UNIVERSITY OF CALIFORNIA

Los Angeles

# Neural and Psychological Coordination in Social Communication and Interaction

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Psychology

by

Shannon Burns

2020

ABSTRACT OF THE DISSERTATION

Neural and Psychological Coordination in Social Communication and Interaction

by

Shannon Burns

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2020

Professor Matthew D. Lieberman, Chair

Dynamic, naturalistic study of social interactions in humans is a small but growing literature. Emerging from this work is the theory that social interaction creates a “merged mind” between interlocutors – they come into psychological, behavioral, and neural alignment in order to better predict each other and coordinate as one social unit. However social interaction is diverse, so more work is needed to understand the specific nature of alignment between people in a variety of interactive contexts. In particular, it’s unclear how heterogeneities among members of an interaction impact their ability to align. This work aims to help address this gap by first evaluating and improving ways to collect neuroimaging data in naturalistic, social settings (Chapter 2). Then, empirical research is presented that examines how personal similarity factors impact the extent of alignment during personal disclosure interactions, where one person speaks and the other listens (Chapter 3). Finally, further empirical research investigates different types of alignment that may be present in a dyadic back-and-forth discussion in a joint decision-making paradigm. How this work contributes to a broader

understanding of the ways people communicate and work together, and how this research can continue with improved methods, is discussed.

This dissertation of Shannon M. Burns is approved.

Richard Alan Clarke Dale

Carolyn M. Parkinson

Jennifer Ashley Silvers

Matthew D. Lieberman, Committee Chair

University of California, Los Angeles

2020

# Table of Contents

# List of Figures

# List of Tables

# Acknowledgments

I would like to thank my committee members Dr. Matthew Lieberman, Dr. Rick Dale, Dr. Carolyn Parkinson, and Dr. Jennifer Silvers for their advice and feedback during development of this dissertation and during my time at UCLA. I would like to thank my past and present lab-mates in the Social Cognitive Neuroscience lab, other members of the Social Psychology program, and other cohort members in the Psychology department for their assistance and social support over the years. I am particularly grateful for the guidance, insight, and friendship of Macrina Cooper-White, Jessica Shropshire, Carrianne Leschak, and Lianne Barnes who made it possible for me to finish this journey. I would also like to thank the team of research assistants who worked on these projects between 2017 and 2020. Without their enthusiasm and dedication, these data would not exist.

Outside of UCLA, I am thankful to my family for their eternal support and love. You taught me how to be curious about the world and how to persist through whatever obstacles arise. I am also grateful to my partner Eli Berg for his coding advice, awful puns, and endless love during a challenging time.

Finally, this research was supported by the United States Department of Defense MINERVA Initiative, the National Science Foundation Graduate Research Fellowship, and the UCLA Dissertation Year Fellowship.

# Select Vita

## EDUCATION

| | | |
|---|---|---|
| Advancement to Doctoral Candidacy | University of California – Los Angeles | March 2019 |
| M.A., Social Psychology | University of California – Los Angeles | 2016 |
| B.A., Psychology | Pomona College | 2015 |

## PUBLICATIONS

**Burns, S.M**., Barnes, L.N., McCulloh, I.A., Dagher, M.M., Falk, E.B., Storey, J.D., & Lieberman, M.D. (2019). Making social neuroscience less WEIRD: Using fNIRS to measure neural signatures of persuasive influence in a Middle East participant sample. *Journal of Personality and Social Psychology, 116(3),* e1-e11. doi: 10.1037/pspa0000144

**Burns, S.M**., Barnes, L.N., Katzman, P.L, Ames, D.L., Falk, E.B., & Lieberman, M.D. (2018). A functional near infrared spectroscopy (fNIRS) replication of the sunscreen persuasion paradigm. *Social Cognitive and Affective Neuroscience, 13(6),* 628-636.

Lieberman, M.D., **Burns, S.M**., Torre, J.B., & Eisenberger, N.I. (2016). Reply to Wager et al: Pain and the dACC: The importance of hit rate-adjusted effects and posterior probabilities with fair priors. *Proceedings of the National Academy of Sciences, 113,* E2476-E2479.

## MANUSCRIPTS IN PREPARATION

**Burns, S.M.** & Lieberman, M.D. (in prep). The use of fNIRS for unique contributions to social and affective neuroscience. Preprint: https://psyarxiv.com/kygbm

**Burns, S.M.,** Barnes, L.N., & Lieberman, M.D. (in prep). Effects of shared experience and perceived similarity on neural and psychological alignment between speakers and listeners.

**Burns, S.M.,** Dale, R., & Lieberman, M.D. (in prep). Linear, nonlinear, and higher-order measures of neural synchrony in conversing dyads.

## ORAL PRESENTATIONS

**indicates presenting author*

***Burns, S.M.,** Barnes, L.N., Dale, R., & Lieberman, M.D. (Sep 2019). Dynamic neural systems in natural social interaction. Invited talk presented at University of Pennsylvania & Princeton University.

***Burns, S.M.,** Cooper-White, M.A., Barnes, L.N., McCulloh, I.A., Dagher, M.M., Falk, E.B., Storey, J.D., & Lieberman, M.D. (Feb 2018). Cross-cultural examination of brain indicators of persuasion. Symposium talk at the *Society for Personality and Social Psychology Annual Convention.* Atlanta, GA.

***Burns, S.M.,** Barnes, L.N., Katzman, P.L, Ames, D.L., Falk, E.B., McCulloh, I.A., Dagher, M.M., Storey, J.D., & Lieberman, M.D. (May 2017). Neuroscience beyond the scanner: Social brain processes in real-world contexts. Invited talk presented at UCLA Psychology Department.

## AWARDS & HONORS

| | |
|---|---|
| Shelley Taylor Dissertation Award – UCLA Social Psychology Area | 2020 |
| UCLA Dissertation Year Fellowship | 2019 |
| Collegium of University Teaching Fellowship | 2019 |
| Harold H. Kelley Award for a meritorious research paper addressing basic issues or theory - UCLA Social Psychology Area | 2018 |
| Society for Personality and Social Psychology Student Travel Award | 2017 |
| National Science Foundation Graduate Research Fellowship | 2016 |
| Distinction in the Senior Exercise - Pomona College | 2015 |
| Pomona College Scholar (Dean's List) | Spring 2015<br>Fall 2014<br>Spring 2014<br>Fall 2012 |

# Chapter 1 - Dissertation Overview

## Background

Over the course of a normal day, one might engage in various social interactions – listening to a friend talk about their vacation, meeting with colleagues at work, or cooperating with a spouse to get children bathed and put to bed. Humans spend much of their waking life interacting with other people in some way, and much of this is done relatively seamlessly and effortlessly. Yet, the psychological abilities needed to manage and coordinate cognition and behavior with others can be quite complex. The occasional social gaffe is a reminder of how much could go wrong.

Traditional social psychology and neuroscience has typically tried to parse the complexities of social cognition and behavior by distilling them into subcomponents that could be investigated individually, summarized simply, and reproduced in a laboratory. This pursuit has offered important insights into processes such as face perception, social information memory, or physical action prediction. Yet an understanding of how these processes naturally operate and interact in real-world scenarios remains incomplete. Further, most research has focused on the level of the individual in a social interaction – how they view social agents around them, regulate their own behaviors, etc. While useful for understanding the psychological experience of an individual, this doesn’t address aspects of the social system as a whole, taking into account the dynamic and emergent phenomena of a social interaction arising between two or more people. Thus the majority of social psychological theory is built on literature about the human mind in isolation rather than in context and interaction with other people.

That isn't to say that investigations of dynamic, real world social interaction are entirely absent, however. Though the majority of social psychology has focused on the individual as the unit of analysis, some researchers have investigated interactive social processes between people. In fact, some of the earliest work in social psychology as a formal field described the concept of "group dynamics," or the social forces that guide behavior. Motivated by the seemingly inhuman atrocities committed during World War II, psychologist Kurt Lewin argued that it was important to move past the description of behavior in individuals and study the dynamic phenomena emerging from a social group context, as some group behaviors could not be comprehended by investigating a person alone (Lewin, 1947). Lewin's theorizing about group behavior also emphasized the dynamical nature of it, describing how fluctuations in racial prejudice might be created in a town, or how work group members may collectively increase aggressive behaviors under certain leadership styles.

In more modern research, advances in statistical analysis have enabled more detailed characterizations of the processes that enable the coordination of thought and action over time, and the creation of shared meaning between people. A noteworthy thread emerging from this work is the observation that people tend to naturally synchronize the oscillations of a host of behaviors such as walking, talking, gesturing, or fidgeting (Lakin, 2013). This often happens without awareness, but produces several downstream effects such as interpersonal liking, emotion regulation, and changes in cognitive executive functioning (Reddish, Fischer, & Bulbulia, 2013; Valdesolo, Ouyang, & DeSteno, 2010). Convergence in language use and problem-solving strategies also seems to occur between conversation partners and facilitates their joint performance (Fusaroli et al., 2012; Hinsz, Tindale, & Vollrath, 1997; Thompson & Fine, 1999). The term "shared cognition" has thus been used to describe the ways in which task representations and information processing strategies are more similar between interacting partners than between disconnected single actors (Tindale & Kameda, 2000; van den Bossche et al., 2011).

Social neuroscience has begun to investigate the neural mechanisms underlying these coordination dynamics as well. Much as the previously mentioned behavior signals do, neural patterns seem to synchronize when people are engaged in behavioral mimicry (Holper, Scholkmann, & Wolf, 2012; Osaka et al., 2015; Pan et al., 2017). Neural synchrony has also been identified in cases of joint task performance (e.g., Dommer et al., 2012; Fishburn et al., 2018), interpersonal storytelling (e.g., Stephens, Silbert, & Hasson, 2010; Yeshurun et al., 2017), and dyadic discussion (Jiang et al., 2015; Sänger, Müller, & Lindenberger, 2013).

Based on this evidence, theorists have suggested that shared cognition reflects mutual attention direction, similar representational mappings, and the ability to predict upcoming behavior between interacting individuals (Dale et al., 2013; Macrae et al., 2008; Pelose, 1987). This enables fluid exchange of information and meaning. In turn, this shared cognition (as represented by neural synchrony) reflects a "merging of minds" that is necessary for shared understanding and joint goal-directed behavior to occur (Kelso, Dumas, & Tognoli, 2013; Thompson & Fine, 1999; Wheatley et al., 2012). Multiple people engaging in the same mental processes enables personally-relevant functions such as learning of new information from others, but also actions like group movement and maintenance that are evolutionarily adaptive for humans.

Yet, this theory that psychological and neural convergence enables social interaction is young and large questions remain. In particular, what about the social interactions where interlocutors don't perform the same action or have the same ideas? In the previously described research, coordination among people was the goal, but preexisting differences among interaction partners weren't assessed that may influence social convergence. In addition, this "merging of minds" is conceived to be an exact reinstantiation of patterns across people, but it is not clear if weaker forms of coherence may still exist in otherwise asynchronous situations. Finally, the potential value of *not* being on the same page as another person is unclear, though could potentially benefit situations where "two minds are better than one."

## This Dissertation

This dissertation aims to help fill some of the holes in our understanding of mental convergence in social interaction by examining the potential phenomenon under varying conditions of heterogeneity among social actors. Across two studies, I and collaborators collected neural, behavioral, and survey data in order to estimate how participants' psychological experience varies within social communication and interaction, and if there are still coherent dynamics among these experiences.

Measuring neural processes within social interactions is particularly difficult, as traditional neuroimaging modalities such as fMRI require participants to be immobilized and thus are not well suited to natural social contexts. A relatively underutilized neuroimaging technology called near-infrared spectroscopy (fNIRS) is more robust to participant motion and thus uniquely situated for research on social interaction, so it was used throughout this dissertation. However, its uncommonness as a neuroscience research tool means that best practices are still in development and data processing approaches are unstandardized. Thus, in Chapter 2 of this dissertation, I explore in depth the functional capacity of fNIRS as a research tool for social neuroscience. In addition, I describe techniques I developed for exploring fNIRS data quality in social interaction research in order to make informed decisions about data inclusion.

In Chapter 3, I investigate neural and psychological alignment in personal disclosure communication and what may modulate its strength. Specifically, this experiment focuses on how similarly listeners of a personal narrative encode the content of that narrative in relation to the speaker sharing that narrative. This type of speaker-listener design has been used before in past investigations of neural alignment (Stephens, Silbert, & Hasson, 2010; Liu et al., 2017), but it is currently unclear how variables like interpersonal similarity may make it easier or more difficult to align in understanding of a communicated narrative. Past research hypothesizes that more similar people and/or people with similar experiences may understand each other better in communication as well as express more empathy to each other (Banks, Berenson & Carkhuff, 1967; Haley & Dowd, 1988; Kirk, Best, & Irwin, 1986; Robiner & Storandt, 1983), so these attributes are investigated in relation to how they may be associated with neural and psychological alignment between a speaker and a listener during the recounting of a narrative. In addition, alignment is also measured between pairs of listeners, to investigate how well they converge on one understanding of the narrative even if that particular understanding does not match the speaker's. The results of this investigation shed light on the effect of interpersonal similarity on social communication, as well as the nature of interpersonal alignment in this type of social interaction more generally.

Finally, in Chapter 4 I explore neural synchronization in a discussion that involves back and forth deliberation between two people to jointly solve a decision-making task. Unlike in speaker-listener scenarios where successful communication can be defined as accurate reinstantiation of one person's mental representation in that of another, successful discussion may involve separate thought patterns between interlocutors as they present alternative view points before eventually converging on a joint decision. Thus, the specific type of alignment that might occur in this sort of social interaction may not be the same as that experienced in communication with one speaker and one listener (Fusaroli & Tylén, 2016; Kelso, Dumas, & Tognoli, 2013; Riley et al., 2011). Therefore this chapter examines different approaches to calculating neural alignment to see which is a better characterization of the neural dynamics within interpersonal discussion, and which better predicts discussion outcomes such as the efficiency of the discussion and the interpersonal feelings it engenders between discussion partners.

Altogether, this work aims to refine research approaches to social interaction topics as well as improve our understanding of whether and how people "merge minds" in natural social interaction.

## References

Banks, G., Berenson, B.G., & Carkhuff, R.R. (1967). The effects of counselor race and training upon counseling process with Negro clients in initial interviews. *Journal of Clinical Psychology, 23(1),* 70-72.

Dommer, L., Jäger, N., Scholkmann, F., Wolf, M., & Holper, L. (2012). Between-brain coherence during joint n-back task performance: A two-person functional near-infrared spectroscopy study. *Behavioural Brain Research*, *234*(2), 212–222.

Fishburn, F. A., Murty, V. P., Hlutkowsky, C. O., Caroline, E., Bemis, L. M., Murphy, M. E., … Perlman, S. B. (2018). Putting our heads together : interpersonal neural synchronization as a biological mechanism for shared intentionality. *Social Cognitive and Affective Neuroscience*, *13*(8), 841–849.

Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., & Tylén, K. (2012). Coming to terms: Quantifying the benefits of linguistic coordination. *Psychological Science*, *23*(8), 931–939.

Fusaroli, R. & Tylén, K. (2016). Investigating conversational dynamics: Interactive alignment, interpersonal synergy, and collective task performance. *Cognitive Science, 40,* 145-171.

Haley, T.J. & Dowd, E.T. (1988). Responses of deaf adolescents to differences in counselor method of communication and disability status. *Journal of Counseling Psychology, 35(3),* 258-262.

Hinsz, V. B., Tindale, R. S., & Vollrath, D. a. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin*, *121*(1), 43–64.

Holper, L., Scholkmann, F., & Wolf, M. (2012). Between-brain connectivity during imitation measured by fNIRS. *NeuroImage*, *63*(1), 212–222.

Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., & Lu, C. (2015). Leader emergence through interpersonal neural synchronization. *Pnas*, *112*(14), 4274–4279.

Kelso, J.A.S., Dumans, G., & Tognoli, E. (2013). Outline of a general theory of behavior and brain coordination. *Neural Networks, 37,* 120-131.

Kirk, W.G., Best, J.B. & Irwin, P. (1986). The perception of empathy in alcoholism counselors. *Journal of Studies on Alcohol, 47(1),* 82-84.

Lakin, J. L. (2013). Behavioral mimicry and interpersonal synchrony. In J. Hall & M. Knapp (Eds.): *Nonverbal Communication*, 539–576. Berlin: Walter de Gruyter.

Lewin, K. (1947). Frontiers in group dynamics: Concept, method, and reality in social science; social equilibria and social change. *Human Relations*, *1*, 5–41.

Liu, Y., Piazza, E.A., Simony, E., Shewokis, P.A., Onaral, B., Hasson, U., & Ayaz, H. (2017). Measuring speaker-listener neural coupling with functional near infrared spectroscopy. *Scientific Reports, 7,* 43293.

McGrath, J.E. (1984). *Groups: Interaction and Performance.* New Jersey: Prentice-Hall, Inc.

Osaka, N., Minamoto, T., Yaoi, K., Azuma, M., Shimada, Y. M., & Osaka, M. (2015). How two brains make one synchronized mind in the inferior frontal cortex: FNIRS-based hyperscanning during cooperative singing. *Frontiers in Psychology*, *6*(NOV), 1–11.

Pan, Y., Cheng, X., Zhang, Z., Li, X., & Hu, Y. (2017). Cooperation in lovers : An fNIRS-based hyperscanning study. *Human Brain Mapping*, *38*(2), 831–841.

Reddish, P., Fischer, R., & Bulbulia, J. (2013). Let's dance together: Synchrony, shared intentionality and cooperation. *PloS One, 8(8),* e71182.

Riley, M.A., Richardson, M.J., Shockley, K., & Ramenzoni, V.C. (2011). Interpersonal synergies. *Frontiers in Psychology, 2,* 38.

Robiner, W.N. & Storandt, M. (1983). Client perceptions of the therapeutic relationship as a function of client and counselor age. *Journal of Counseling Psychology, 30(1),* 96-99.

Sänger, J., Müller, V., & Lindenberger, U. (2013). Directionality in hyperbrain networks discriminates between leaders and followers in guitar duets. *Frontiers in Human Neuroscience*, *7*(June), 234. http://doi.org/10.3389/fnhum.2013.00234

Stephens, G.J., Silbert, L.J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences, 107(32),* 14425-14430.

Tindale, R. S., & Kameda, T. (2000). 'Social sharedness' as a unifying theme for information processing in groups. *Group Processes & Intergroup Relations*, *3*(2), 123–140.

Thompson, L. & Fine, G.A. (1999). Sociall shared cognition, affect, and behavior: A review and integration. *Personality and Social Psychology Review, 3(4),* 278-302.

Valdesolo, P., Ouyang, J., & DeSteno, D. (2010). The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of Experimental Social Psychology, 46(4),* 693-695.

van den Bossche, P., Gijselaers, W., Segers, M., Woltjer, G., & Kirschner, P. (2011). Team learning: Building shared mental models. *Instructional Science*, *39*(3), 283–301.

Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C. J., & Hasson, U. (2017). Same story, different story. *Psychological Science*, 095679761668202.

# Chapter 2 - Methods of fNIRS data quality assessment for naturalistic social experiments

## Introduction

Over the last two decades, social neuroscience has grown into a robust field studying the psychological, neural, and physiological processes of humans in a social context – how they think about and interact with other humans, and how the social world influences human development and function. Yet, testing these social phenomena in their most naturalistic form within the lab has been a struggle for past neuroimaging work. The technological limitations of popular neuroimaging methods like fMRI and EEG require that research subjects be mostly immobile while participating in experiments, and mostly alone while doing so. Subjects can view socioemotional stimuli on a computer screen or push keyboard buttons to engage with a task, but this is still far removed from the richness of the real-world situations these experiments try to emulate.

Functional near infrared spectroscopy (fNIRS) is an emerging neuroimaging method that can address these concerns. Yet, it is also a relatively underdeveloped technology compared to other neuroimaging modalities, so there is still room for improvement in processing and analyzing pipelines to best use fNIRS for truly social neuroscience research. In this chapter, I will review how fNIRS works compared to other neuroimaging modalities, why these differences give fNIRS a unique position in social neuroscience research, and ultimately why it was used as the neuroimaging technology for this dissertation. Then, I will describe signal evaluation developments I have made to increase fNIRS' utility for research into social communication and interaction.

## How fNIRS Works

Functional near infrared spectroscopy is a noninvasive neuroimaging device that tracks and records brain activity. Similar to fMRI, fNIRS relies on the BOLD response to do this – the blood oxygen level dependent signal that occurs when localized populations of neurons fire, consume oxygen, and thereby require more oxygen to be pumped to the area in order to continue operating. This oxygen arrives by hitching a ride on the back of a protein called hemoglobin. Both fMRI and fNIRS can detect changes in hemoglobin concentrations caused by neuronal firing, due to physical differences in the oxygenated and deoxygenated hemoglobin molecules (HbO and Hb, respectively).

In fMRI, the differences are magnetic - Hb distorts a magnetic field created by the MR scanner more than HbO does, so recording the strength of this field across the brain enables the researcher to find when an increased ratio of HbO to Hb is present at each area of the brain as a mark of brain activity. In fNIRS, the important property of HbO and Hb is optical. HbO and Hb have different absorption spectra, meaning that when light of some wavelength is projected through a medium composed of one of those compounds, different amounts of light are absorbed depending on which compound the light is passing through. Light in the visible red and near infrared wavelength range (~700-900 nanometers) can pass through skin and bone fairly easily, so by projecting this sort of light into the head and measuring how much is reflected back out, fNIRS can detect concentrations of both HbO and Hb independently. This process is akin to a pulse oximeter used at the doctor's office, but scaled up to many different light emitters and detectors spread across the head to record HbO and Hb concentrations at many brain locations. A more detailed discussion of the biophysics of fNIRS can be found in Ferrari, Mottola, & Quaresima (2004) and Scholkmann et al. (2014).

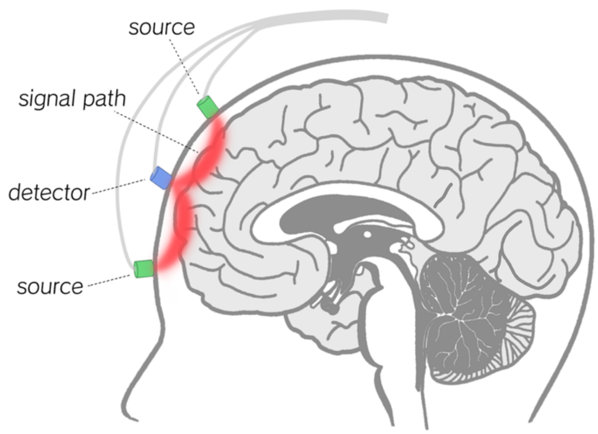



**Figure 1** – Schematic of a simple fNIRS set up, with infrared source and detector wires set up over the prefrontal cortex. Light travels through the cortex from sources to detectors in a banana-shaped path, recording a channel of data between each source-detector pair.

The fNIRS machine itself is generally the shape of a small box, varying in size between that of a textbook to that of a microwave. The machine interfaces with the human head via bundles of cables called optodes – wires that emit near infrared light (sources), and wires that detect incoming light (detectors). One source-detector pair creates a "channel," in which one HbO/Hb time series recording is made of whichever brain area is located between the source and detector.  Channels can be thought of as large MRI voxels located at that fNIRS channel. In order to facilitate the projection of light into head and detection of light reflected back out, optodes are typically held in place on the head via a stretchy cap or a rigid frame. The number of channels in an fNIRS recording is determined by the optode layout on the head and the number of optodes available on a particular machine. This number typically ranges from around a dozen in studies of targeted neural areas to more than one hundred in high density configurations.


**Figure 2 –** Examples of fNIRS head layouts and machines from various production companies. A) High density, full brain coverage with the ETG-4000 unit from Hitachi (hitachimed.com). B) A lightweight, prefrontal-specific layout with the OctaMon unit from Artinis (artinis.com). C) A hyperscanning study imaging multiple concurrent participants with the NIRScout unit from NIRx (nirx.net). D) Backpack design of the mobile LIGHTNIRS unit from Shimadzu (shimadzu.com).

## Advantages of fNIRS

Due to its technological design, fNIRS differs from other neuroimaging modalities in important ways. Table 1 summarizes how fNIRS compares to other common techniques with respect to a variety of research concerns. Notably, fNIRS is uniquely tolerant of participant motion. Optodes are affixed to a participant's head, and thus move through real space in the same way the head moves through real space. This makes any head motion irrelevant to the measurement system between the source and detector. In contrast, the MRI and PET environments depend on a magnetic field or positron camera system surrounding the outside of a person's head and thus require the head position to remain stationary in order to localize activity within the field to a particular brain location. Additionally, head muscle movements will not overwhelm the cerebral signal in fNIRS as it does in EEG. So long as any body motion does not shift an optode's secured position on the head, participants can sit up, talk, gesture, and even walk or exercise while wearing an fNIRS apparatus.

**Table 1 – Comparison of fNIRS to other imaging modalities**

| | **fNIRS** | fMRI | EEG | PET |
|---|---|---|---|---|
| *Signal depth:* | ~1.5cm into cortex | Full brain | Cortex | Full brain |
| *Spatial resolution:* | ~1cm | ~1mm | 4-8cm | ~4mm |
| *Sampling rate:* | 1-200Hz | <= 1Hz | 200-1000Hz | Minutes to hours |
| *Cost:* | $10k-400k | $1-7 million, + several hundred $ per scan | $5k-200k | $1-3 million, + several hundred $ per scan |
| *Portability:* | Portable machine, few accessories | Stationary | Portable machine, many delicate accessories | Stationary |
| *Motion sensitivity:* | Only sensitive if optodes move on scalp | Participants cannot move head | Participants cannot move muscles | Participants cannot move head |

| | | | | |
|---|---|---|---|---|
| | | | in/around head | |
| *Participant comfort:* | Snug cap, but participants can move around, no safety risk | Participants must remain still, loud machine, safety risks | Participants must hold upper body still, messy gel applied to head, no safety risk | Participants must remain still, loud machine, injection required, safety risks |

Another major advantage of fNIRS is its general usability, which can be broken down into its cost, portability, and comfort to the participant. It is an order of magnitude more affordable to acquire than fMRI or PET, and does not require any additional costs to record data besides reimbursement to a subject for their research participation, as in a behavioral study. No specific qualifications are needed to operate fNIRS beyond research training with the system, so no additional doctor or specialized technician is required on site. The machine is portable with limited accessories required for traveling – simply a computer for recording data, a power source, and the head cap or other optode positioning system. In contrast, EEG can be portable, but high quality data collection requires a dense and delicate electrode cap with associated conductive paste and cleaning materials. The more user-friendly EEG systems trade spatial resolution and signal quality for increased usability. PET and MRI machines are very large and must be secured to the floor in specially designed rooms. Finally, the only potentially uncomfortable aspect of fNIRS imaging is the tightness of the caps used to hold the optodes to the head. Participants are able to move comfortably during an experiment, and no gel or other liquid needs to be added to their head to improve optode-scalp contact as in EEG. There is also no substance to inject into a participant like with PET, and no safety concerns such as gamma radiation in PET or ferromagnetic implants in an MRI.

fNIRS is limited in terms of which areas of the brain it can record signal from - data in fNIRS is restricted to 1-2cm of surface cortex due to the fact that light scatters and dissipates too much to be usable in deeper layers of tissue. This is more than adequate to record functional areas on the cortical surface such as medial prefrontal cortex, motor cortex and the temporoparietal junction, but fNIRS is not sensitive to activity in deeper structures like the limbic system, orbitofrontal cortex, and cingulate cortex. The spatial resolution and sampling rate of fNIRS are also potential limitations, depending on the requirements of a particular experiment. While better than EEG, fNIRS is less spatially resolved than fMRI and PET. The signal can be reliably localized to about 1 centimeter, with 3D tomographic mapping capabilities only available in very dense optode layouts. Additionally, the sampling rate of fNIRS can range between 1-200Hz, which is better than all but EEG. Yet, fNIRS still measures the hemodynamics response, which is an inherently slower signal than direct neuronal firing.

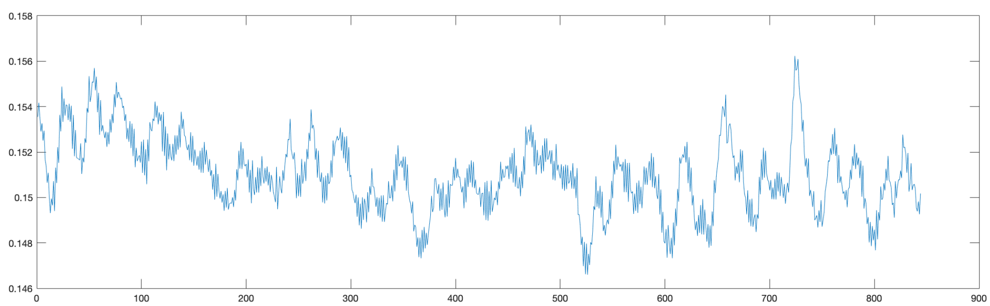
Due to these characteristics, fNIRS can excel in particular niches of social neuroscience that require participant motion, such as studies of social interactions as done in this dissertation.

## Challenges in fNIRS Data Processing

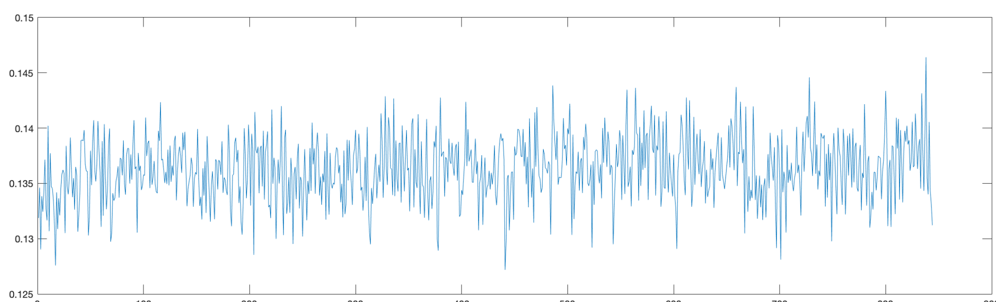
fNIRS holds great promise for improving the external validity of social neuroscience research. However, the technology is relatively underutilized and underdeveloped compared to other neuroimaging modalities such as fMRI and EEG – according to Web of Science, it wasn't until 2012 when the number of fNIRS research publications per year exceeded 100 (in contrast, fMRI sees thousands of publications each year). Due to this fact, open questions still remain about the best way to clean and prepare fNIRS data for analysis.

For example, the quality of fNIRS data suffers greatly when detected infrared light signal coming out of the head is weak, and/or when that signal is contaminated by ambient light from the surrounding environment. This problem does not occur when optodes make unobstructed connection with the scalp, but participant hair sometimes gets in the way of this connection. The thicker and darker these strands of hair are, the more infrared light is blocked (Katus et al., 2019; Khan et al., 2012). When too much infrared signal is blocked this way, the resulting data is a time series of white noise instead of hemodynamic fluctuations (Figure 3b). Engineering efforts to increase fNIRS signal intensity and optode design are ongoing, but currently this sort of signal disruption is not uncommon in fNIRS research, as an estimated 57% of the world population has dark brown to black hair on the head (Panhard, Lozano, & Loussouarn, 2012).

Another major challenge for analyzing fNIRS data is the presence of motion artifacts. As previously mentioned, fNIRS is robust to participant motion, but only so long as the optodes do not shift position on a participants' head. If this occurs, large shifts in light intensity are recorded as the detector optode momentarily lets in ambient light or either optode changes in amount of connection with the scalp (Figure 3c). Best fNIRS practices involve tightly fitting optodes to the head, stabilizing optode cables so that their movement doesn't shift optode position, and/or using wireless fNIRS devices. However, motion artifacts may still occur and negatively impact data quality. The most egregious artifacts are large signal variations over very short periods of time – signal patterns that do not occur in the relatively slow drifting hemodynamic signal.  However, while the presence of an artifact is usually easy to detect, automatic removal of these patterns from data time series in order to reveal true signal beneath them is not trivial. Extensive work on detecting and removing these artifacts from collected data has been conducted (e.g., Barker, Aarabi, & Huppert, 2013; Chiarelli et al., 2016; Fishburn et al., 2019; Molavi & Dumont, 2012), but as of yet no standard process has emerged that reliably removes motion artifacts from all datasets. All methods attempted for this dissertation either left some large motion artifacts unaffected, or imposed occasional new perturbations that do not correspond to brain signal.



A



B

C

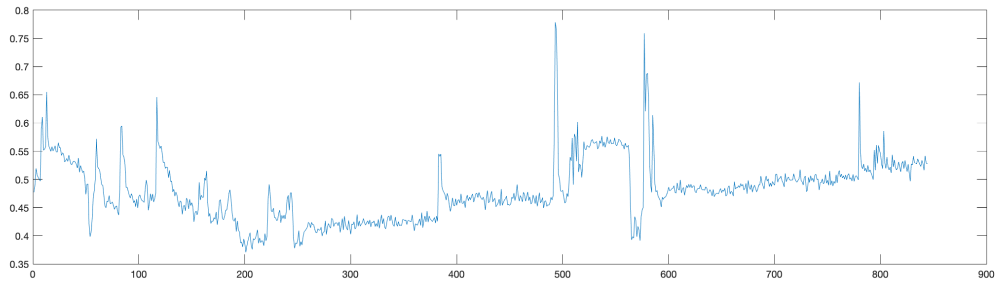



**Figure 3** – Examples of raw fNIRS data time series. A) High quality data is relatively autocorrelated with no quick perturbations. In particularly clean data, a heartbeat signal is visible, corresponding to small fluctuations at ~ 1Hz. B) When optodes have poor connection to the scalp or hair is blocking light travel between source and detector, the predominant signal evident in the recorded data is white noise – random high frequency fluctuations with few to no slow frequencies present. C) In the event an optode shifts position on the scalp, motion artifacts may be introduced that look like brief spikes (large change with return to previous baseline), discontinuities (large change with change in new baseline), or periods of volatility change (variance in data over time changes). The amplitude of these artifacts is usually several times that of the standard deviation of the data without motion, but not always. The primary characteristic of a motion artifact is that these changes occur very quickly (within a couple samplings of data), while real hemodynamic signal typically takes several seconds to change +/- 1 standard deviation.

It is outside the scope of this dissertation to develop additional signal processing methods for removing these sorts of artifacts in fNIRS data. Instead, we followed an alternative approach – evaluate how adverse the possibility of these artifacts may be for our data, and reject data channels based on a data susceptibility threshold. While the goal of signal cleaning is to eliminate the need for dropping any data, the state of fNIRS preprocessing is not mature enough to ensure this outcome for many real datasets, and thus further evaluation methods are needed to judge how good of a job the preprocessing did before the data can be trusted as a true representation of participant brain activity. The following methods have been implemented in open source Matlab and Python code called *preprocessingfNIRS* along with other batch preprocessing functions, available at github.com/smburns47/preprocessingfNIRS.

## Evaluating Noisy Channels Before Preprocessing

When there is a poor connection between an optode and the scalp, the recorded signal will either be swamped by ambient light, or not be recorded at all such that ambient white noise is the only signal present. This results in a data time series dominated by high frequency noise from which real signal cannot be reliably recovered. Thus, it is important to detect and exclude noisy channels like this from analysis.

Though the difficulties of collecting quality fNIRS signal in thick dark hair are well known, there are few explicit guidelines in the literature about how to tell if a channel has enough strength in the recorded hemodynamic signal. Thus, we developed a method of automatically evaluating how noisy a channel was based on the shape of its Fourier power spectra. It is well known from early work with fMRI that the power spectral density of a typical hemodynamic response in human cortex is $1/f$, or inversely proportional to the frequency of the signal (Zarahn, Aguirre, & D'Esposito, 1997). In other words, slow frequencies are more prominent in hemodynamic signal than fast frequencies (Figure 4a). In contrast, white noise as in noisy fNIRS data channels has a flat power spectrum (Figure 4c). Thus, we can calculate the power spectrum of an fNIRS data channel and evaluate the extent to which it resembles the canonical $1/f$ density shape in order to judge the likelihood that this channel has recoverable brain data. To do this, we calculate a modified version of the quartile coefficient of dispersion (Bonett, 2006). The power spectral density plot of a channel is divided into quartiles, which are then summed. Then the variation between the slowest frequency quartile and fasted frequency quartile is computed as

$$C = (Q_{slow} - Q_{fast}) / (Q_{slow} + Q_{fast})$$

This coefficient C will approach 1 as the magnitude of $Q_{slow}$ exceeds that of $Q_{fast}$, and will be 0 if the magnitude of the summed frequency densities in these quartiles are the same. Thus, a larger coefficient C means that the data channel resembles that of a true hemodynamic signal, and a smaller coefficient C means the data is likely to be white noise.

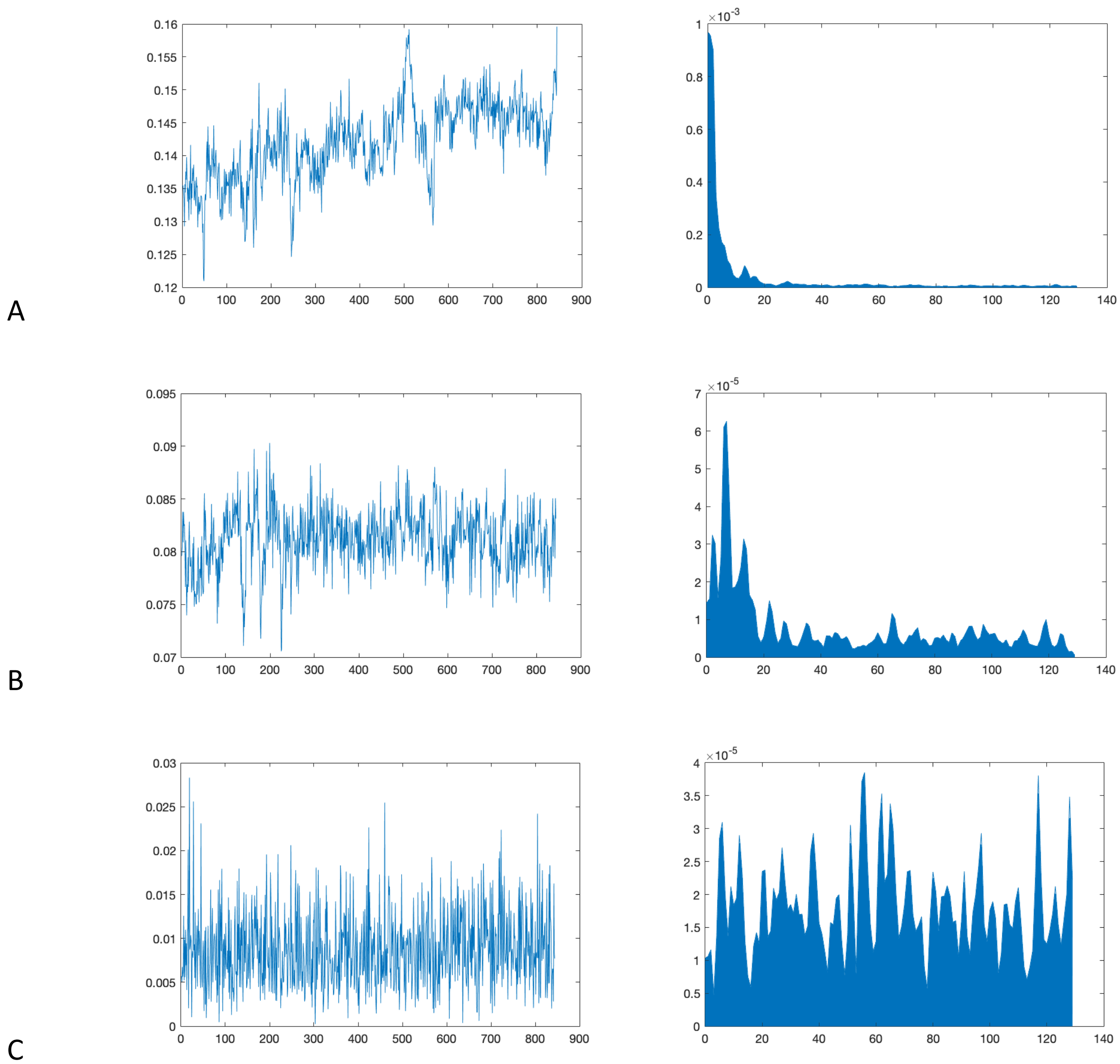
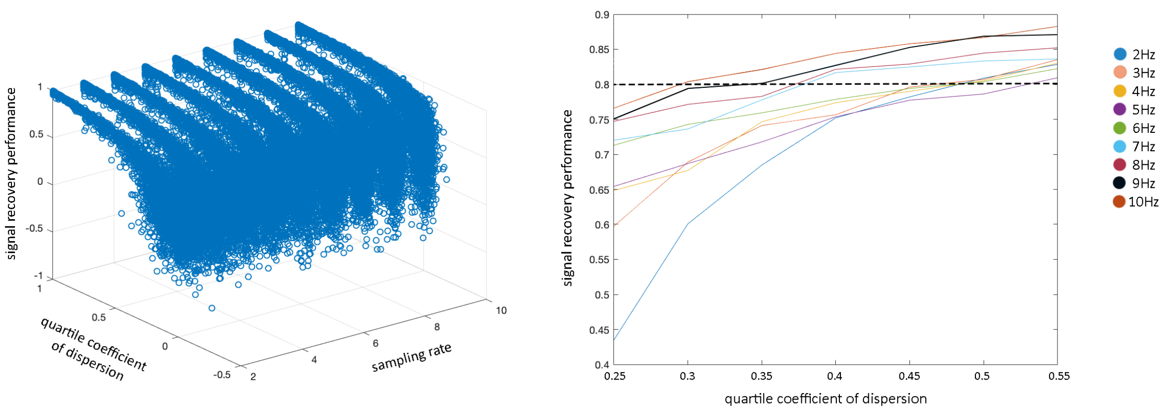



**Figure 4** – Raw data time series and power spectral density (PSD) plots for three different fNIRS channels, expressed in terms of signal power (y axis) per Nyquist frequency (x axis). A) Clean data, where a clear 1/*f* shape can be seen in the PSD plot; quartile coefficient of dispersion C = 0.99. B) Data with some noise, but from which signal can still be recovered as lower frequencies still dominate in the PSD plot; C = 0.66. C) Data that is completely saturated by noise, with seemingly random values throughout the PSD plot; C = 0.05. The C threshold for this data sampling rate was 0.54.

We set a default threshold of

$$C_{thresh} = 0.6 - 0.03*\text{sampling rate}$$

for determining if a channel still had recoverable signal or not based on simulations of how much noise could be imposed on a time series before the true signal could no longer be reliably recovered at correlation r=0.8 with bandpass filtering (Figure 5). If a channel has a coefficient C larger than this threshold, it is accepted as usable data for further preprocessing. If not, it is rejected from any further analysis.




A B

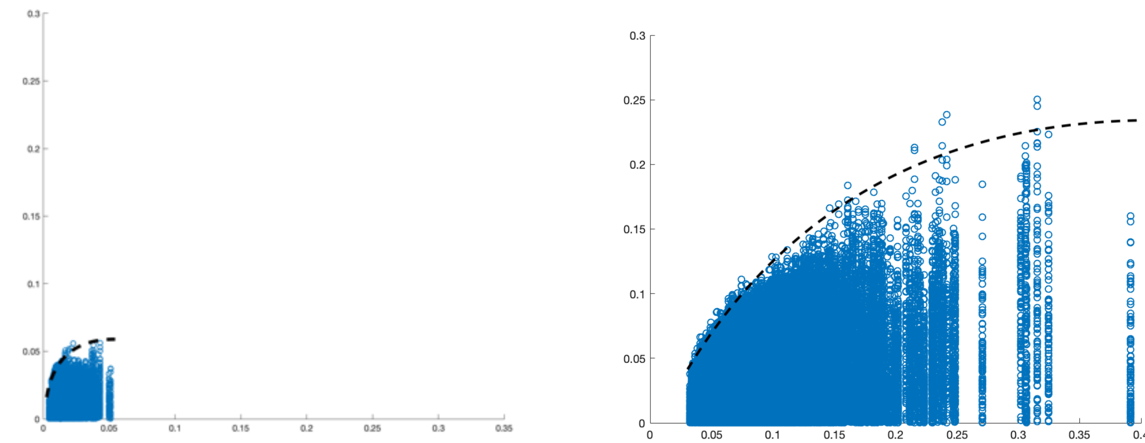
**Figure 5** – A) Over 50,000 simulation iterations, noise was superimposed over clean neural data of various sampling rates. The quartile coefficient of dispersion C was computed for each of these noisy time series, and then the data was bandpass filtered to try and recover the true signal under the noise. B) Lower 95% confidence bound of recovery strength for each sampling rate at various values of the quartile coefficient. It was generally easier to recover signal in time series of higher sampling rate.

Applying this method to the datasets in this dissertation, approximately 20% of channels in the data from Chapter 3 were rejected as too noisy, while approximately 10% of the channels in the data from Chapter 4 were rejected as too noisy. Channels were more often removed at the lower back of the head, where human hair tends to grow the thickest, and were rarely removed over the forehead. This means that analyses done on more dorsal and anterior channels in these datasets have somewhat more statistical power than more ventral and posterior channels.

## Evaluating Questionable Channels After Preprocessing

After identifying clean and noisy data channels, we preprocess our data with the most common filtering approaches – detecting motion artifacts based on a >5 SD change in value over a 1-second timespan, rescaling these detected spikes/discontinuities/periods of volatility to match that of the preceding data, and then bandpass filtering to a window 0.008-0.2 Hz that encompasses the frequencies of hemodynamics activity associated with cognitive function (Zuo et al., 2010). This is a conservative filtering approach in that it removes fluctuations that are certainly not attributable to brain activity, but may leave in data features that are of questionable origin. We use this approach to ensure that no real brain data is inadvertently removed during filtering, but spurious artifacts might still remain in the data. Again, there is currently no explicit method in the literature for evaluating how well a motion filtering pipeline worked within a specific dataset, so we developed a way to do this based on the impact that removing these potential artifacts might have on later statistical analyses.

To do so, we first identify questionable artifacts remaining in preprocessed data. Since definite motion artifacts were previously defined as changes >5 SD in 1 second, a questionable artifact here may be a weaker feature: >3SD in 2 seconds. Once those questionable artifacts have been identified, we proceed with the assumption that the effect of their presence is localized in the data time series, and removing them within this contained window may change the statistics computed on the otherwise clean data to variable degrees. It is not possible to measure the extent of this change when a true synchrony value is unknown, but it can be measured by estimating the extent to which it changes the correlation of the time series with itself, with and without the artifact. In a simulation comparing this kind of change in autocorrelation to a change in correlation with another time series when fake artifacts were injected into otherwise clean time series, we found that small spikes that vary within the range of the rest of the data have a negligible effect on a synchrony estimation – the autocorrelation rarely changes by more than 0.1, and even when it does, correlation with another time series does not change by more than 0.08 (Figure 6a). However, the effect of large spikes that exceed the range of the rest of the data do meaningfully affect a synchrony estimate. In these instances, if the autocorrelation doesn't change much, then the synchrony estimate doesn't change much either. But both of these *can* change by quite a bit (Figure 6b & 6c). The same can be repeated for other statistical estimates such phase synchrony between time series or the beta parameter within a general linear model. It is worth noting for this dissertation that in general, the same artifacts affect a self-self and self-other phase synchrony estimate less than they do a correlation estimate. This is likely because phase synchrony as a calculation is more robust to outlier values within a time series. However, the same general conclusions exists here, in that small autocorrelation changes don't result in meaningful changes in the statistical estimate of interest, but large autocorrelation changes might.



A – effect of small spikes B – effect of large spikes

**Figure 6 –** The effect of inserting one artifact of various types in data on the autocorrelation of that data (x-axis) and the correlation of that data with another clean time series (y-axis). Dotted lines represent the maximum the correlation may change given an observed change in autocorrelation. A) The effect of a small spike within a data time series, defined as a change over 2 seconds with an amplitude matching the rest of the data's range. When a small spike is imputed into a time series, neither the autocorrelation nor correlation value with another time series change by much. B) The effect of a large spike, defined as a change over 2 seconds with an amplitude of 2 times the range of the rest of the data. This effect of this type of artifact is sometimes negligible on the autocorrelation of the data, but can sometimes change the autocorrelation by quite a bit. If the autocorrelation doesn't change by much, then the correlation with another time series doesn't change by much. But if the autocorrelation does change by a meaningful amount, then it is possible that the correlation with another time series changes by a meaningful amount as well.

Thus, within a time series where questionable artifacts are detected, we can smooth over these artifacts to remove them and then measure how much the data's autocorrelation changed based on this smoothing. The density plot in Figure 5b was used as a guide to identify a threshold autocorrelation change that would be allowed given that we do not want the synchrony estimate to change by more than some value (0.1 used as default). If removal of a questionable artifact resulted in a large change of autocorrelation, then that channel is marked as too uncertain – this artifact may or may not be real signal we are interested in, but the consequences of making the wrong call on that decision are too great. Uncertain channels are thus not included in later statistical models.

In the dataset in Chapter 3, because phase synchrony was used as the neural synchrony estimate of interest, few channels were identified as questionable. On average, less than 1% of a participants data channels were removed in this way. In the dataset in Chapter 4 this number was slightly higher at 2% of channels per person, likely due to the greater motion involved in a natural conversation. Ultimately however, these outcomes seem to suggest that the motion filtering pipeline performed relatively well in our datasets.

## Conclusion

Among neuroimaging modalities, fNIRS is particularly well suited for nautralistic experiments of social interactions. However, given that optimal filtering techniques for fNIRS data are still in development, for this dissertation we aimed to develop approaches to measuring the quality of a data time series and the performance of a filtering pipeline on that time series in order to make judgements about which data channels can be used in analysis. The outcomes of these evaluation approaches should make the analyses in the later chapters of this dissertation less subject to signal noise, and thus more accurate at inferring neural dynamics patterns in social communication and interaction. In the future we plan to compare the analyses of datasets with and without these approaches in order to measure overall what impact such evaluation techniques make on the inferences that can be derived from fNIRS data.

## References


Barker, J.W., Aarabi, A., & Huppert, T.J. (2013). Autoregressive model based algorithm for correcting motion and serially correlated errors in fNIRS. *Biomedical Optics Express, 4(8),* 1366.

Bonett, D.G. (2006). Confidence interval for a coefficient of quartile variation. *Computational Statistics and Data Analysis, 50(11),* 2953-2957.

Chiarelli, A.M., Maclin, E.L., Fabiani, M., & Gratton, G. (2016). A kurtosis-based wavelet algorithm for motion artifact correction of fNIRS data. *NeuroImage, 112,* 128-137.

Ferrari, M., Mottola, L., & Quaresima, V. (2004). Principles , techniques , and limitations of near infrared spectroscopy. *Canadian Journal of Applied Physiology*, *29*(4), 463–487.

Fishburn, F.A., Ludlum, R.S., Daidya, C.J., & Medvedev, A.V. (2019). Temporal derivative distribution repair (TDDR): A motion correction method for fNIRS. *NeuroImage, 184,* 171-179.

Katus, L., Hayes, N.J., Mason, L., Blasi, A., McCann, S., Darboe, M.K., de Haan, M. Moore, S.E., Lloyd-Fox, S., & Elwell, C.E. (2019). Implementing neuroimaging and eye tracking methods to assess neurocognitive development of young infants in low- and middle-income countries. *Gates Open Research, 3,* 1113.

Khan, B., Wildey, C., Francis, R., Tian, F., Delgado, M. R., Liu, H., … Alexandrakis, G. (2012). Improving optical contact for functional near-infrared brain spectroscopy and imaging with brush optodes. *Biomedical Optics Express*, *3*(5), 1820–1828.

Molavi, B. & Dumont, G.A. (2012). Wavelet-based motion artifact removal for functional near-infrared spectroscopy. *Physiological Measurement, 33,* 259-270.

Panhard, S., Lozano, I., & Loussouarn, G. (2012). Greying of the human hair: A worldwide survey, revisiting the '50' rule of thumb. *British Journal of Dermatology,167,* 865-873.

Scholkmann, F., Kleiser, S., Metz, A. J., Zimmermann, R., Mata Pavia, J., Wolf, U., & Wolf, M. (2014). A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *NeuroImage*, *85*, 6–27.

Zarahn, E., Aguirre, G. K., and D'Esposito, M. (1997). Empirical analysis of BOLD fMRI statistics. I. Spatially unsmoothed data collected under null-hypothesis conditions. *Neuroimage 5*, 179–197.

Zuo, X., Di Martino, A., Kelly, C., Shehzad, Z.E., Gee, D.G., Klein, D.F., Castellanos, F.X., Biswal, B.B., & Milham, M.P. (2010). The oscillating brain: Complex and reliable. *NeuroImage, 49(2),* 1432-1445.

# Chapter 3 - Effects of similar experience and identity on accurate interpersonal understanding

## Introduction

When sharing a personal story with a friend, they may appear attentive and caring. However, it may not always be the case that they actually understand what happened to you and what it means to you. Instead, they may interpret your story and your feelings in an entirely different way than you intend. Folk wisdom asserts that in order to really understand what someone is going through, you need to “walk a mile in their shoes” and “see the world through their eyes.” Within these metaphors, the message is that you come to understand someone through shared experience and first-hand knowledge about that experience. In accordance with this guidance, there is substantial evidence that people prefer to disclose personal events to and seek social support from others who have had the same experience (Hoyt et al., 2010; Simich, Beiser, & Mawani, 2003; Suitor, Pillemer, & Keeton, 1995). People with a shared experience also express greater empathy compared to people without the shared experience (Barnett, 1984; Barnett, Tetreault, & Masbad, 1987; Batson et al., 1996; Eklund, Andersson-Straberg, & Hansen, 2009; Hodges et al., 2010; Preis & Kroener-Herwig, 2012).

Yet, evidence is lacking about whether shared experience meaningfully impacts how *accurate* of an understanding one can develop in response to another’s self-disclosure. Most research on shared experience focuses on affiliation behaviors or empathy expression rather than communication success or empathic accuracy. For those that do, the operationalization of accurate understanding is often loosely defined as simply a subjective rating of whether the listener in a conversation "got it," rather than any quantitative measurement of how much and what kind of content understanding the listener received (Banks, Berenson & Carkhuff, 1967; Haley & Dowd, 1988; Kirk, Best, & Irwin, 1986; Robiner & Storandt, 1983).

Furthermore, the current literature rarely makes a distinction between shared experience and shared identity. I.e., can an increase in empathy and understanding be explained because I have gone through the same events and you and know how I reacted to it, or because I am a similar person to you and therefor better know how you will react to things? Ostensibly over time, accumulated shared experiences between people can make them more similar to each other. For example, a shared history of alcoholism can extend beyond just a set of similar events that happened to two people, but can be a dominating life experience that shaped the people's perspectives and habits in particular ways (Kirk, Best, & Irwin, 1986). Yet on the level of an individual interaction, where there is an accuracy advantage of similarity between people, it is not clear if that advantage is driven by the shared experience of a particular event being described, or is driven by a more general identification by the listener with the speaker (whether because shared identity affords more accurate inferences about the speaker, or simply motivates the listener to attend better). Prior work on the relationship between shared identity and empathy/understanding is mixed, (Grover & Brockner, 1989; Heinke & Louis, 2009; Krebs, 1975; Pietromonaco, Rook, & Lewis, 1992; Verhofstadt et al., 2008; Westmaas & Silver, 2006), but this has not been directly compared to the effect of similar experience.

Therefore, the current study aims to directly compare the effect of similar experience and similar identity on a listener's ability to understand and empathize with a speaker's personal emotional stories. Similar experience and similar identity will be measured explicitly and separately in order to parse which may drive an effect of understanding and/or empathy. We aim to assess shared understanding in three ways – 1) the extent to which a listener can retell all the semantic content of the narrative they just heard; 2) the similarity between the speaker and a listener's moment-to-moment ratings of the story's emotional content; 3) the synchronization of brain activity between the speaker and a listener while the story is being told and encoded, as recent developments in neuroscience illustrate how shared understanding of a narrative between people can be identified via the amount of synchronous brain activity between them (Dikker et al., 2014; Honey et al., 2012; Liu et al., 2017; Nguyen, Vanderwal, & Hasson, 2018; Stephens, Silbert, & Hasson, 2010; Yeshurun et al., 2017). We will also measure the extent to which a listener experiences the same emotions from the story as the speaker, how much empathy they feel toward the speaker, and their empathic accuracy for how the speaker reported feeling. By assessing similar experience and identity in tandem and measuring shared understanding in more quantitative fashion than previous literature, we hope to discover what about someone can enable them to understand and empathize with another person more effectively.

## Methods

### *Participants*

*Speaker* - Recruitment for this study was conducted in two phases. In the first phase, 12 initial participants were recruited with a university-wide email distribution to record a variety of personal stories on camera in the lab while their brain activity was recorded. We wanted the speakers to talk about personal events that varied in terms of how strong identity cues within the story may be, so we recruited speakers who had experienced both an event with low identity salience (death of a pet or romantic break up), and an event with high identity salience (status as a sexual minority being outed without permission). Six of these speakers identified as female, one as male, two as transgender, and three as gender nonbinary. Five identified as white, two as Hispanic/Latinx, three as Asian/Pacific Islander, and two as mixed ethnicity. The average speaker age was 21.67 (*SD* = 4.05). From this set, one speaker was chosen to show to listeners whose videos were relatively similar in length, similar in the strength of self-reported negative affect, and whose brain activity recordings during the stories were high quality. This speaker identified as a white female and was 21 years old. The negative personal stories she told recounted the time her dog died and a time her sexual identity was outed without her permission.

*Listeners* - In phase 2 of the study, 120 participants were recruited with a university-wide email distribution to watch the chosen speaker's recorded stories while their own brain activity was recorded. The recruitment process asked participants questions relating to their experiences with pets and their own sexual identity to ensure a variety of experiences and identities relative to the speaker were represented in the listener sample. To eliminate any cross-gender communication effects in this study, all recruited listeners identified as female. The average age was 20.27 years (*SD* = 2.07). Forty-one listeners identified as Asian/Pacific Islander (34.16%), 26 as White (21.66%), 26 as mixed ethnicity (21.66%), 22 as Hispanic/Latinx (18.33%), 3 as Black (2.5%), and 2 as North African/Middle Eastern (1.66%).

*Materials*

Neural activity in both the speaker and listeners was recorded using functional near infrared spectroscopy (fNIRS). This imaging method uses infrared light to measure oxygenated blood flow in the brain's cortical surface as an indirect measure of brain activity (see Chapter 2 of this dissertation for more information). The specific equipment used was a NIRScout imaging unit from NIRx Technologies (nirx.net). This unit has 32 source and 32 detector optodes, which were secured into stretchy head caps and positioned to create 108 separate data measurement channels covering nearly the entire head (Figure 1). This positioning was standardized over all participants using the 10-10 UI external positioning system. Light intensity data was collected at wavelengths 760 and 850nm, with a sampling rate of 1.95Hz.

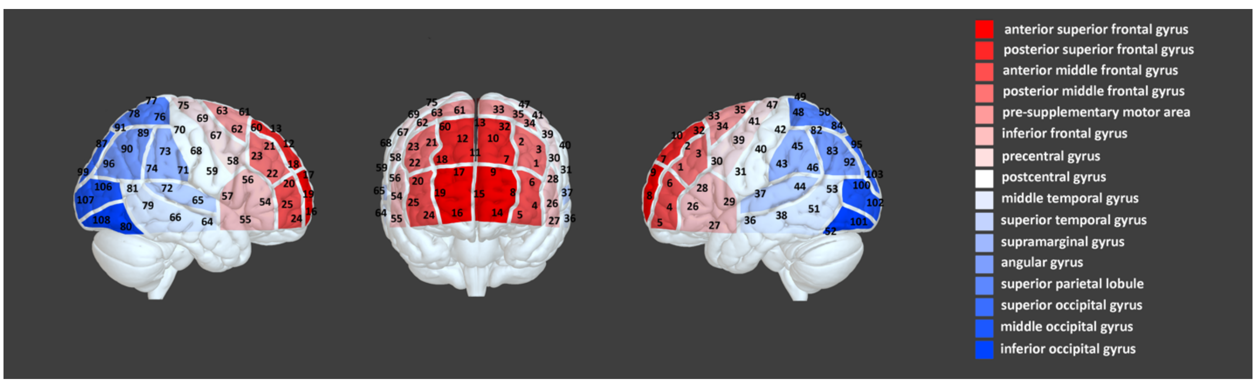



**Figure 1** – Channel location projected to the cortical surface, and groupings of channels to form 32 anatomical regions of interest.

Audiovisual videos of speakers' stories were recorded on a Panasonic LUMIX G7 Mirrorless 4k camera and displayed to listeners in full screen resolution on a 21.5" Apple iMac desktop computer. Speakers' and listeners' continuous affect ratings were recorded on a separate Lenovo IdeaPad laptop running the real-time media annotation software CARMA

(Girard, 2018). All other experimental responses were recorded via survey hosted digitally on Qualtrics.

*Procedure*

*Speaker* - Upon arrival to the lab, speakers were seated in front of a computer and video camera and were fitted with the fNIRS equipment. Once fitted, the speakers were asked to recall 5 different stories on different topics, ignoring a topic if they had not experienced that event before. These topics were chosen by the researchers to elicit differently-valenced emotions, and to prompt different perceptions of shared experience/personal similarity in listeners – their first day at college, a time their pet died, a time they went through a romantic break up, a time their sexual identity was outed, and a time they were shown great kindness. For this study however, analyses were only planned for the negatively-valenced stories. Before each story, the speakers were given instructions on which story to tell and how long to make it (3-8 minutes), and then were left alone in the experiment room to plan how they would tell their story. Once decided, they notified the experimenter that they were ready to begin by ringing a desk bell in the room. At this point, the experimenter began the camera and fNIRS recordings, and then left the room again while the speaker told their story. Once finished, the speaker again notified the experimenter with the desk bell. In between each story, the speaker completed the short form of the Positive and Negative Affect Schedule survey (PANAS; Mackinnon et al., 1999) for how they were feeling right now after recounting the story, and how they felt at the time of the story. After finishing all stories, the fNIRS equipment was removed and the speakers re-watched their videos on a laptop computer. During this rewatching, they provided continuous affect ratings of the stories. Specifically, they were instructed to rate how positive or negative the content of the story was at each moment on a scale from +100 (positive) to -100 (negative).

*Listeners* - Upon arrival at the lab, listeners were seated in front of a computer and were fitted with the fNIRS equipment. Once fitted, listeners watched the four videos the chosen speaker had recorded, in the following fixed order – first day at college, death of a pet, outing of their sexual identity, and an act of kindness. This order was chosen so that the identity-salient video would not contaminate reactions to the pet loss video (no salient identity cues), and so that the session could end with a positive story. During video watching, participants' brain activity was recorded with fNIRS. They were left alone in the experiment room during the videos and rang a desk bell at the end to indicate to the experimenter when they were finished. In between each video, listeners answered survey questions in which they completed the PANAS for themselves, for what they thought the speaker felt at the time of the story, and for what they thought the speaker felt at the time of the video recording. Listeners were also asked to rate how much empathy they felt for the speaker on a 1-7 scale, how similar of an experience this was to one they've had before, how similar the speaker seems to them, and were asked to retype the story they just heard in as much detail as possible. After watching the videos for the first time, the fNIRS equipment was removed and the listeners again watched the videos a second time, during which they provided continuous affect ratings in the same manner as the speakers.

*Neural Data Processing*

Data collected with fNIRS was subjected to a preprocessing pipeline that progressed as follows – 1) automatic identification and removal of noisy channels with frequency spectra approach (see Chapter 2); 2) detection of motion artifacts within remaining channels, defined as periods of data with a greater than 5 standard deviation change in less than 1 second; 3) rescaling of data in artifact segments to remove spikes/discontinuities/volatility changes within the data; 4) bandpass filtering to 0.008-0.2Hz; 5) conversion of light intensity values to percent change in oxygenated hemoglobin concentration using the Modified Beer Lambert Law; 6) evaluation and rejection of any channels with large remaining spikes/volatility changes.

*Variable Definition and Statistical Analysis*

This study aimed to measure shared understanding, as defined by alignment across 3 different domains – semantic, affect perception, and neural. Each type of alignment calculation is a measure of similarity within that measurement domain between two participants – either the speaker and a listener, or a pair of listeners. Semantic alignment was calculated on the text of the story the speaker told and the text of each listener's retelling. The software package ALIGN was used to do this calculation by locating the position of each text within a pre-trained high-dimensional semantic space and then calculating the distance between these positions (Duran, Paxton, & Fusaroli, 2019). This semantic space was previously trained on 3 billion words in the freely available Google News corpus. The inverse of the resulting distance is thus a similarity score between the semantic content of each text and represents how much content within one text is present within the other (allowing for variation is highly similar vocabulary used to express the same meaning).

Alignment in affect perception was calculated as a correlation between the timecourses produced by each participant's continuous affect ratings. Thus, this measure does not tap into affective judgements of the overall video, but on the perceived affective valence and intensity of the content in each moment of the story. Similar fluctuations in these ratings indicate that two people similarly perceived the affective dynamics of the story, and would result in high affective perception alignment scores between them.

Neural data was used to compute another alignment measure. First, time courses of brain activity were recorded in each fNIRS channel on a participants head. To reduce the number of independent statistical tests performed on this data and increase the signal-to-noise of neural measurement, these time courses were combined to create 32 new time courses representing 32 anatomical regions of interest (ROIs) (Figure 1). Then, the similarity between the neural time course in one participant's ROI to that of the matching ROI in another participant was calculated using phase synchrony. While Pearson's correlation is a more common method of neural synchrony calculation (Nastase et al., 2019), phase synchrony is more robust to momentary time course artifacts that may occur in neural data while an individual is speaking and can provide moment-to-moment estimations of coherence (Glerean et al., 2012).

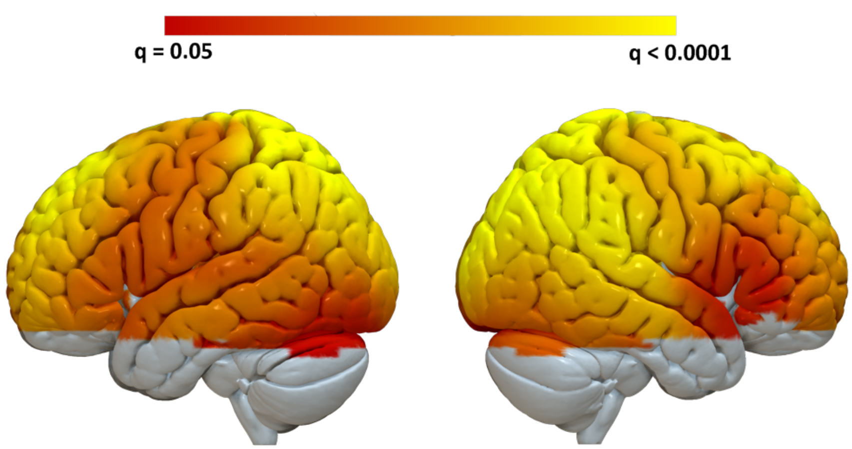
The final variables used in this study were four measures of relevant empathy outcomes – expressed empathy, emotional state matching between speaker/listener or listener/listener, and empathic accuracy. The empathy rating question within the survey was used as the measure of expressed empathy for each listener. Next, because the PANAS questionnaires completed by the speaker and listeners consisted of rating scales for ten different emotions, a correlation distance between the 10-item vectors of these ratings can be used to investigate similarity in emotion states. Emotion matching was defined as the correlation between the speaker's and listeners' self PANAS ratings, while empathic accuracy was defined as the correlation between the speaker's self PANAS ratings and those provided by the listeners on behalf of the speaker.

Each of these alignment and empathy scores were computed within each video. In the speaker-listener analyses, scores were computed between the listener and speaker for each listener. In the listener-listener analyses, scores were computed for each possible pair of listeners. Relationships between variables were determined using cross-classified multilevel models treating participant and video as random factors. For these analyses, only the data from the two negative stories were used, and any missing data was excluded from the model pairwise. The Satterthwaite method was used to calculate degrees of freedom and p-values were corrected for multiple comparisons and converted to q-values using the Benjamini-Hochberg approach (significance scores denoted by *q*). Effect sizes are reported as the marginal $R^2$ (variance explained by the fixed effect).

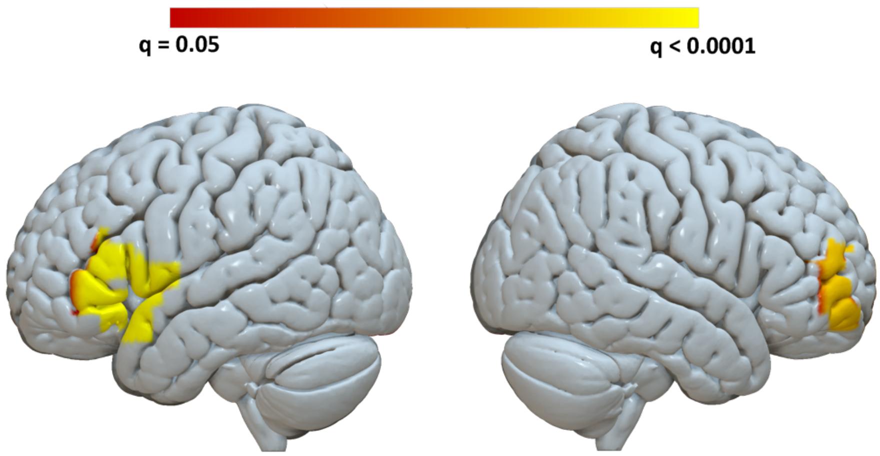
## Results

### *Baseline Neural Synchrony*

First, we examined the baseline amount of neural synchrony for all negative videos between listeners, and between the speaker and listeners, in order to evaluate how much neural processing was shared on average. Synchrony between each listener pair was computed, and the average compared to a bootstrapped null distribution of phase-randomized neural time courses. Between listeners, there was a significant amount of synchrony all across the brain, with the strongest synchrony occurring in the posterior superior frontal gyri, superior parietal lobules, and right superior occipital gyrus (Figure 2a). Between the speaker and listeners, average neural synchrony was computed with a lag, as speakers' neural encoding of their message tends to precede the verbal delivery of it (Liu et al., 2017; Stephens, Silbert, & Hasson, 2010). A lag of 4-10s was investigate, based on which value maximized synchrony for a given ROI. Speaker-listener synchrony was maximized at 8-10s in the significantly synchronous prefrontal areas. Baseline synchrony here was less extensive, but still significant in the left inferior frontal gyrus and right anterior middle frontal gyrus (Figure 2b).




A

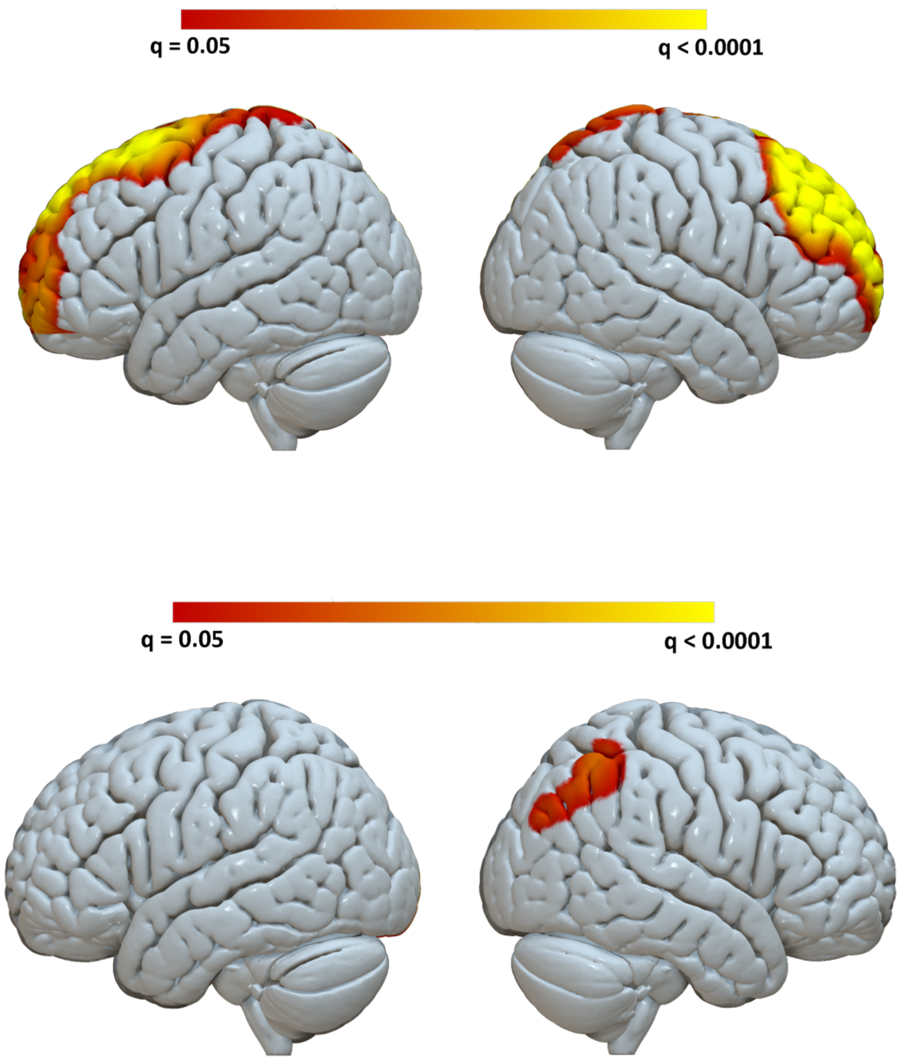



B

**Figure 2** – Average neural synchrony in all negative emotion videos A) between all listeners, and B) between the speaker and all listeners.

*Neural Synchrony and Psychological Alignment Measures*

Next, we examined how related the measures of semantic alignment, affect perception alignment, and neural alignment were to each other to see how much independent information they captured and to validate that neural alignment reflected shared understanding of the narrative. Across all listener pairs, semantic alignment was not related to affective perception alignment ($t(1, 10404) = -0.68$, $p = 0.50$). In other words, listeners who exhibited similar understanding of the semantic content they heard did not necessarily have a similar affective understanding of the story. However, semantic alignment between listeners did predict neural alignment across large areas of the prefrontal cortex, left pre-SMA, and right superior parietal lobule (Figure 3a). This indicates that listeners who had more similar semantic understanding of the stories had more synchronized neural activity to each other in these areas. There was also a significant relationship between affect perception alignment and neural synchrony in the right angular gyrus (Figure 3b).




A

B

**Figure 3** – Brain areas where amount of neural synchrony between pairs of listeners significantly corresponded with A) amount of semantic alignment, and B) amount of affect perception alignment.

Examining speaker-listener pairs, we did not find any significant relationship between neural synchrony and the psychological alignment measures. High alignment with the speaker in semantic or affective perception of the story did not predict strength of neural synchrony with the speaker.

*Effect of Similarity*

Lastly, we sought to investigate how much similar experience drove alignment and empathy scores, compared to similar identity. In other words, we tested the effect of listeners having a similar experience to the speaker against the effect of listeners perceiving themselves to be similar people to the speaker. While these variables are conceptually related, they were only moderately correlated in these data ($r = 0.42$), and can thus be interrogated as separate effects. First, we investigated the effect of similarity on alignment between listeners – i.e., did listeners who saw themselves as more similar to the speaker converge on a more canonical understanding of the story. Both predictors (similar experience and similar identity) were run in the same model in order to identify which type of similarity significantly related to the outcome variables over and above the effect of the other type. For semantic alignment, both similarity of experience and similarity of identity significantly predicted semantic alignment, but in differential ways (Table 2). Listeners with a more similar identity to the speaker had significantly more semantic alignment with each other, but listeners with more similar experiences to the speaker had significantly less semantic alignment with each other. In other words, listeners with a similar experience in their past had a more idiosyncratic understanding of the stories' semantics. Similar experience did not relate to affective perception alignment between listeners. When investigating the effect of similarity on alignment between the speaker and listeners, there were no significant relationships.

**Table 1 – Effects of similar experience  and similar identity on semantic and affective alignment**

| | Semantic alignment between listeners | Affective alignment between listeners | Semantic alignment between speaker-listener | Affective alignment between speaker-listener |
|---|---|---|---|---|
| Similar experience | **$t$(12629) = -6.95 $q$ = 1.17e-11*** $R^2$ = 0.0039** | $t$(10280) = 0.65 $q$ = 0.52 | $t$(189) = -1.99 $q$ = 0.095 | $t$(193) = -0.52, $q$ = 0.90 |
| Similar identity | **$t$(12770) = 16.74 $q$ < 1e-8 *** $R^2$ = 0.022** | $t$(10302) = 0.85 $q$ = 0.39 | $t$(222) = 0.35 $q$ = 0.87 | $t$(191) = 1.35, $q$ = 0.54 |

Listeners with a more similar identity to the speaker had more convergent neural synchrony with each other than listeners with less similar identity in the prefrontal cortex (Figure 4a). This means that if two listeners regarded themselves as very similar to the speaker, their brain activity was more likely to resemble each other in the prefrontal cortex (Figure 5). Thus, perceptions of similar identity seemed to make brain activity converge on a canonical response to the stories. In contrast, listeners with a more similar experience to the speaker had more neural synchrony with each other only in the right occipital cortex. In the bilateral temporal cortex and right posterior middle frontal gyrus, the more two listeners had a similar experience with the speaker, the less neural synchrony they had with each other (Figure 4b). Similar to the behavioral responses, these sorts of listeners seemed to have more idiosyncratic neural activity in these areas while listening to the stories.

There were no significant relationships between similar experience / similar identity and speaker-listener neural synchrony.

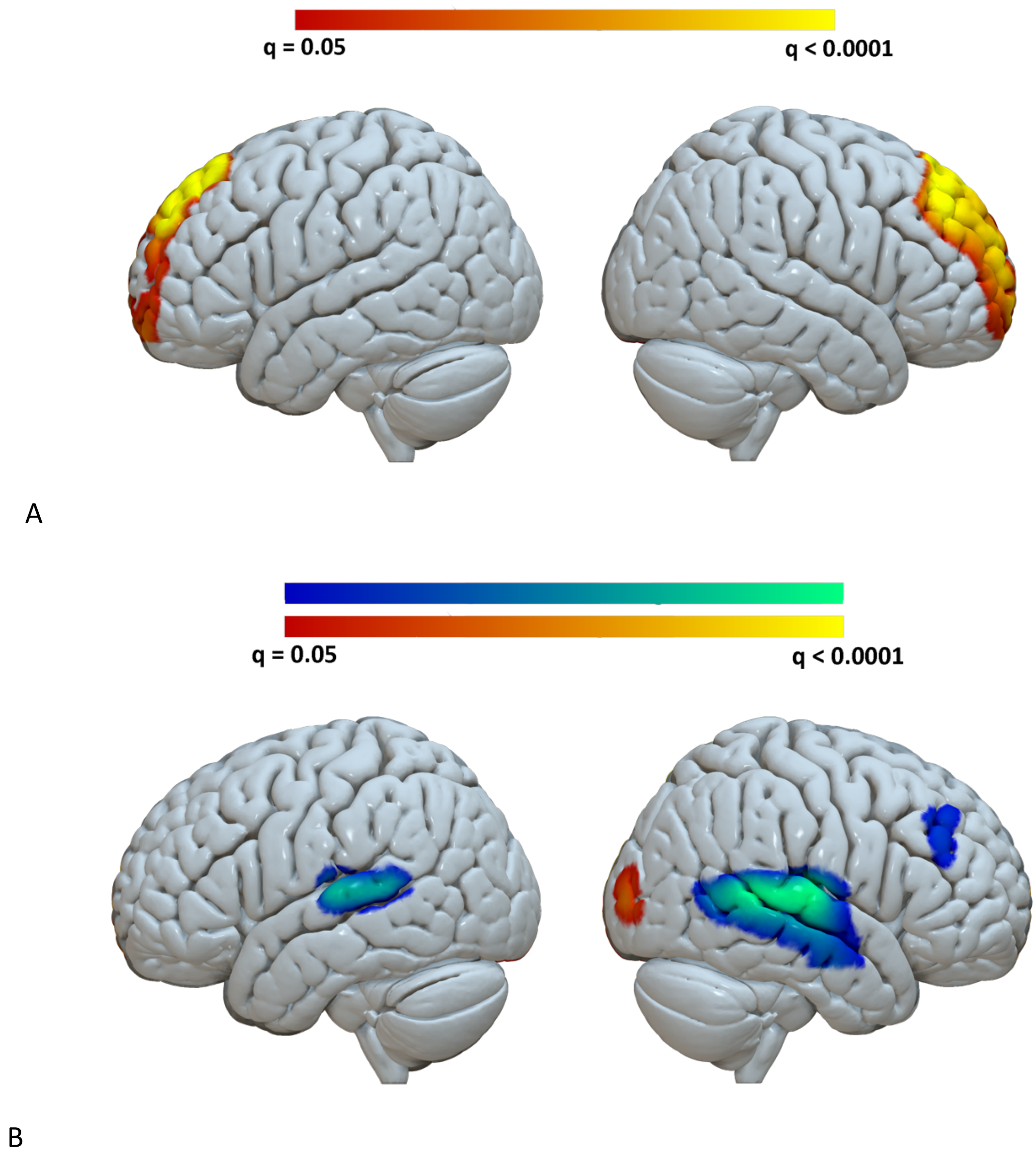



**Figure 4** – A) Areas of greater neural alignment between listeners of high similar identity rating. B) Areas of greater neural alignment (red) and less neural alignment (blue) between listeners of high similar experience rating.

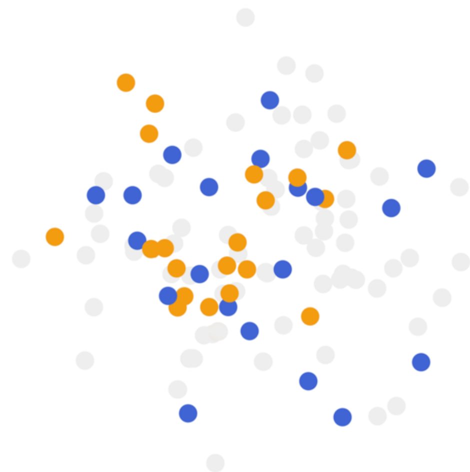


**Figure 5** – Visualization of convergent brain patterns within the prefrontal cortex between listeners who rated themselves as highly similar to the speaker, and listeners who rated themselves as different. The relative location of the dots represents the similarity between brain responses, such that dots that are close together represent more similar neural responses than dots that are far apart. Orange dots are the 20 listeners who rated themselves as most similar to the speaker, while blue dots are the 20 listeners who rated themselves as least similar to the speaker. The orange dots cluster tighter in space, indicating that high-similarity listeners converged more on a canonical brain response to the story.

Lastly, we investigated the effects of similarity on the various empathy measures – emotion state matching with other listeners, emotion state matching with the speaker, empathic concern, empathic accuracy during the retelling, and empathic accuracy for the time of the event. Similar experience was positively related to emotion state matching with the speakers, but negatively related to emotion state matching with each other. That is, the stories seemed to evoke more divergent emotions in listeners who had had the same experience while listeners without the experience had similar emotional responses to each other. However, the set of emotional responses amongst the similar experience listeners still resembled that of the speaker more than the emotional responses of the listeners without a similar experience. Listeners who perceived themselves to have a similar identity to the speaker had more similar emotional experiences to each other, but similar identity was not related to how well their emotional experiences matched the speaker. Both similar experience and similar identity were positively associated with expressed empathy for the speaker independently. Neither similar experience nor similar identity was significantly associated with empathic accuracy.

**Table 2 – Effect of similar experience and similar identity on expressions of empathy**

| | Emotion state matching with other listeners | Emotion state matching with speaker | Empathic concern | Empathic accuracy (telling) | Empathic accuracy (event) |
|---|---|---|---|---|---|
| Similar experience | **$t(12770) = -4.51$ $p = 9.99e-6$*** $R^2 = 0.0017$** | **$t(214) = 2.96$, $q = 0.010$ $R^2 = 0.034$** | **$t(187) = 3.25$, $q = 0.0083$ $R^2 = 0.032$** | t(212) = 0.19, $q = 0.85$ | $t(214) = 0.42$, $q = 0.81$ |
| Similar identity | **$t(12770) = 7.16$ $p = 1.27e-12$*** $R^2 = 0.0039$** | t(217) = -0.61, $q = 0.81$ | **$t(218) = 5.90$, $q = 8.16e-8$*** $R^2 = 0.17$** | $t(218) = 0.17$, $q = 0.87$ | $t(216) = 0.66$, $q = 0.99$ |

## Discussion

Common wisdom suggests that someone who has experienced the same thing as someone else should be able to more effectively understand and empathize. However, quantitative evaluations of whether this similar experience actually improves accurate interpersonal understanding are missing from the literature, and it is unclear how much a positive effect may be driven by similar experience versus more general similar identity between a speaker and listener. This study aimed to resolve these questions by measuring semantic, affective, and neural alignment between speakers and listeners of two negative emotion stories.

Perhaps contrary to expectations, listeners who had had a similar experience as the speaker did not have significantly better or worse understandings of the story contents than listeners who had not had a similar experience. There was no detectable effect of similar experience on either semantic alignment, affect perception alignment, or neural alignment between the speaker and listeners. This suggests that when encoding the details of a story, the personal history of a listener is not related to how well they understand the story. However, amongst each other, similar experience did have an effect. In particular, pairs of listeners who said they both had a similar experience to the speaker had more divergent semantic understandings of the corresponding story. The emotion state the story evoked within them was also more idiosyncratic if they had had a similar experience. Additionally, while listening to the story, listeners with a similar experience had less neural alignment with each other in areas including bilateral superior temporal cortex. Given this area is associated with the perception and decoding of spoken language (Buchsbaum, Hickok, & Humphries, 2001; Chang et al., 2010; Howard et al., 2000), this result seems to suggest that people with a similar experience were encoding the story in a more idiosyncratic way at the lowest levels of auditory perception. Such divergence in narrative understanding may be a result of self-projection while listening to the story – if certain details remind you of events in your past, this could trigger vivid and unique mental associations and memories while listening to the story that would not exist in people without similar past experience. These associations may distract the listener from attending to the specific details of the story, but still evoke similar emotions. This may be why neural and semantic alignment were negatively associated with similar experience, but these listeners still had relatively more similar emotional reactions to the speaker than listeners did without the similar experience.

In contrast, perceptions of similar identity between the listeners and speaker were generally positively associated with how the story was encoded. While this variable was still not related to how well listeners understood the specific semantic and affective details communicated by the speaker, listeners who saw themselves as more similar to the speaker seemed to converge on some canonical understanding of the stories. There was greater semantic alignment in their retellings of the story compared to listeners who did not see themselves as similar to the speaker, and their neural activity was more convergent in the prefrontal cortex. This suggests they were processing the narratives in a particular way. This may be because generally similar people can more accurately predict each others' mental states and actions (Stinson & Ickes, 1992). Alternatively, this may be a story of motivation. Perceptions of similarity between people increases motivation to attend to them (Westmaas & Silver, 2006), and higher motivation to understand someone increases empathic accuracy for their emotions (Klein & Hodges, 2001; Nelson, Klein, & Irvin, 2010). Without detailed personal experiences to refer to, then, people with a similar identity to the speaker may be motivated to listen and then fall back on a more stereotypical mental representation of what they heard. We expect the motivation story is more likely, as similar identity in this study was defined by the listener's perceptions of similarity based on a short video clip (which might be inaccurate) and not any specific measures of similarity dimensions between the speaker and listeners.

A lack of associations between similarity variables and alignment with the speaker could signify that these attributes don’t affect how well one understands the story another person is sharing. If so, this would be good news for anyone hoping to be an effective support giver but who does not share a history of certain events with the speaker. However, this could also be a result of study design – only one speaker was used in this study, so it is possible that meaning was not well communicated in this particular instance. This design was chosen because it allowed for pairwise analysis between listeners of the same narrative, but it does reduce the statistical power of the analyses between the speaker and listeners. In order to verify that similar experience and identity are not associated with accurate interpersonal understanding, future research should include more speakers and story examples. Altogether, these results illustrate how similar experience and similar identity operate differently to influence the way we process socially shared information.

## References


Banks, G., Berenson, B.G., & Carkhuff, R.R. (1967). The effects of counselor race and training upon counseling process with Negro clients in initial interviews. *Journal of Clinical Psychology, 23(1),* 70-72.

Barnett, M.A. (1984). Similarity of experience and empathy in preschoolers. *The Journal of Genetic Psychology, 145(2),* 241-250.

Barnett, M.A., Tereault, P.A., & Masbad, I. (1987). Empathy with a rape victim: The role of similarity of experience. *Violence and Victims, 2(4),* 255-262.

Batson, C.D., Sympson, S.C., Hindman, J.L., Decurz, P. Todd, R.M., Weeks, J.L., Jennings, G., & Burris, C.T. (1996). "I've been there, too": Effect on empathy of prior experience with a need. *Personality & Social Psychology Bulletin, 22(5),* 474-482.

Buchsbaum, B.R., Hickok, G., & Humphries, C. (2001). Role of left posterior superior temporal gyrus in phonological processing for speech perception and production. *Cognitive Science, 25,* 663-678.

Chang, E.F., Rieger, J.W., Johnson, K., Berger, M.S., Barbaro, N.M., & Knight, R.T. (2010). Categorical speech representation in human superior temporal gyrus. *Nature Neuroscience, 13(11),* 1428-1433.

Dikker, S., Silbert, L.J., Hasson, U., & Zevin, J.D. (2014). On the same wavelength: Predictable language enhances speaker-listener brain-to-brain synchrony in posterior superior temporal gyrus. *The Journal of Neuroscience, 34(18),* 6267-6272.

Duran, N.D., Paxton, A., & Fusaroli, R. (2019). ALIGN: Analyzing Linguistic Interactions with Generalization techNiques – A Python library. *Psychological Methods, 24(4),* 419-438.

Eklund, J. Andersson-Str˚aberg, T., & Hansen, E.M. (2009). "I've also experienced loss and fear": Effects of prior similar experience on empathy. *Scandinavian Journal of Psychology, 50,* 65-69.

Girard, J.M. (2018). CARMA: Software for continuous affect rating and media annotation. *Journal of Open Research Software, 2(1),* e5.

Glerean, E., Salmi, J., Lahnakoski, J.M., Jääskeläinen, I.P., & Sams, M. (2012). Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity. *Brain Connectivity, 2(2),* 91-101.

Grover, S.L. & Brockner, J. (1989). Empathy and the relationship between attitudinal similarity and attraction. *Journal of Research in Personality, 23(4),* 469-479.

Haley, T.J. & Dowd, E.T. (1988). Responses of deaf adolescents to differences in counselor method of communication and disability status. *Journal of Counseling Psychology, 35(3),* 258-262.

Heinke, M.S. & Louis, W.R. (2009). Cultural background and individualistic-collectivistic values in relation to similarity, perspective taking, and empathy. *Journal of Applied Social Psychology, 39(11),* 2570-2590.

Hodges, S.D., Kiel, K.J., Kramer, A.D.I., Veach, D., & Villanueva, B.R. (2010). Giving birth to empathy: The effects of similar experience on empathic accuracy, empathic concern, and perceived empathy. *Personality & Social Psychology Bulletin, 36(3),* 398-409.

Honey, C.J., Thompson, C.R., Lerner, Y., & Hasson, U. (2012). Not lost in translation: Neural responses shared across languages. *The Journal of Neuroscience, 32(44),* 15277-15283.

Howard, M.A., Volkov, I.O., Garell, P.C., Noh, M.D., Granner, M., Damasio, H., Steinschneider, M., Reale, R.A. Hind, J.E., & Brugge, J.F. (2000). Auditory cortex on the human posterior superior temporal gyrus. *Journal of Comparative Neurology, 416(1),* 79-92.

Hoyt, T., Pasupathi, M., Smith, B.W., Yeater, E.A., Kay, V.S., & Tooley, E. (2010). Disclosure of emotion events in groups at risk for posttraumatic stress disorder. *International Journal of Stress Management, 17(1),* 78-95.

Kirk, W.G., Best, J.B. & Irwin, P. (1986). The perception of empathy in alcoholism counselors. *Journal of Studies on Alcohol, 47(1),* 82-84.

Klein, K.J.K. & Hodges, S.D. (2001). Gender differences, motivation, and empathic accuracy: When it pays to understand. *Personality & Social Psychology Bulletin, 27(6),* 720-730.

Krebs, D. (1975). Empathy and altruism. *Journal of Personality & Social Psychology, 32,* 1134-1146.

Liu, Y., Piazza, E.A., Simony, E., Shewokis, P.A., Onaral, B., Hasson, U., & Ayaz, H. (2017). Measuring speaker-listener neural coupling with functional near infrared spectroscopy. *Scientific Reports, 7,* 43293.

Mackinnon, A., Jorm, A.F., Christensen, H., Korten, A.E., Jacomb, P.A., & Rodgers, B. (1999). A short form of the Positive and Negative Affect Schedule: Evaluation of factorial validity and invariance across demographic variables in a community sample. *Personality and Individual Differences, 27(3),* 405-416.

Nastase, S.A., Gazzola, V., Hasson, U., & Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. *Social Cognitive and Affective Neuroscience, 14(6),* 667-685.

Nelson, D.W., Klein, C.T.F., & Irvin, J.E. (2010). Motivational antecedents of empathy: Inhibiting effects of fatigue. *Basic and Applied Social Psychology, 25(1),* 37-50.

Nguyen, M. Vanderwal, T., & Hasson, U. (2018). Shared understanding of narratives is correlated with shared neural responses. *NeuroImage, 184,* 161-170.

Pietromonaco, P.R., Rook, K.S., & Lewis, M.A. (1992). Accuracy in perceptions of interpersonal interactions: Effects of dysphoria, friendship, and similarity. *Journal of Personality & Social Psychology, 63(2),* 247-259.

Preis, M.A. & Kroener-Herwig, B. (2012). Empathy for pain: The effects of prior experience and sex. *Therapeutic Studies & Psychology, 16(9),* 1311-1319.

Preston, S.D. & Hofelich, A.J. (2012). The many faces of empathy: Parsing empathic phenomena through a proximate, dynamic-systems view of representing the other in the self. *Emotion Review, 4(1),* 24-33.

Robiner, W.N. & Storandt, M. (1983). Client perceptions of the therapeutic relationship as a function of client and counselor age. *Journal of Counseling Psychology, 30(1),* 96-99.

Simich, L., Beiser, M., & Mawani, F.N. (2003). Social support and significance of shared experience in refugee migration and resettlement. *Western Journal of Nursing Research, 25(7)*, 872-891.

Stephens, G.J., Silbert, L.J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences, 107(32),* 14425-14430.

Stinson, L. & Ickes, W. (1992). Empathic accuracy in the interactions of male friends versus male strangers. *Journal of Personality and Social Psychology, 62,* 787-797.

Suitor, J.J., Pillemer, K., & Keeton, S. (1995). When experience counts: the effects of experiential and structural similarity on patterns of support and interpersonal stress. *Social Forces, 73(4),* 1573-1588.

Verhofstadt, L.J., Buysse, A., Ickes, W., Davis, M., & Devoldre, I. (2008). Support provision in marriage: The role of emotional similarity and empathic accuracy. *Emotion, 8(6),* 792-802.

Westmaas, J.L. & Silver, R.C. (2006). The role of perceived similarity in supportive responses to victims of negative life events. *Personality & Social Psychology Bulletin, 32(11),* 1537-1546.

Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C.J., & Hasson, U. (2017). Same story, different story: The neural representation of interpretive frameworks. *Psychological Science, 28(3),* 307-319.

# Chapter 4 - Using linear, nonlinear, and higher-order approaches to identifying neural synchrony in conversing dyads

## Introduction

Emerging theory of social interaction posits that in order to coordinate with other people effectively, separate minds mutually adapt to each other in a dynamic fashion to converge on overlapping mental representations and synchronous behavior (Chatel-Goldman et al., 2013; De Jaegher, Di Paolo, & Gallagher, 2010; Hasson et al., 2012; Kelso, Dumas, & Tognoli, 2013; Wheatley et al., 2012). By adapting to the dynamic patterns of someone else (their behavior, their language use, etc.), one can better anticipate the behavior of other people, form joint goals, and operate as cohesive social units.

An important piece of evidence for this theory is neural data demonstrating alignment in mental representations between interacting people. In fMRI experiments, researchers have found that spatial patterns representing particular mental concepts are reliable and distinguishable (Devereux et al., 2013; Diedrichsen & Kriegskorte, 2017; Thornton & Mitchell, 2017), and that listeners of a narrative story have meaningfully convergent spatial patterns of brain activity with each other in the default mode network when listening to the same narrative (Chen et al., 2017; Regev et al., 2019). These patterns match those elicited when speaking about the narrative as well (Chen et al., 2017; Zadbood et al., 2017). Over time, fluctuations in these patterns map onto each other across people the more people share a basis of knowledge for processing the narrative – for example, when knowing the language the narrative is in (Honey et al., 2012; Liu et al., 2017) or having particular background knowledge or perspective on the narrative (Lahnakoski et al., 2014; Nguyen, Vanderwal, & Hasson, 2018; Yeserun et al, 2017). Outside of the well-controlled lab environment, researchers have also used EEG to find convergent brain processes among high school students the more engaged they were with a lesson (Dikker et al., 2017). This work has shown that synchronized neural activity underlies similar mental representations among people receiving a particular message, and successful communication brings listeners mentally in line with each other.

What about in continuous dialog, where interlocutors may be expected to mutually converge with each other based on ongoing feedback? Researchers have sought to identify synchronous neural patterns in these sorts of situations as well. However, this effort has experienced more difficulty. While several experiments have documented “neural synchrony in social interaction” as a general overarching result, the specifics of the data are more conflicting than in single-person narrative studies. In particular, this observed synchrony seems to occur in only small areas that vary widely across experiments. These studies sometimes identify it in the temporoparietal junction (Jiang et al., 2012; Jiang et al., 2015; Tang et al., 2015), sometimes in medial prefrontal cortex (Holper, Scholkmann, & Wolk, 2012), sometimes in the pre-supplementary motor area (Pan et al., 2017), and sometimes not between any two corresponding areas but instead between different locations across dyad members (Lu & Hao, 2019).

It is difficult to say at this time why these results are not more coherent. One reason may be that social interactions are diverse experiences in terms of interactive dynamics and joint goals, and variation in these factors might affect how the neural synchronizing process progresses (McGrath, 1984). Two people may be engaged in a joint discussion where the goal is to come to an agreement, but along the way different ideas must be explored and thus the partners must evoke divergent mental representations before arriving at a joint conception. In another case where mutual performance is the goal, two partners may be expected to converge neurally as they do the same action, or they might be expected to consistently have divergent activity if the overall goal requires separate actions done in parallel. This diversity of experience makes it difficult to investigate the neural dynamics of dyadic social interaction generally.

In addition, there is need for more explicit acknowledgement of different types of neural alignment findings in this work. Most use the phrase "neural synchrony" to refer to evidence of similar neural patterns between people as they interact in real time. However, the specific mathematical approach to calculating this "similarity" varies substantially, and have large implications for the theoretical interpretation of that similarity. In some work, including most of the literature addressing shared viewing experience of narratives, similarity in neural activity is defined as a linear function between each person's data streams (e.g., Liu et al., 2017; Stephens, Silbert, & Hasson, 2010). That is, given a time course vector $\underline{x}$ of one person's neural data, the time-corresponding values $\underline{y}$ of another person's neural data can be expressed as

$$\underline{y} = \beta_1\underline{x}$$

An important aspect of this sort of relationship is that the effect of x is modulated by a constant $\beta_1$ – any point $x_t$ in $\underline{x}$ can be transformed into the corresponding point $y_t$ in $\underline{y}$ by applying the same scalar transformation $\beta_1$. This kind of relationship means there is a one-to-one mapping between points in $\underline{x}$ and $\underline{y}$, and this mapping is consistent across time. In practice, this means that under linear synchrony, two brain signals would be exhibiting the same amount of activity at the same time (accounting only for a possible difference in scaling; Figure 1).

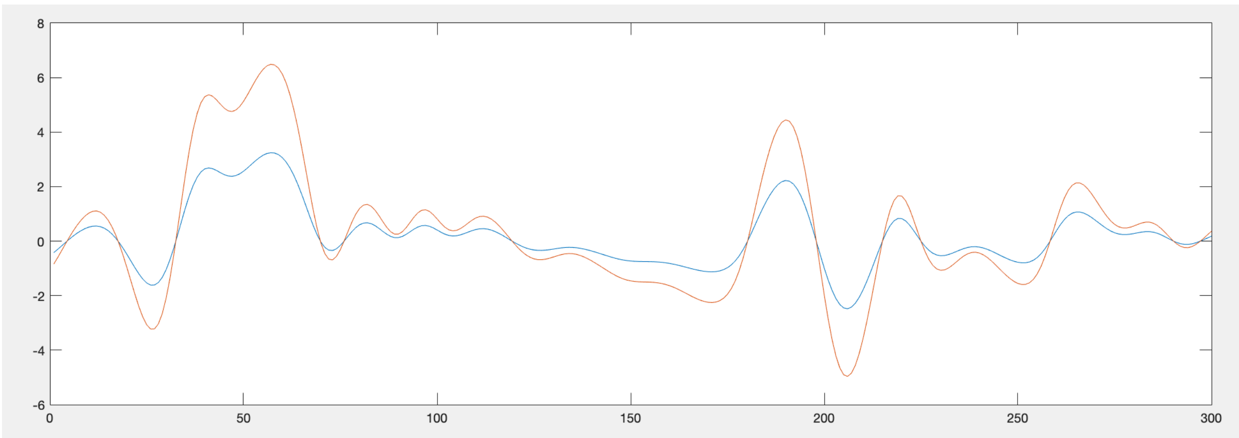



**Figure 1** – Two hypothetical time courses from separate brains that are linearly dependent – similar fluctuations at similar times.

However, two signals could still be dependent even if the requirements for one-to-one, time-invariant mapping of points are relaxed. This would qualify the signals for nonlinear dependence. Some experiments search for this kind of nonlinear relationships between separate people's brain activity (e.g. Holper, Scholkmann, & Wolf, 2012; Jiang et al., 2015). These include a multitude of specific approaches such as bicoherence, wavelet coherence, distance correlation, and mutual information. The exact ways of calculating these approaches vary, but in general they don't search for a single constant scalar that transforms a time course $\underline{x}$ into $\underline{y}$. Instead, there is typically some sort of nonlinear transformation done on the data first in order to find a different basis space to identify dependence. For example, in the cross-wavelet transform, similarity between time courses is identified in the frequency domain of the signals (having converted them to frequency power spectra). This is expressed as

$$W_{f1}\, \underline{y} = \Psi_{xy}\, W_{f2}\, \underline{x}$$

where $W_f$ is itself a function that maps the values of a time course to a particular shape, in this case the Morlet mother wavelet, and $\Psi_{xy}$ is the cross-wavelet power that describes how much the transformed signal $\underline{x}$ is present in the power spectra of the transformed signal $\underline{y}$ (Issartel et al., 2015; Maraun & Kurths, 2004). The implication here is that there is no single transformation that can be applied to all values of $\underline{x}$ to get the right corresponding value of $\underline{y}$ – where $\underline{x}$ increases by a little, $\underline{y}$ might increase by a little or decrease by a lot. But the fluctuation shapes that $\underline{x}$ goes through are also identifiable in $\underline{y}$, but perhaps at a time delay. This might occur in systems where one signal partially causes another at a time delay, or where the two signals are mutually causal.

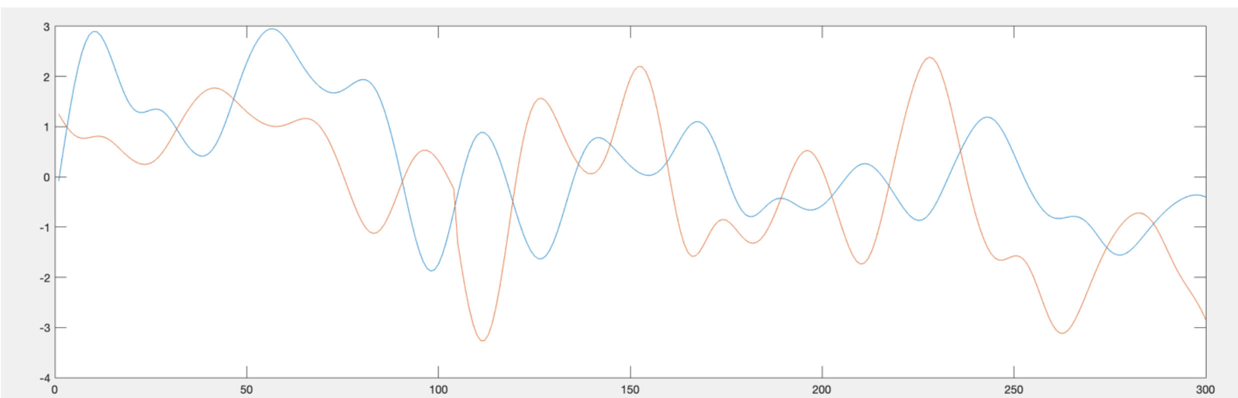



**Figure 2** – Two hypothetical brain time courses with time-variant nonlinear dependence. The blue signal exhibits the same fluctuations, but at a delay of 15 seconds and at varying scales over time. Linear dependence measures such as Pearson's correlation would miss this dependence, but nonlinear synchrony approaches can detect it.

Two signals can also be dependent in time, but not deterministically so. This means that a particular value of x does not predict a closed set of possibilities in y. Put another way, knowing where a certain brain goes next in brain state space may not give information about where another mutually interacting brain goes in state space. However, these signals may still be related through their higher-order statistics – perhaps the specific values of the signals don't correspond, but how much that value changes over time in each signal does. For example, an event happening in the environment two people are a part of may cause their brain activity to shift at the same time in response, but may not cause them to shift in the same way. This means there is not a linear relationship between $x_t$ and $y_t$, but there is in the distance metrics computed between each points t and t+1. This can be expressed as

$$||y_t - y_{t+1}|| = \beta_1 ||x_t - x_{t+1}||$$

where $||\bullet||$ denotes a distance measurement, such as Euclidean distance or correlation. In this way there is no defined solution set for mapping signal x to y, but there is a defined set for mapping the changes in values over time (Figure 3). While this approach has been used to identify the way individual brains reliably chunk viewed narratives into discrete encodings (Baldassano et al., 2017; Chang et al., 2018), it has not yet been used to document coherence between people in live social interactions.

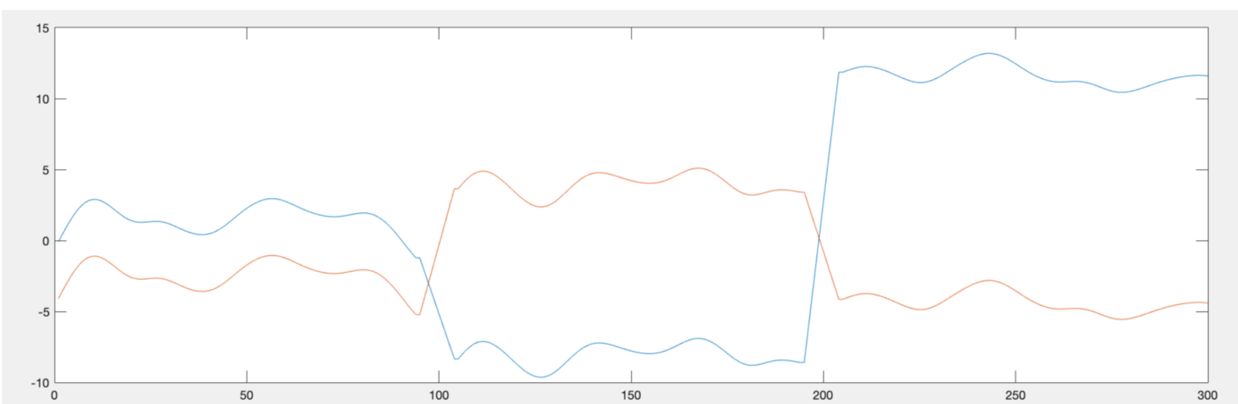



**Figure 3** – Two hypothetical brain time courses that have higher-order synchrony – the value of one signal is not predictable from the value of the other, but a small *change* in one brain predicts a small change in the other, while a large change in one brain predicts a large change in the other.

The psychological implications of these similarity measures are very different. In a linear relationship, this would mean two interlocutors are having the same mental experience across time. In a nonlinear relationship, they may not be having the same experience at once, but similar experiences that one person had are also had by the other at some delay. Lastly, in a higher-order relationship based on changes rather than values, two people may never have the same experience but there is some constraining factor about the conversation that determines the timing of when peoples' experiences change. Currently, because there have not yet been any intentional investigations of what kind of dynamics better explain neural activity in various types of social interaction, our understanding of this overall phenomenon is lacking.

In the present study, then, we aim to directly compare these different measures of neural synchrony in a dataset of conversing dyads in order to investigate what sort of similarity best explains the neural dynamics of that interaction. Specifically, we will be comparing the linear measurement Pearson's correlation, the nonlinear approach wavelet transform, and the state change coherence measurement described previously. The data in this study consists of dyads in a joint decision-making paradigm, where one might expect the moment-to-moment synchrony of the dyads to be low as they present and analyze different opinions on a problem, but for more rigid synchrony to emerge by the end as a joint solution and perspective on the problem is converged upon. Therefore, we expect nonlinear and higher-order similarity measures to better describe the neural dynamics within these conversations than the linear method, and to be more associated with psychologically meaningful attributes about the conversation such as its length, pre-existing similarity in discussion starting position, and subjective ratings of personal similarity / interaction quality / common understanding /

perceived overlap in background information. These results should help clarify the specific nature of neural synchrony within a joint decision context, which would contribute to a better understanding of social interaction overall.

## Methods

### *Participants*

Participants for this study were recruited from the student population of a large California university between fall 2017 and spring 2018. Two hundred and twenty total people participated, paired into 110 same-gender dyads (71 female-female dyads, 39 male-male dyads). The average age of the participant sample was 20.32 years (*SD* = 2.60). Sixty-three identified as White/Caucasian (28.64%), 49 as Biracial/Mixed (22.27%), 48 as Hispanic/Latinx (21.82%), 42 as Asian/Pacific Islander (19.09%), 10 as Middle Eastern/North African (4.55%), 7 as Black/African-American (3.18%), and 1 as Native American (0.45%).

### *Materials*

Neural activity in participants was recorded using functional near infrared spectroscopy (fNIRS). This imaging method uses infrared light to measure oxygenated blood flow in the brain's cortical surface as an indirect measure of brain activity (see Chapter 2 of this dissertation for more information). The specific equipment used was a NIRScout imaging unit from NIRx Technologies (nirx.net). This unit has 32 source and 32 detector optodes, which were split evenly over the two interlocutors' heads during the experiment. These optodes were secured in stretchy head caps and positioned over the prefrontal cortex and bilateral temporoparietal junction to create 35 separate data measurement channels per person (Figure 4). These positions were chosen because they correspond to default mode network areas that can be reached on the surface cortex, and these areas are most commonly identified in past research as locations of coherence in communication. This positioning was standardized across participants using the 10-10 UI external positioning system. Light intensity data was collected at wavelengths 760 and 850nm, with a sampling rate of 3.91Hz.

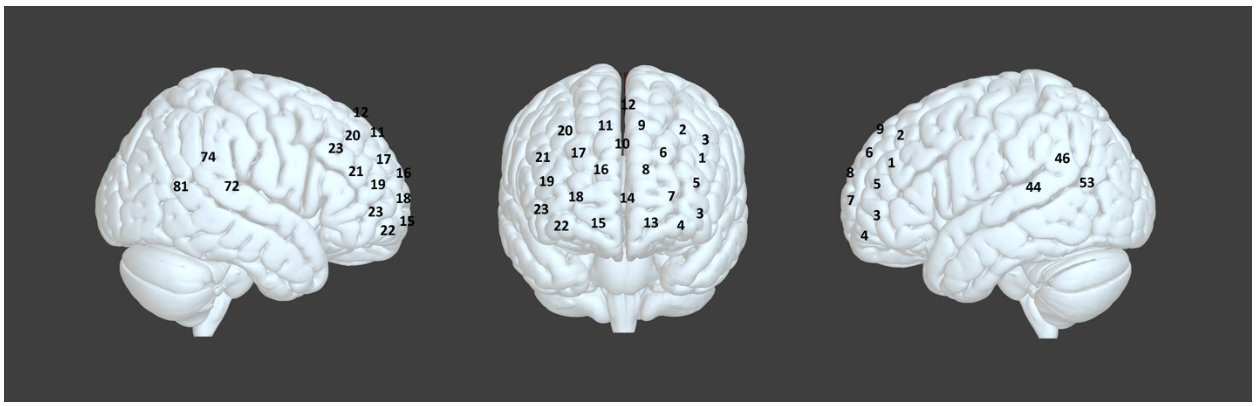


**Figure 4** – Positions of channels on each participant's head.

Prior to the experiment, participants reviewed a short description of a public health issue (the then contemporary Zika epidemic) and descriptions of 5 different approaches to helping people affected by this public health issue. They were then prompted to distribute a hypothetical $100 million of grant money amongst these options (Appendix A). During the experiment when participants were discussing, they again completed this resource allocation task, but togheter as a dyad. After the discussion, participants answered a set of post-study questionnaires to assess their opinions of how the discussion went. The reduced set of questions analyzed for this study are denoted with a * in Appendix B. Audiovisual videos of the participants' discussions were recorded on two Panasonic LUMIX G7 Mirrorless 4k cameras.

*Procedure*

The data in this paper comes from a study that was initially planned to investigate discourse differences between people who believed they were the same or different political identity. Therefore, before coming to the lab, participants completed a political identity questionnaire and then were paired to randomize their respective political positions (while being told they held the same or different positions, a treatment that was orthogonal to dyads' actual political differences). The experimental procedure was modeled after the joint resource allocation task describe in Keltner & Robinson (1993). Upon arrival to the lab, participants were welcomed to different rooms, given informed consent, and then asked to complete the resource allocation task by themselves. Participants were then brought into the same room together, fitted with the fNIRS equipment, and then told they either had the same or different political views, or weren't told anything at all. They were also instructed to complete the resource allocation task again, but this time as a group that needed to arrive at a joint solution. They were asked to discuss the issue in depth, as if they were making a decision about the allocation of real money. After instructions, the experimenters turned on the cameras, started the fNIRS recording, and left the room for the duration of the discussion. This way, the cameras captured both the onset timing of the brain activity recording and the onset of the conversation. Participants were given as much time as they needed to come to a joint solution to the task, and then rang a desk bell at the end to signal to the experimenters that they were finished. Lastly, participants were again separated into different rooms and asked to complete the discussion quality questionnaire. After completion of this survey, participants were debriefed about their actual political alignment and the aims of the study, were given compensation, and allowed to leave.

*Video Coding*

Video recordings of the conversation were trimmed to start and end with the dyads' conversations, and then were analyzed in three ways. First, the videos were run through the opensource Python software FlowAnalyzer (https://flowanalyzer.readthedocs.io/) in order to extract measurements of the participants' frame-by-frame motion (Barbosa et al., 2009). These were downsampled to the sampling rate of the fNIRS acquisition, and then used as motion regressors for the neural data. Likewise, research assistants watched the videos and coded when each participant was speaking. These data were used as speaking turn regressors for the neural data. Lastly, transcripts of the conversations were automatically generated with the online tool Temi (temi.com), and any mistakes in the auto transcripts were corrected by research assistants. These transcripts will be analyzed at a later date.

*Neural Data Preprocessing*

Data collected with fNIRS were subjected to a preprocessing pipeline that progressed as follows – 1) automatic identification and removal of noisy channels with frequency spectra approach (see Chapter 2); 2) Removal of motion and speaking regressors; 3) detection of motion artifacts within remaining channels, defined as periods of data with a greater than 5 standard deviation change in less than 1 second; 4) rescaling of data in artifact segments to remove spikes/discontinuities/volatility changes that do not occur in neural patterns and which might negatively impact further preprocessing/analysis; 5) bandpass filtering to 0.008-0.2Hz; 6) conversion of light intensity values to percent change in oxygenated hemoglobin concentration using the Modified Beer Lambert Law; 7) evaluation and rejection of any channels with large remaining spikes/volatility changes. Regression of motion and speaking regressors was performed prior to later preprocessing and analysis because these actions might introduce patterns within the data due only to the offset and onset of physically speaking, rather than underlying neural representations.

## Results

### *Linear Synchrony Between Dyads*

First, we analyzed dyads' neural data using Person's correlation to check if we could detect any linear dependencies between interlocutor's brain data (Nastase et al., 2019). Specifically, we took neural activity time courses from the same data channel in each of the discussion partners' recordings and calculated the correlation between these. We then tested these correlation values against a bootstrapped null distribution, generated from 10,000 random pairings of participants' data who were not in the same conversation together. Controlling for multiple comparisons, we found that dyads exhibited significant linear synchrony in the medial prefrontal cortex during their discussion (Figure 5). However, there was no significant increase in this synchrony over the course of the discussion, in this area or other brain areas, as measured by comparing synchrony estimates in the first and last two minutes of the conversation. There also were no significant associations between correlation synchrony and any behavioral measures.

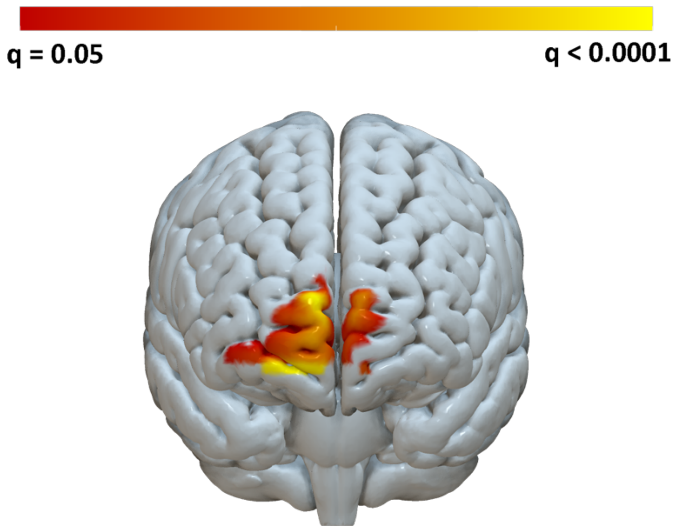
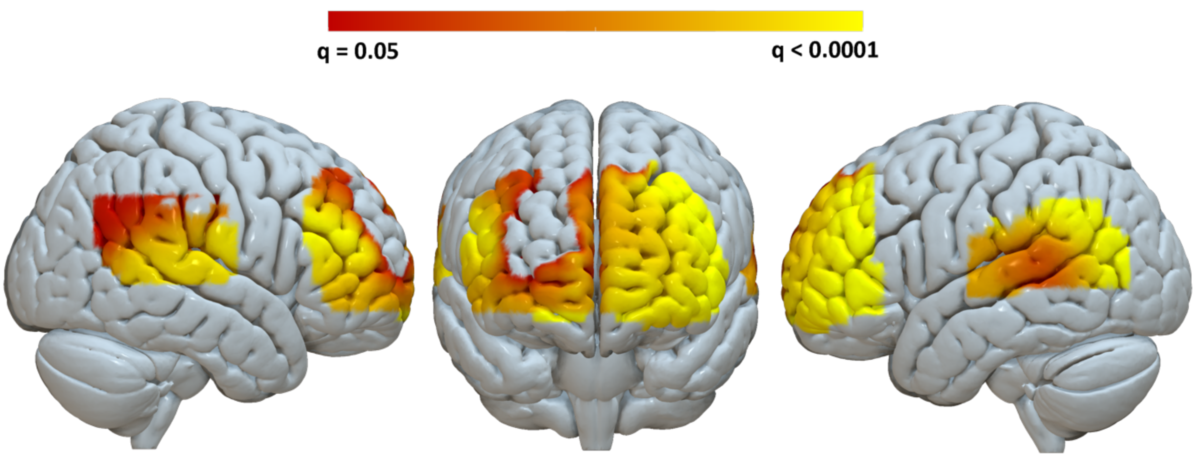


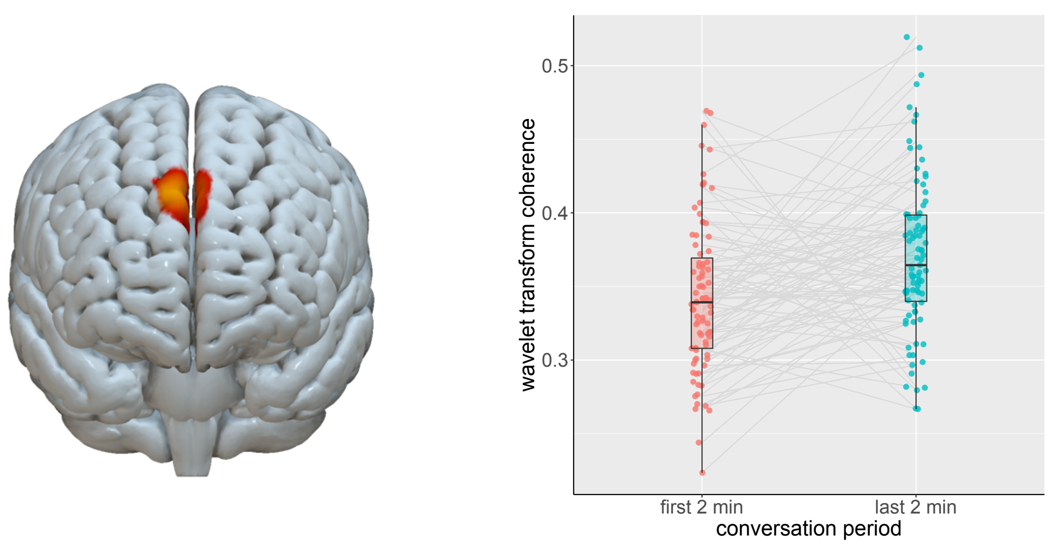

**Figure 5** – Location of significant neural time series correlation during dyad conversations.

*Nonlinear Synchrony Between Dyads*

Next, we analyzed dyads' neural data using the cross-wavelet transform procedure to check if we could detect any linear dependencies between interlocutor's brain data (code from Grinsted, Moore, & Jevrejeva, 2004). Moving window sections of neural activity time courses from the same data channel in each of the discussion partners' recordings were convolved with Morlet mother wavelets of varying scale, and the cross wavelet power was then estimated for the frequency bin corresponding to each scale. This produces a matrix of time and frequency-resolved estimates of wavelet coherence between the two signals, which was then averaged across the time period of interest for each analysis. We then tested these coherence values against a bootstrapped null distribution, generated from 10,000 random pairings of participants' data who were not in the same conversation together. Controlling for multiple comparisons, we found that dyads exhibited much more extensive nonlinear coherence during the discussion than in the linear coherence analysis. Significant nonlinear synchrony was present in nearly the entire prefrontal cortex, as well as in bilateral temporoparietal junction (Figure 6a). There was also a significant increase between the first and last two minutes of the conversation within the dorsomedial prefrontal cortex ($t(206,1) = 3.70$, $p = 0.00039$; Figure 6b), though there were still no associations with the behavioral measures.




A



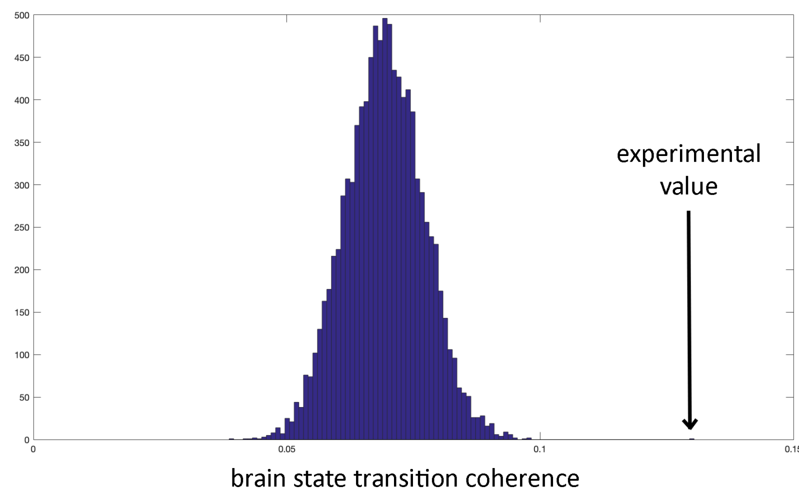

B

**Figure 6** – A) Areas of significant wavelet transform coherence between dyad partners during conversation. B) There was a significant increase in wavelet transform coherence between dyad partners from the first two minutes of the conversation to the last two minutes.

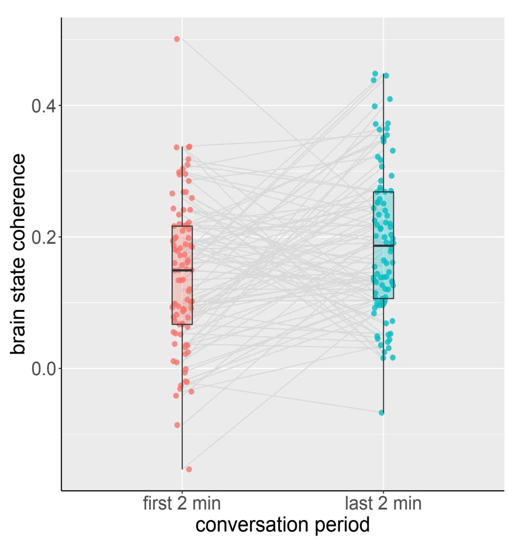
*State Change Synchrony Between Dyads*

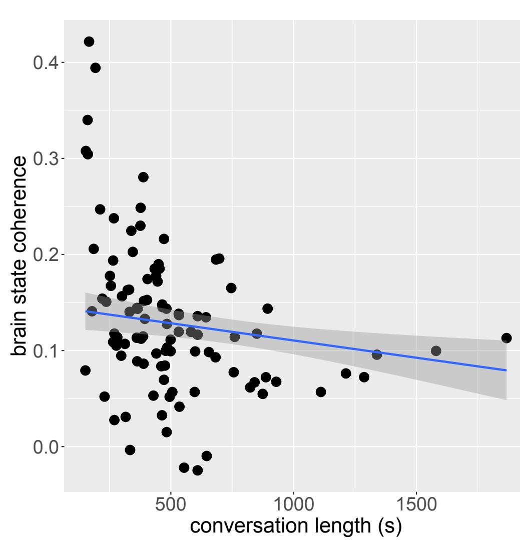
Finally, we analyzed how well the timing of dyads' whole-brain state changes matched each other ("whole-brain" here meaning the inclusion of all data measurement channels over the default mode network). Past publications that used this method relied on Hidden Markov Models to match neural data to predefined state change events (corresponding to scene changes in movies) in order to estimate alignment of the location of these events between people. In our case, it was not possible to predefine the timing or even number of expected brain state changes, as a natural conversation is not as experientially discrete as an audiovisual movie. Thus, we developed an alternative measure that tracked relative magnitude of brain state changes over time, and then identified the correlation between these magnitudes for the separate discussion partners. First, within one subject, the vector of whole-brain activity at every time $t$ was correlated with the vector of whole-brain activity at every other point in time to create an auto-similarity matrix. The off-diagonal $k$ for this matrix thus represents the magnitude of brain state change that occurred between time $t$ and $t+k$. These off-diagonals were extracted for the signal windows $t$+5 through $t$+20, and corresponding off-diagonals from each interaction partner were then correlated. The resulting correlation values for every off-diagonal examined were then averaged to derive one state change coherence value. We then tested these coherence values against a bootstrapped null distribution, generated from 10,000 random pairings of participants' data who were not in the same conversation together.

Using this procedure, we found a significant amount of neural state change coherence in conversing dyads ($t(103) = 7.97$, $p < 0.0001$; Figure 7a). Between the first and last two minutes of the conversation, there was also a significant increase in neural coherence ($t(206,1) = 3.17$, $p = 0.0021$; Figure 7b). Lastly, the amount of neural coherence over the course of the conversation was significantly associated with how long the conversation lasted, such that dyads with greater state change coherence arrived at a joint solution to the task more quickly exhibited higher neural coherence ($t(102) = -2.85$, $p = 0.0053$; Figure 7c).




A




B

C

**Figure 7** – A) Bootstrapped null distribution of brain state transition coherences among non-interacting participants, compared to the experimental sample mean. B) This brain state transition coherence measure significantly increased over the first and last two minutes of the participants' conversation. C) The extent of this brain state coherence over the conversation was significantly associated with conversation length, such that quicker discussions exhibited great brain state coherence.

## Discussion

The theory of shared cognition in social interactions posits that the coordination of neural and mental states in interactional communication enables mutual coordination and understanding. However, until now it has been difficult to establish this phenomenon in real life dyads due to the complexities of back-and-forth conversation. This study aimed to investigate the nature of alignment that might exist within a mutual decision-making conversation by comparing different conceptualizations of synchrony.

We found evidence of linear synchrony between discussion partners within the medial prefrontal cortex during their discussion, such that during the conversation the activity patterns within this area are more entrained than they would be between non-interacting dyads. Specifically, this entrainment was of the type where certain temporal patterns of fluctuations within one person matched on to the same patterns in the other person. This suggests that somewhat similar mental processes within this area were engaged in at the same time between discussion partners. The medial prefrontal cortex is consistently associated with reasoning about decision options and the mental states of others (Frith & Frith, 2006; Overwalle & Baetens, 2009; Spiers & Maguire, 2006), so given that linear synchrony was identified here, this may suggest that discussion partners tend to engage reasoning in concert with one another in conversation.

We also found more extensive time-variant nonlinear synchrony across wide areas of the medial and lateral prefrontal cortex, as well as bilaterally in the temporoparietal junction. According to these results, activity in these areas did not necessarily match in amount between discussion partners, but particular kinds of activity in one person was reliably associated with some other kind of activity in the other. While fNIRS is likely not spatially resolved well enough to detect differences in spatial patterns within a brain area that would differentiate the content of information represented here, this sort of widespread synchrony between partners might speak to a see-sawing, communication-generation vs. communication-decoding pattern between alternating speakers and listeners. This is a weaker form of synchrony than the above reported linear coherence results, but it still suggests that some amount of mutual adaption occurred within real discussion partners. This entrainment increased in strength between the beginning and end of the conversation in the dorsomedial prefrontal cortex in particular, perhaps indicating that trading roles between sharing one's thoughts and interpreting another's thoughts became a more efficient process over the course of the conversation.

Finally, across the default mode network, we found evidence that movement between different brain states in one person significantly predicted movement to new brain states within an interacting partner. This higher-order synchrony increased from beginning to end of the interaction, and the strength of this synchrony over the interaction was associated with how efficiently the interaction progressed. The analysis method used did not assume that these states would match across participants, or even that one particular state within a participant would ever reoccur at any point in another person. Thus, these results speak not so much to any shared mental representations between people, but does suggest that they were tuned into a particular temporal progression through the conversation – when one person switched to a new topic mentally, the other person was not still stuck in an old representation or daydreaming independently of the other person's communication. They progressed through the conversation in a coordinated fashion, and were detectably within the same social system experiencing the same temporal events. Where this coordination was weaker, interpersonal decision-making took longer to complete, perhaps signifying some process loss between people who weren't progressing toward a solution at the same rate.

All together, these results suggest that while there may be some minimal linear synchrony and temporal pattern matching between people in a live decision-making dialog, nonlinear and higher-order cohesion describe the neural dynamics of this sort of interaction more fruitfully. This means that discussion partners do not simply mirror each other's neural and mental states during conversation, but are still coordinated in terms of what they think next and when they think it. Yet, only one of these measures was significantly associated with only one behavioral measure of interaction quality, suggesting the identified results may be more a property of joint decision-making generally and can't differentiate between subjective assessments of how the interaction progressed. In the future, it would be worthwhile to validate these results in a more spatially resolved neuroimaging method like hyperscanning fMRI, or in conversations where agreement is never reached. It would also be worthwhile to compare these types of synchronization in other interaction contexts where psychological convergence might be expected for the entire conversation, and which might more closely predict interaction quality – e.g., joint action, or interpersonal instruction. In these cases, perhaps linear synchronization and mental representation matching is more prominent.

In addition, while the above analysis methods were applied largely in a global fashion across large sections of time in the interaction, sliding window or time localized versions can also be used to pinpoint particular moments of high synchrony within an interacting dyad. Social units are hypothesized to be weak oscillators (Mayo & Gordon, 2020), which typically exhibit instability in terms of the amount of cohesion experienced at any one moment. It would be particularly interesting to investigate if, say, periods of linear synchrony might emerge out of general nonlinear patterns at key moments of clarity between people, or if length/strength of these temporary periods increase in more successful interlocutors. By continuing to acknowledge and explore the various kinds of synchrony that may occur within social interactions, we can grow our understanding of how people’s minds coordinates across a diversity of social interaction contexts.

Appendices

*Appendix A: Discussion Prompt:*

**The CDC has issued an alert about the dangers of the Zika virus. This disease has affected hundreds of thousands of people worldwide and tens of thousands in the United States and Territories. It is spread mostly via mosquito bites. While Zika is usually not fatal, it can induce fever, rash, joint and muscle pain, and headache. Most notably, if a pregnant woman becomes infected, her fetus may develop birth defects like microcephaly, which causes the baby to have a smaller-than-normal head size, intellectual disability, hearing and vision problems, and/or seizures. There is currently no vaccine for Zika. These facts make Zika one of the more concerning health risks currently known to the CDC.**

Imagine you are a member of a grant committee. You are in charge of $100 million in charity money for the purpose of helping people affected by the Zika virus. Below are a number of programs that have proposed different solutions. Please decide how you would allocate the money among the following programs in order to best help people affected by Zika:

1. Scientific research and development to find a vaccine ____________________
2. Public education about how best to avoid contracting the disease ____________________
3. Research and development for improving the life quality of babies already born with microcephaly ____________________
4. Subsidize healthcare for affected families ____________________
5. Research and implement mosquito control/eradication strategies ____________________

*Appendix B: Post-Discussion Questionnaire*

**My partner was very cooperative**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**My partner's input was very useful**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**I liked my partner as a person**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**My partner seemed very willing to work together to decide on the solution**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**My partner had good ideas**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**The reasons my partner gave for their opinions made sense to me**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**I think my partner would agree with me on other important issues**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**I would want to work with my partner again in the future**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

***I found the interaction to be comfortable**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

***I thought the interaction was very difficult to get through**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

***This interaction made me feel very stressed**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**I was motivated to find the best solution**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Very False | Somewhat False | Unsure | Somewhat True | Very True |

**My partner was motivated to find the best solution**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Very False | Somewhat False | Unsure | Somewhat True | Very True |

**I listened carefully to my partner**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Very False | Somewhat False | Unsure | Somewhat True | Very True |

**My partner listened carefully to me**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Very False | Somewhat False | Unsure | Somewhat True | Very True |

**The inputs from my partner and I complimented each other**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Very False | Somewhat False | Unsure | Somewhat True | Very True |

**We drew conclusions together**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Very False | Somewhat False | Unsure | Somewhat True | Very True |

**My partner and I handled differences of opinions by addressing them directly**

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Very False | Somewhat False | Unsure | Somewhat True | Very True |

**My partner gave me equal say in the decision**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**I tried to give my partner equal say in the decision**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

***My partner and I had a common understanding of what the problem was**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

***My partner and I had a common understanding of how to solve the problem**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

***My partner and I had very different background information about the topic at the start of our interaction**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

***My partner and I are very different types of people**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**We solved the problem in a way we both agree on**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**I am satisfied with the decision we came up with**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**My partner is satisfied with the decision we came up with**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**I think our solution is the best solution to the problem at hand**

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**My partner thinks our solution is the best solution to the problem at hand**

| | | | | | | |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Strongly Disagree | Moderately Disagree | Slightly Disagree | Neither Agree nor Disagree | Slightly Agree | Moderately Agree | Strongly Agree |

**How much do you care about solving health issues related to the Zika virus?**

| | | | | |
|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 |
| None At all | A little | A moderate amount | A lot | A great deal |

## References


Baldassano, C., Chen, J., Zadbood, A., Pillow, J.W., Hasson, U., & Norman, K.A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron, 95,* 709-721.

Barbosa, A., Munhall, K.G., Yehia, H.C., & Vatikiotis-Bateson, E. (2009). Quantifying the time-varying coordination between performer and audience. *The Journal of the Acoustical Society of America, 125,* 2574.

Chang, L.J., Jolly, E., Cheong, J.H., Rapuano, K., Greenstein, N., Chen, P.H.A., & Manning, J.R., (2018, preprint). Endogenous variation in ventromedial prefrontal cortex state dynamics during naturalistic viewing reflects affective experience. *bioRxiv*, doi:

Chen, J., Leong, Y.C., Honey, C.J., Yong, C.H., Norman, K.A., & Hasson U. (2017). Shared memories reveal shared structure in neural activity across individuals. *Nature Neuroscience, 20(1),* 115-125.

Chatel-Goldman, J., Schwartz, J.-L., Jutten, C., & Congedo, M. (2013). Non-local mind from the perspective of social cognition. *Frontiers in Human Neuroscience*, *7*(April), 107.

De Jaegher, H., Di Paolo, E., & Gallagher, S. (2010). Can social interaction constitute social cognition? *Trends in Cognitive Sciences*, *14*(10), 441–447.

Devereux, B.J., Clarke, A., Marouchos, A., & Tyler, L.K. (2013). Representational similarity analysis reveals commonalities and differences in the semantic processing of words and objects. *The Journal of Neuroscience, 33(48),* 18906-18916.

Diedrichsen, J. & Kriegeskorte, N. (2017). Representational models: A common framework for understanding encoding, pattern-component, and representational-similarity analysis.

*PLoS Computational Biology, 13(4),* e1005508.

Dikker, S., Wan, L., Kaggan, L., Oostrik, M., McClintock, J., Rowland, J., Michalareas, G., Van Bavel, J.J., Ding, M., & Poeppel, D. (2017). Brain-to-brain synchrony tracks real-world dynamic group interactions in the classroom. *Current Biology, 27(9),* 1375-1380.

Frith, C.D. & Frith, U. (2006). The neural basis of mentalizing. *Neuron, 50(4),* 531-534.

Glerean, E., Salmi, J., Lahnakoski, J.M., Jääskeläinen, I.P., & Sams, M. (2012). Functional magnetic resonance imaging phase synchronization as a measure of dynamic functional connectivity. *Brain Connectivity, 2(2),* 91-101.

Grinsted, A., Moore, J.C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics, 11,* 561-566.

Honey, C.J., Thompson, C.R., Lerner, Y., & Hasson, U. (2012). Not lost in translation: Neural responses shared across languages. *The Journal of Neuroscience, 32(44),* 15277-15283.

Hasson, U., Ghazanfar, A. A., Galantucci, B., Garrod, S., & Keysers, C. (2012). Brain-to-brain coupling: A mechanism for creating and sharing a social world. *Trends in Cognitive Sciences*, *16*(2), 114–121.

Holper, L., Scholkmann, F., & Wolf, M. (2012). Between-brain connectivity during imitation measured by fNIRS. *NeuroImage, 63,* 212-222.

Issartel, J., Bardainne, T., Gaillot, P., & Marin, L. (2015). The relevance of the cross-wavelet transform in the analysis of human interaction – a tutorial. *Frontiers in Psychology, 5,* 1566.

Jiang, J., Chen, C., Dai, B., Shi, G., Ding, G., Liu, L., & Lu, C. (2015). Leader emergence through interpersonal neural synchronization. *Procedings of the National Academy of Sciences, 112(14),* 4274-4279.

Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., & Lu, C. (2012). Neural synchronization during face-to-face communication. *The Journal of Neuroscience, 32(45),* 16064-16069.

Kelso, J.A.S., Dumas, G., & Tognoli, E. (2013). Outline of a general theory of behavior and brain coordination. *Neural Networks, 37,* 120-131.

Keltner, D. & Robinson, R. (1993). Imagined ideological differences in conflict escalation and resolution. *International Journal of Conflict Management, 4(3),* 249-262.

Lahnakoski, J.M., Glerean, E., Jääskeläinen, I.P., Hyönä, J., Hari, R., Sams, M., & Nummenmaa, L. (2014). Synchronous brain activity across individuals underlies shared psychological perspectives. *NeuroImage, 100,* 316-324.

Liu, Y., Piazza, E.A., Simony, E., Shewokis, P.A., Onaral, B., Hasson, U., & Ayaz, H. (2017). Measuring speaker-listener neural coupling with functional near infrared spectroscopy. *Scientific Reports, 7,* 43293.

Lu, K. & Hao, N. (2019). When do we fall in neural synchrony with others? *Social Cognitive and Affective Neuroscience, 14(3),* 253-261.

Maraun, D. & Kurths, J. (2004). Cross wavelet analysis: Significance testing and pitfalls. *Nonlinear Processes in Geophysics, 11(4),* 505-514.

Mayo, O. & Gordon I. (2020). In and out of synchrony – Behavioral and physiological dynamics of dyadic interpersonal coordination. *Psychophysiology, 57(6),* e13574.

McGrath, J.E. (1984). *Groups: Interaction and Performance.* New Jersey: Prentice-Hall, Inc.

Nastase, S.A., Gazzola, V., Hasson, U., & Keysers, C. (2019). Measuring shared responses across subjects using intersubject correlation. *Social Cognitive and Affective Neuroscience, 14(6),* 667-685.

Nguyen, M. Vanderwal, T., & Hasson, U. (2018). Shared understanding of narratives is correlated with shared neural responses. *NeuroImage, 184,* 161-170.

Overwalle, F.V. & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage, 48,* 564-584.

Pan, Y., Cheng, X., Zhang, Z., Li, X., & Hu, Y. (2017). Cooperation in lovers: An fNIRS-based hyperscanning study. *Human Brain Mapping, 38(2),* 831-841.

Regev, M., Simony, E., Lee, K., Tan, K.M., Chen, J., & Hasson, U. (2019). Propagation of information along the cortical hierarchy as a function of attention while reading and listening to stories. *Cerebral Cortex, 29(10),* 4017-4034.

Spiers, H.J. & Maguire, E.A. (2006). Spontaneous mentalizing during an interactive real world task: An fMRI study. *Neuropsychologia, 44(10),* 1674-1682.

Stephens, G.J., Silbert, L.J., & Hasson, U. (2010). Speaker-listener neural coupling underlies successful communication. *Proceedings of the National Academy of Sciences, 107(32),* 14425-14430.

Tang, H., Mai, X., Wang, S., Zhu, C., Krueger, F., & Liu, C. (2016). Interpersonal brain synchronization in the right temporo-parietal junction during face-to-face economic exchange. *Social Cognitive and Affective Neuroscience, 11(1),* 23-32.

Thornton, M.A. & Mitchell, J.P. (2017). Consistent neural activity patterns represent personally familiar people. *Journal of Cognitive Neuroscience, 29(9),* 1583-1594.

Wheatley, T., Kang, O., Parkinson, C., & Looser, C. E. (2012). From mind perception to mental connection: Synchrony as a mechanism for social understanding. *Social and Personality Psychology Compass*, *6*(8), 589–606.

Yeshurun, Y., Swanson, S., Simony, E., Chen, J., Lazaridi, C., Honey, C.J., & Hasson, U. (2017). Same story, different story: The neural representation of interpretive frameworks. *Psychological Science, 28(3),* 307-319.

Zadbood, A., Chen, J., Leong, Y.C., Norman, K.A., & Hasson, U. (2017). How we transmit memories to other brains: Constructing shared neural representations via communication. *Cerebral Cortex, 27,* 4988-5000.

# Chapter 5 – General Discussion

## Overview of Findings

Social neuroscience as a field has developed largely through highly controlled investigations of socially isolated individuals. This means that a large frontier of research remains on real-world and naturalistic social processes like social communication and coordination. While emerging theory suggests that people's mental states begin to merge during an interaction and that this shared cognition enables more efficient coordination, the current literature does not examine if this can still happen when there are underlying difference between people, or when the social interaction involves heterogenous thought patterns among people.

This dissertation aimed to help answer these questions. First, in Chapter 2, I discussed how traditional methods of neuroimaging like fMRI and EEG are not best suited for social interaction research, and why the relatively underutilized method fNIRS should be given more consideration in this application. fNIRS is more robust to participant motion and provides for more flexible use cases, meaning it is less restrictive on what a participant or participants can be doing during an experiment. However, due to its underutilization, optimal data processing strategies are still in development. Thus, I showed two ways in which fNIRS data collected during dynamic social communication experiments can be evaluated for quality in order to improve measurement accuracy and inference quality.

In Chapter 3, fNIRS was used to evaluate how neural patterns between a speaker and listener in a personal disclosure situation may be synchronous, reflecting shared understanding of the communicated content. This study also evaluated how similarity of experience and similarity of identity between the speaker and listener may modulate the extent of this synchronization. The results showed that there was relationship between these similarity measures and how well listeners of a narrative understood it the way the speaker did, but similarity of experience and identity did associate with how these listeners understood the narratives in general. More specifically, similar experience led to a more idiosyncratic understanding of the semantic content in the stories as well as more divergent neural patterns, suggesting that having personal memories related to the narrative may trigger various mental representations that are unique to the listener. In contrast, similarity of identity brought people more towards a single canonical understanding of the stories they listened to. This work highlights factors within a social communication that can influence the extent to which people mentally coordinate and converge.

The research in Chapter 4 aimed to probe synchronous neural dynamics as well, this time within mutually interacting discussion partners during a joint decision-making task. In addition, rather than focusing on the particular kind of synchrony of Chapter 3 where similar mental states are expected between people, this study compared different conceptualizations of synchrony to see which might best describe the nature of neural coordination in dyadic interaction – linear synchrony, nonlinear synchrony, and higher-order coordination. The results of this work show that while some linear synchrony was identifiable in interacting dyads, nonlinear and higher-order synchrony was more descriptive of the neural dynamics in the overall conversation, identified temporal increases, and was associated with conversation efficiency. It also highlighted the need to be theoretically and analytically specific about what type of coherence of a social system is being sought in a research study.

## Implications

While not a complete evaluation of all social communication and coordination contexts, this work contributes to our understanding of how humans successfully interact in social contexts. Where it is important that one message is communicated and understood between people, synchronized neural activity reflects how well that mental representation is reinstantiated in a new person. Yet, how accurately one encodes the particulars of that message does not necessarily predict how much empathy one feels towards the communicator, or even how well one infers the speakers feelings about that message. Thus, while remembering the specific details of a message may be related to specific patterns of brain activity, asynchronous activity and understanding may still lead to socially desirable outcomes such as empathy and emotion interpretation.

Relatedly, an exact match in neural patterns is not the only type of coordination that can exist in a social interaction. Interlocutors can also exhibit synergistic states that are not the same, but that still predict each other reliably, as well as higher-order coordination tied to the temporal progression of events within a conversation. This nicely illustrates that social minds might not necessarily “merge” in the strictest sense when the goals of the interaction include comparing and discussing different sources of information and opinion, but are still mutually dependent within a coherent social system.

This work also makes contributions to the wheelhouse of methods for studying natural social interaction. By improving our approach to evaluating fNIRS data, we can make rich and externally valid social neuroscience more common, which would in turn improve our general understanding of the human brain. In addition, this work contributes a new approach to analyzing dynamic interaction data collected in this way, which refines existing theory about how brains coordinate in conversation and provides new avenues to explore within other types of social interactions.

## Future Directions

Due to the richness of social communication data, there is room for further exploration within the data of the presented studies. Chapter 3 explored factors about the speaker and listeners that affect the extent of neural synchronization between them, but the messages themselves can also be explored as variable sources of entrainment. Within a narrative, certain moments may bring listeners particularly in line with each other or with the speaker. These could potentially be attributed to periods of high emotionality, naming of personal emotions rather than just detail description, or other unexpected factors.

This temporally-sensitive approach to analyzing data could also be applied to the study in Chapter 4. While linear synchrony was not a very accurate model of the relationship between dyad members' neural activity in conversation, there could be moments in time where this is more pronounced. It would be valuable to the study of social interaction to know what these moments may mean – e.g., periods of speaker-listener role-taking, mutual insight, coordinated action, etc. The stability of these possible periods could also be investigated as evidence that a weakly-coupled social system is developing stronger connections. Alternatively, the direction that signals seem to be leading each other at various points in time could be used to predict who is currently leading a conversation, enabling investigation of turn-taking behaviors in dyadic discussions.

Beyond this work, there is still extensive space for exploring shared neural dynamics is a variety of different social contexts, and the factors that may modulate those dynamics. Situations like negotiation, interpersonal instruction, or team task work are understudied in social neuroscience due to the complexities of these events and the difficulties of collecting data in them. The approaches described in this dissertation could be brought to bear on investigating these poorly-understood social contexts. Further, if enough descriptive research is performed, it could be possible to track the extent of coherence between people in social interactions as a predictor of successful communication. This would be valuable for arranging teams within organizations, identifying points of strength of weakness within instructional lessons, and for training of social support figures like therapists.

## Final Conclusions

Social interaction is one of the most difficult contexts to study empirically, due to its complexity and diversity in the real world. Yet it is also one of the most important to understand, as so much of human life is lived in and optimized for social communication and coordination. The research in this dissertation contributes new insights into the neural and psychological dynamics underlying social interaction in a few different contexts. It also demonstrates how these dynamics relate to heterogeneities among members of the interaction, such as their background knowledge or individual opinions. Lastly, it provides methodological contributions for continuing this line of work. This all will hopefully benefit the pursuit of understanding human psychology and behavior in context, the natural way in which we all live.