to advise on new health care initiatives

Saturday Neighborhood Health Clinic, Washington University School of Medicine: Coordinator (October 2014 – May 2015), Volunteer (October 2016 – present)
- Helped coordinate WUSM's community service clinics for a year
- Worked longitudinally with patients to help them obtain follow-up care
- Served as an interpreter for Hindi and Bengali speaking patients
- Continuing to volunteer at clinics, assessing patients under the guidance of an attending physician, and mentoring first-year and second-year medical students

Hospital of the University of Pennsylvania: Volunteer (September 2011 – July 2012)
- Engaged in >100 hours of service: assisting patients and their families; liaising between patients, visitors, care providers and other hospital staff; performing various administrative tasks; training new volunteers through the Volunteer Mentor Program

Musical performance: Band manager, lead vocalist, and rhythm guitarist; dance team member (September 2014 – present)
- Recruited musicians to form a medical student band; led rehearsals and performances at social events, cultural shows, fundraisers, memorial services, hospitals, and nursing homes; wrote, recorded, and edited original music, covers, and music videos
- Performed with dance teams at cultural shows and fundraisers