

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Psychology

2020

Shared and idiosyncratic neural processing of naturalistic communication in the default mode network

Nguyen, Mai Lin

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp011g05ff53g>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

Shared and idiosyncratic neural processing of naturalistic communication in the default mode network

Mai Lin Nguyen

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE

BY THE DEPARTMENT OF PSYCHOLOGY

Advisors: Uri Hasson & Alexander Todorov

June 2020

© Copyright by Mai Lin Nguyen, 2020. All rights reserved.

Abstract

Humans have a unique ability to share knowledge of the world with each other via communication. A finite set of spoken sounds and written symbols can be used in as diverse situations as telling a personal story, ordering coffee, or sharing scientific results. However, the remarkable flexibility in the form and content of communication is matched with inherent ambiguity in its interpretation. A listener might misunderstand a speaker, lack the appropriate context or background to process the information, or simply interpret the same information in a different way than is intended. This dissertation aims to elucidate neural processes underlying shared and idiosyncratic interpretation of naturalistic communication. We argue that shared understanding of information across forms of communication is reflected in the shared dynamics of neural responses in the default mode network, both among receivers of information and between senders and receivers. Using a novel method, intersubject representational similarity analysis, we first show that regions of the DMN are sensitive to individual differences in the interpretation of an abstract, ambiguous narrative: across individuals, similarity in narrative interpretation is correlated to similarity in neural activity (Chapter 2). At the same time though, we find that the DMN is remarkably invariant to the modality of the narrative. The same narrative presented in strikingly different forms elicits similar neural responses in the DMN, but only when interpretation of the narrative is shared across forms (Chapter

3). We then demonstrate that shared neural responses in DMN emerge from the interaction of brains in communication in a non-narrative context. We find that the neural activity of teachers and students learning abstract technical information are coupled during communication and that the magnitude of this coupling is related to learning outcomes (Chapter 4). We ultimately argue for a new view of the DMN as a system for integrating internal representations with incoming, external information, resulting in both shared and idiosyncratic neural representations over long timescales. This unique interplay between extrinsic and intrinsic forces enables us to both shape and be shaped by other minds and provides a mechanism for negotiating a shared neural code underlying communication (Chapter 5).

Acknowledgements

These are immensely difficult times. The COVID-19 pandemic has radically altered our way of life, and we still do not know how this will end. If there is one bright side though, the weeks and months of uncertainty and social isolation have driven home how immensely fortunate I am in being surrounded and supported by profoundly kind, funny, intelligent, generous, and inspiring family, friends, and colleagues. I am truly grateful for all you have done, all you are, and your presence in my life.

I would like to start by thanking my advisor, Uri Hasson, whose work, ideas, support, and patience utterly changed my scientific interest and thinking and who made all this work possible. Seven years ago, I thought letting people watch movies (looking wherever they want!) and talk in the scanner was nothing less than absurd; now I cannot imagine studying psychology or cognition any other way. I also thank my secondary advisor, Alex Todorov, for putting up with me for six years, feeding me Balkan burgers, and of course giving me his support, time, and encouragement. I am grateful to my reading committee, Adele Goldberg and Ken Norman, and my oral committee, Alex Todorov and Diana Tamir, for their time, feedback, and support. Finally, an enormous thank you to the people who set me on this path to graduate school in the first place: Kalanit Grill-Spector and Davie Yoon at Stanford who gave a chance to a freshman who knew absolutely nothing and prepared me for a successful graduate career; Nathan Witthoft, also at Stanford, who taught me as much about life as about science and who

was truly there when I needed help; Greg Davis and Kate Plaisted-Grant at University of Cambridge, who welcomed and supported me in an adventure to a new country; and Jonathan Winawer at NYU, whose meticulousness, careful and deep thought, and generosity I strive to emulate in my own work.

Thank you to the entire Hasson Lab for comradery, insight, support, caring, commiserating, knowledge, mentorship, and generosity which sustained me through this PhD. A special thank you especially to Hassonian lab mates, Janice Chen, Chris Baldassano, Yaara Yeshurun, Chris Honey, Liat Hasenfratz, Meir Meshulam, Sam Nastase, Ariel Goldstein, Claire Chang, and Hanna Hillman and my fellow graduate students, Asieh Zadbood and Mor Regev. I also thank my immense rotating cast of officemates ($N=13$): Yaara, Ido Davidesco, Michael Chow, Claire, Hanna, Ariel, Jamal Williams, Aditi Rao, Chris B., Bobbi Aubrey, Zaid Zada, Gina Choe, and Eric Ham. Thank you as well to the entire Todorov Lab, especially Aaron Kurosu, DongWon Oh, Joel Martinez, and Stephan Uddenberg. Thank you to my amazing summer undergraduate research assistants, Ashley Chang and Emily Micciche.

Outside of my labs, I am deeply appreciative of the support from many people in the PNI/Psych community not already mentioned: my entire cozy cohort (past and present), Aaron, Brandy Briones, Angela Radulescu, Nina Rouhani, Cameron Ellis, Victoria Ritvo, Jessie Schwab, Kyonne Isaac, and Yaritza Perez-Hooks; professors Casey Lew-Williams, Justin Junge, Ting Qian; the Princeton Computational Memory Lab; the Princeton Social Neuroscience Lab; the MRI support staff, Nick DePinto, Leigh Nystrom, and Mark Pinsk; the main office staff, especially Keisha Craig and RoseMarie Stevenson; the IT support staff; my dozens and dozens of scanning buddies who even

came in even on weekends, early mornings, and Friday nights; data archivists Neggin Keshavarzian and Arlene Lormestoire; and Remy/Dez, Fawkes/Rinn, Eurodora/Joni, Vervain/Terry, Glaam, Golden Eyes, Islay, Daana, Spugnoir, and Samuel A. Nastase, for not leaving Lady Adamaris/Vorgak the Virtuous for dead (evil which cannot be removed must be eliminated).

Graduate Student Government played an unexpected, but deeply rewarding, role in making my time at Princeton a rich and meaningful one. I thank all of my fellow GSG Executive Committee members, especially Jonathan Balkind, Ashkay Mehra, Emily Kern, Emily Cuddy, Sarah-Marie Bruno, Noah Apthorpe, Laura Bustamante, Karina Alventosa, Lauren Feldman, and Brandy Briones. I also thank the members of the Graduate Mental Health Initiative, Sarah-Marie, Mira Nencheva, Brian Kraus, Brianna Romer, Abigail Novick, Victoria Ritvo, Judith Mildner, and Sarah Wilterson. I am grateful for support of many university administrators that I had the pleasure to work with: Dean Sarah-Jane Leslie, Dean Lisa Schreyer, Dean Nicole Barkley, Dean Cole Crittenden, Vice President Rochelle Calhoun, Dr. Laura Murray, Dr. Calvin Chin, Vice Provost Michele Minter, Charles Tennyson, and Dorian Johnson.

Finally, to all of my amazing, incredible, and brilliant friends in graduate school, undergrad, and the Jeesh, much love and a heartfelt thank you. To my family, Mom, Dad, Aimee, and Tata, I could not have done it without you. Your constant love and support mean more to me than anything. To my partner, Shawn Westerdale, let's touch eyeballs.

Mai Lin Nguyen

April 2020

To M. Butt, who supported me for all these years

Author's Note

Chapters 1 and 5 were adapted from a manuscript in preparation:

Yeshurun, Y. *, Nguyen, M. *, & Hasson, U. (2020). Beyond intrinsic: a new view of the default mode network.

Chapters 2 and 3 were adapted from publication in a peer-reviewed journal:

Nguyen, M., Vanderwal, T., & Hasson, U. (2019). Shared understanding of narratives is correlated with shared neural responses. *NeuroImage*, 184: 161-170.

Chapter 4 was adapted from a manuscript submitted to a preprint server:

Nguyen, M., Chang, A., Micciche, E., Meshulam, M., Nastase, S., & Hasson, U. (2020). Teacher-student neural coupling during teaching and learning. *bioRxiv*.

* Denotes joint first authorship

April 2020

Table of Contents

Abstract	i
Acknowledgements.....	iii
Author's Note.....	vii
List of Figures	xii
1 Introduction.....	1
1.1 The default mode network and internally-oriented thought	3
1.2 DMN integrates internal and external information over long timescales	5
1.2.1 A hierarchy of timescale processing	5
1.2.2 Integration in DMN depends on content, not form.....	7
1.2.3 DMN is sensitive to context.....	8
1.3 DMN is coupled between senders and receivers during naturalistic communication	10
1.4 Outline of dissertation	11
2 Shared narrative interpretation is correlated with shared neural responses	14
2.1 Methods.....	16
2.1.1 Subjects	16
2.1.2 Stimuli and experimental design.....	16
2.1.3 Stimulus presentation.....	17

viii

2.1.4	MRI acquisition	17
2.1.5	Behavioral analysis	18
2.1.6	MRI data analysis	19
2.2	Results	25
2.2.1	Behavioral results.....	25
2.2.2	Neural results	27
2.3	Discussion.....	30
2.3.1	Processing in the default mode network	31
2.3.2	Processing outside the DMN	33
2.3.3	Limitations and alternate interpretations	34
2.3.4	Intersubject RSA.....	35
2.4	Conclusions.....	36
3	Shared default mode network responses across modality depend on shared interpretation.....	37
3.1	Methods.....	39
3.1.1	Subjects.....	39
3.1.2	Stimuli and experimental design.....	40
3.1.3	Stimulus presentation.....	42
3.1.4	MRI acquisition	42
3.1.5	Behavioral data analysis	43
3.1.6	MRI data analysis	43
3.2	Results	45
3.2.1	Behavioral results.....	45

3.2.2	Neural results	47
3.3	Discussion.....	52
3.3.1	Shared processing of narratives across modality	53
3.3.2	Shared processing of visual features across modality.....	54
3.3.2	Shape animations and mentalizing.....	54
3.4	Conclusions.....	55
4	Teacher-student neural coupling during teaching and learning	56
4.1	Methods.....	59
4.1.1	Subjects	59
4.1.2	Experimental design.....	59
4.1.3	MRI acquisition	62
4.1.4	MRI data analysis	63
4.2	Results	67
4.2.1	Behavioral results.....	67
4.2.2	fMRI results	69
4.3	Discussion.....	75
4.3.1	Limitations and future directions	78
4.4	Conclusions.....	80
5	General discussion	81
5.1	A new view of the default mode network.....	82
5.1.2	Individual differences and the DMN: integrating internal representations with ongoing information	83

5.1.3	Communication and DMN: integrating self-input with input from other minds	87
5.2	Importance of naturalistic paradigms	90
5.3	Future directions and open questions	92
5.3.1	Brain-to-brain coupling over time	92
5.3.2	Intersubject representational similarity analysis and individual differences.....	94
5.3.3	Towards the neuroscience of social interaction	95
5.4	Concluding remarks.....	99
References.....		100
Appendices.....		121
Appendix 1	Animate shapes movie	121
Appendix 2	Theory of Mind audio story.....	122
Appendix 3	Physical audio story.....	125
Appendix 4	fMRI video lesson.....	129
Appendix 5	fMRI video review.....	146
Appendix 6	fMRI quiz questions	150

List of Figures

Figure 1.1	Default mode network using functional connectivity during movie watching	3
Figure 2.1	Shapes movie stimulus	17
Figure 2.2	ISC and intersubject RSA procedure	21
Figure 2.3	LSA results among shape movie recalls	25
Figure 2.4	ISC among movie subjects with similar versus dissimilar recalls	26
Figure 2.5	Intersubject RSA among movie subjects	28
Figure 2.6	ISC and intersubject RSA on motion residuals	29
Figure 3.1	Theory of mind and physical audio stimuli	41
Figure 3.2	LSA recall similarity results for movie and audio recalls	46
Figure 3.3	Cross-modal ISC between Shapes Movie and Audios	48
Figure 3.4	Differences in cross-modal ISC between similar and dissimilar recall groups	50
Figure 3.5	Cross-modal intersubject RSA	51
Figure 4.1	Teaching and learning fMRI task design	60
Figure 4.2	Behavior: Learning scores and attention	68
Figure 4.3	Student-Student and Teacher-Student ISC during teaching and learning	70
Figure 4.4	Lagged ISC for Student-Student and Teacher-Student coupling	71
Figure 4.5	Teacher-Student ISC in Intact versus Unscrambled Lesson	72

Figure 4.6	Teacher-Student ISC correlates with learning outcomes	74
Figure 5.1	DMN integrates internal and external information over long timescales.....	84

Chapter 1

Introduction

Human communication is remarkably flexible. We use the same finite set of sounds and written symbols to share stories, order coffee, tell jokes, collaborate with coworkers, pay taxes, and more. Indeed, communication subserves almost every aspect of modern human life. The near effortlessness of everyday communication, however, belies its complexity. On the one end, successful communication requires a person who is communicating something (“senders”) to take information represented in the neural encoding of their brain and convert it to an external signal. Typically senders might do so by modulating air pressure through their larynx via contraction and retraction of vocal chords (speech) or by inscribing specific lines on flat materials via fine motor control of hands and tool use (writing). People receiving information (“receivers”) must then input these signals via auditory and visual senses and process this information in their own brains (Hasson et al., 2012). Despite the flexibility in what is communicated and how it is communicated, the process of communication is thus fraught with ambiguity and potential pitfalls. Receivers may misunderstand the sender, lack the appropriate context or background to process the information, or simply interpret the same information in a different way than is intended.

Communication therefore requires ongoing, continuous interpretation that may result in divergence of understanding among receivers or between senders and receivers.

Researchers are only beginning to elucidate the neural processes and mechanisms that enable humans to process and communicate information with each other. Much of this work has focused on how isolated sounds, words, or short phrases are represented in individual brains during either production or comprehension of spoken or written language and other forms of communication. While such research has identified a distinct network of regions that process linguistic information (e.g. Price, 2010; Friederici and Gierhan, 2013; Ardila et al., 2016), the typical experimental design is highly artificial with little resemblance to real-life communication: words and phrases are temporally isolated from each other, stimuli lack coherent interpretation, and individuals are separated from each other and the broader context. These artificial designs may in fact “control away” aspects of neuropsychological paradigms that are of significant interest (Ben-Yakov et al., 2012; Hasson and Honey, 2012; Hasson et al., 2012; Schilbach et al., 2013; García and Ibáñez, 2014; Liu and Pelowski, 2014; Kasai et al., 2015; Nastase et al., 2020).

This dissertation explores the neural processes underlying both shared and idiosyncratic interpretation of information communicated in naturalistic contexts. We argue that shared understanding of information, among groups of receivers and between senders and receivers, is reflected in the shared dynamics of neural responses in the default mode network, irrespective of communication form. Ultimately, we put forward a new view of the DMN as one that integrates information from the past with the present to form rich, context-dependent, idiosyncratic models of events as they dynamically unfold

over time.

1.1 The default mode network and internally-oriented thought

In 1997, Shulman et al. observed that activity in widespread regions of the brain outside of sensory cortices showed lower neural activity during passive viewing of images than during baseline rest. Following this discovery (Shulman et al., 1997; Gusnard and Raichle, 2001; Raichle et al., 2001; McKiernan et al., 2003), spontaneous fluctuations in activity in these regions were shown to be highly correlated with each other, suggesting that the regions form a distinct functional network (Greicius et al., 2003; Yeo et al., 2011; Power et al., 2012; Schaefer et al., 2018). Based on these findings, researchers suggested that cortical areas are divided into an “extrinsic” system that processes incoming sensory information (e.g. visual or auditory inputs) and an “intrinsic” system, also termed the default mode network (DMN), that specializes in internally-oriented or non-stimulus induced thought (Raichle et al., 2001;

Golland et al., 2006; Andrews-Hanna et al., 2014; Buckner and DiNicola, 2019).

While there are some variations in its definition, the DMN is widely agreed to include posterior medial cortex (PMC), medial prefrontal cortex (MPFC), dorsolateral prefrontal cortex (DLPFC), temporoparietal junction (TPJ) and angular gyrus (ang), and in some cases,

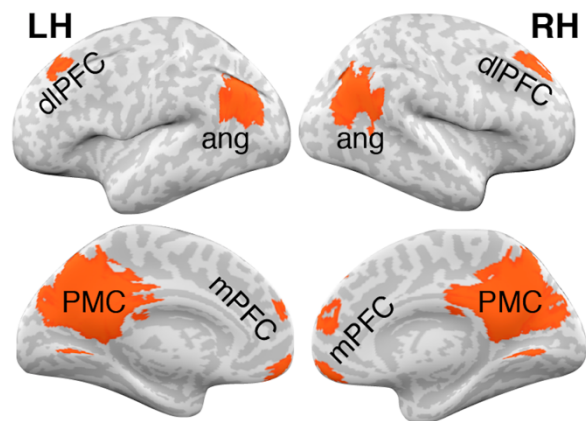


Figure 1.1 Default mode network defined using functional connectivity during movie watching. The DMN includes posterior medial cortex (PMC), ventral medial prefrontal cortex (vmPFC), dorsolateral prefrontal cortex (dlPFC), and temporal parietal junction/angular gyrus (ang).

temporal poles (Andrews-Hanna et al., 2014; Raichle, 2015; Buckner and DiNicola, 2019).

The function of the DMN is a matter of substantial interest and debate. The DMN was initially conceptualized as a “task-negative” network, active only when the brain is not otherwise occupied by a task (Raichle et al., 2001; Buckner et al., 2008). Some have suggested that this deactivation during tasks meaningfully allows for increased performance on goal-directed tasks (Anticevic et al., 2012). Indeed, deactivation of the DMN scales with cognitive demand of tasks (McKiernan et al., 2006; Jansma et al., 2007; Singh and Fawcett, 2008; Vatansever et al., 2015; Krieger-Redwood et al., 2016). In addition, activity in DMN is negatively correlated with activity in the Dorsal Attention Network (Fox et al., 2005; Fransson, 2005), which is recruited during externally oriented tasks (Corbetta and Shulman, 2002).

However, later research showed that the DMN is activated by a range of tasks that are characterized by internally-oriented thought (Spreng et al., 2009; Andrews-Hanna et al., 2014; Christoff et al., 2016; Buckner and DiNicola, 2019; Mildner and Tamir, 2019). These tasks include episodic memory retrieval (Buckner and Carroll, 2007; Schacter et al., 2007; Buckner et al., 2008; Andrews-Hanna et al., 2010), prospection and future planning (Schacter et al., 2007; Spreng et al., 2009; Vatansever et al., 2018), scene construction (Hassabis and Maguire, 2007, 2009; Costabile, 2016), and semantic memory (Binder et al., 2009). Others have highlighted substantial overlap between DMN and the mentalizing network, a set of regions that are recruited during theory of mind and social cognition (Iacoboni et al., 2004; Schilbach et al., 2008; Spreng et al., 2009; Mars et al., 2012; Spunt et al., 2015). Indeed, some have asserted that the “default mode” of the brain

is thinking about others and other minds (Schilbach et al., 2008). Moreover, researchers have noted that during so-called “rest” periods, subjects engage in rich internal cognition (Andrews-Hanna et al., 2010; Spreng et al., 2012), and the DMN is recruited during these periods of spontaneous thought or mind wandering (Mason et al., 2007; Christoff et al., 2009; Fox et al., 2015; Kucyi et al., 2016).

In this dissertation we challenge the view of the DMN as an intrinsic system that is engaged only during internally-directed or non-stimulus induced cognition. Rather, we suggest that the DMN plays an active and dynamic role in integrating incoming extrinsic information with intrinsic prior information represented in memory. The DMN ultimately forms rich, context dependent, and idiosyncratic models of situations as events unfold over time, enabling individuals to both make sense of the world and to shape and share it with others.

1.2 DMN integrates internal and external information over long timescales

1.2.1 A hierarchy of timescale processing

We begin by showing that the DMN integrates external, stimulus-ongoing information with internal representations of the stimulus past. Such integration over time is necessary for understanding temporally extended, multidimensional information in real-life. Unlike the typical experiment in psychology and neuroscience that utilizes temporally isolated and punctate stimuli, events in real-life take place over multiple timescales.

Understanding the “now” thus requires integrating incoming information with internally stored information from seconds, minutes, hours, and more from the past.

Matching the hierarchical structure of real-life events, widespread regions of the brain are organized in a topographic hierarchy of increasing temporal receptive windows (TRWs), the window of time in which past information affects neural processing of new information (Hasson et al., 2008; Jaaskelainen et al., 2008; Kiebel et al., 2008; Wilson et al., 2008; Honey et al., 2012a; Murray et al., 2014; Hasson et al., 2015; Ding et al., 2016). The hierarchy of timescale processing was identified by temporally scrambling naturalistic stimuli such as movies, TV shows, or spoken stories, at different levels (word/phoneme, sentence, paragraph). This manipulation controls the amount of prior information that is available to the subject at every moment of the stimulus. Regions of the brain exhibiting significantly correlated neural responses across individuals at each level of scrambling were then identified using intersubject correlations (ISC). ISC is calculated by correlating each subject’s response timecourse in a given brain area during a stimulus with the response timecourse of other subjects in the same area. Because subject-specific signals and noise will not be correlated across individuals, correlated responses suggest that activity in a particular cortical regions is driven by shared processing of stimulus features (Hasson et al., 2004, 2008; Finn et al., 2019; Nastase et al., 2019a). Brain regions that are not sensitive to the stimulus or features of the stimulus will not show significantly correlated responses across individuals.

Early sensory areas such as A1+ and V1+ show significant ISC regardless of the level of scrambling, consistent with short timescales on the order of tens of milliseconds. These sensory areas therefore predominantly process fast-changing, low-level stimulus

features such as sound amplitude or visual edges. In contrast, regions of the DMN only respond reliably to intact, meaningful stimuli, suggesting that they integrate information from minutes to hours or longer in order to make sense of the overall narrative (Hasson et al., 2008; Lerner et al., 2011; Honey et al., 2012a). These findings were extended using intersubject functional connectivity (ISFC), which correlates the response timecourse of subjects across regions rather than in the same region as in ISC. Using ISFC, researchers showed that stimulus-locked DMN connectivity increases with increased story coherence. Network dynamics within the DMN are therefore sensitive to temporal structure of naturalistic stimuli (Simony et al., 2016).

1.2.2 Integration in DMN depends on content, not form

Shared neural responses in the DMN are locked to the content of the stimuli and are relatively invariant to even dramatic changes in low-level stimulus properties. Despite substantial differences in visual and/or acoustic features, the DMN shows similar timecourses of activity in response to the same narrative presented in written text versus speech (Regev et al., 2013; Wilson et al., 2018), in English versus Russian language (Honey et al., 2012b), in audiovisual movie versus verbal descriptions (Zadbood et al., 2017; Baldassano et al., 2018; Tikka et al., 2018), and using different synonyms (Yeshurun et al., 2017b). A similar pattern of results was also identified using encoding models across written and spoken narratives (Deniz et al., 2019) and decoding models across three languages (Dehghani et al., 2017). In addition, the size of the TRW scales with information rate: compressing or dilating the rate of speech in a story produces corresponding scaling of TRW in neural responses (Lerner et al., 2014).

At the same time, a relatively small change to the stimulus that changes its interpretation can produce increasingly large differences in neural responses across the timescale hierarchy. Yeshurun, Nguyen, and Hasson (2017a) manipulated a small number of words in a spoken story (e.g. changing “laughing” to “sobbing”), resulting in two very different stories. The difference in neural response elicited by the two stories was significantly correlated with TRW: early auditory cortex showed little differences between the stories, reflecting the relatively intact low-level acoustic properties of the stimuli, while DMN regions showed large differences, reflecting the large difference in story meaning.

1.2.3 DMN is sensitive to context

Going beyond the stimulus, the DMN integrates surrounding contextual information that enables or even changes interpretation of a narrative. Researchers have exploited stimuli such as TV shows, stories, and descriptions that are difficult to understand without prior context encoded in memory. Chen et al. (2015a), for example, scanned subjects in fMRI watching the second half of an episode of the *Twilight Zone*, which was difficult to understand without the context of the first half of the episode. One group of subjects watched the first half of the episode prior to scanning, while a second group did not. Chen et al. found greater DMN intersubject synchronization among participants who had the background knowledge to interpret the episode compared to subjects who did not. Similar findings have been observed using movies (van Kesteren et al., 2010; Oren et al., 2017) and verbal descriptions of scenarios (Ames et al., 2015), suggesting that DMN

activity is sensitive to contextual information that enables the coherent interpretation of the stimulus.

Beyond simply enabling interpretation, other studies manipulated context to change the interpretation of a stimulus (Yeshurun et al., 2017b; Chien and Honey, 2020). Yeshurun et al. (2017), for example, presented two groups of subjects with the same ambiguous short story by J.D. Salinger, “Pretty Mouth and Green My Eyes.” The two groups were biased with different contextual information prior to listening to the story, significantly changing its interpretation: one group was told the main character’s wife was cheating on him, while the second group was told the main character was irrationally paranoid and had a faithful wife. Subjects showed significantly more similar DMN responses within interpretation group than between. Although the two groups heard the exact same narrative, they interpreted it differently and so showed different patterns of neural activity in DMN. These findings demonstrate that the DMN represents interpretation and shared understanding, not merely semantic content. Similar findings were observed by manipulating perspective (Lahnakoski et al., 2014; Bacha-Trams et al., 2017) or focus (Cooper et al., 2011), as well as when context led participants to interpret the same situation as ironic (Uchiyama et al., 2012), trying to “save face” (Bašnáková et al., 2014), or intentionally harmful (Koster-Hale et al., 2013). A recent study suggests that regions within the DMN are involved in reinstatement of the context of a storyline, which facilitates convergence of the plot of two storylines into one (Chang et al., 2020). The DMN is therefore sensitive not just to the content within a stimulus, but to the surrounding context.

1.3 DMN is coupled between senders and receivers during naturalistic communication

Research has shown that shared activity in DMN is driven by both internal and external sources of information. That external information took the form of physical stimuli, such as spoken words, written texts, gestures, and audiovisual movies. But the brain can also be driven by other brains: stories, memories, and other ideas are generated by one brain and then transmuted into physical signals such as words, sounds, facial expressions or gestures. These signals are detected by sensory receptors in a second brain, eliciting neural activity that is driven by, or coupled to, activity in the first brain. This “sender-receiver” model of communication therefore extends neural alignment among receivers of information to neural coupling between individuals sending and receiving information.

Brain-to-brain coupling was first demonstrated during unidirectional verbal communication. In a landmark study, Stephens et al. (2010) scanned a speaker telling a personal story in fMRI. The story was then presented to listeners during a later fMRI scan. The speaker’s neural responses were significantly correlated with the listener’s neural responses at many levels of the timescale processing hierarchy. Speaker-listener coupling in early auditory regions reflected shared processing of low-level acoustic properties of stimulus, such as the amplitude envelope. In contrast, alignment in higher-order areas, including the DMN, reflected communication and shared understanding of the narrative. Speaker-listener coupling in DMN, but not early auditory regions, was disrupted when communication was disrupted by speaking only nonsense words (Stephens et al., 2010; Silbert et al., 2014), telling a story in a foreign language that the listener does not understand (Stephens et al., 2010), or temporally scrambling the

stimulus (Nguyen et al., 2020b). A recent study suggested that only the processing of disambiguated, coherent discourse with an intelligible situation model representation involved shared activation in DMN between the speaker and the listener (Heidlmayr et al., 2019). Moreover, speaker-listener coupling in regions of the DMN were correlated with success of communication: the better the listener understood the speaker, the closer their brain matched the speaker's (Stephens et al., 2010; Dikker et al., 2014; Silbert et al., 2014; Nguyen et al., 2020b). These findings demonstrate that comprehension-based processing in the listener's brain is driven by, or coupled to, production-based processing in the speaker's brain.

Speaker-listener coupling has been observed using a variety of social narratives and imaging modalities including fMRI (AbdulSabur et al., 2014; Zadbood et al., 2017; Heidlmayr et al., 2019; Smirnov et al., 2019; Nguyen et al., 2020b), fNIRS (Piazza et al., 2004; Liu et al., 2017; Hirsch et al., 2018), and EEG (Kuhlen et al., 2012). Sender-receiver coupling has also been observed during non-verbal communication such as gestures (Schippers et al., 2010) and facial expressions (Anders et al., 2011). Subsequent work showed that communication serves to transmit memories and construct shared representations of naturalistic events (Chen et al., 2016b; Zadbood et al., 2017). Together, this work suggests that brain-to-brain coupling may underlie successful communication of social narratives.

1.4 Outline of dissertation

Previous research has highlighted the role of DMN in narrative comprehension and communication. Correlated neural responses across individuals in these regions are

locked to processing of external, temporally extended stimuli. This dissertation builds on this work to show that the DMN is sensitive to individual, idiosyncratic differences in interpretation or understanding during narrative and non-narrative communication in different forms. We ultimately argue that the DMN is not internally-oriented or intrinsic, but that it facilitates integration of internal representations with incoming information over long timescales, enabling the formation of a shared world.

In Chapter 2, we show that regions of the DMN are sensitive to individual differences in the interpretation of abstract, ambiguous narratives. We suggest that spontaneous, internal representations can modulate the processing of external information. We also introduce a novel method, intersubject representational similarity analysis (iRSA), for identifying individual differences in the processing of naturalistic stimuli.

In Chapter 3, we extend these findings to show that the same narrative presented in different forms elicits similar neural responses in the DMN as a function of shared interpretation. We show striking modality invariance in the processing of narratives in the DMN, but that this shared processing across modality is contingent on shared interpretation of different features of the stimulus.

In Chapter 4, we then demonstrate that shared neural responses in DMN emerge from the interaction of brains in communication. In the context of teaching and learning, we show that teacher and student brains are aligned during non-narrative communication and that the extent of this alignment is related to learning outcomes.

Finally, in Chapter 5, we synthesize these findings in the context of a new view of the function of the DMN, highlight the importance of naturalistic, temporally extended experiments, and put forward proposals for future directions.

Chapter 2

Shared narrative interpretation is correlated with shared neural responses

In daily life, stories and narratives are often ambiguous, necessitating ongoing processing and interpretation to make sense of events. As the result, people may not understand a narrative in the same way and may interpret it differently than was intended. How are individual differences in the interpretations of complex, ambiguous narratives represented in the brain?

Previous work has shown that narratives elicit correlated neural responses across individuals in regions of the default mode network (DMN), including temporal parietal junction (TPJ), angular gyrus, temporal poles, posterior medial cortex (PMC), and medial prefrontal cortex (mPFC) (Hasson et al., 2008, 2010; Jaaskelainen et al., 2008; Wilson et al., 2008; Lerner et al., 2011; Ben-Yakov et al., 2012; Simony et al., 2016; Nastase et al., 2019). The shared response in high-level regions is thought to be driven by the content and interpretation of the narrative, rather than low-level stimulus properties. Reliably correlated neural responses only emerge in DMN regions during processing of intact, coherent narratives (Hasson et al., 2008; Lerner et al., 2011; Chen et al., 2015b; Farbood et al., 2015; Simony et al., 2016) and are insensitive to dilation or compression of the rate

of speech in a narrative (Lerner et al., 2014). Moreover, when the interpretation of a narrative is manipulated using contextualizing information, the shared neural responses in DMN are greater among subjects who shared interpretations than between subjects with contradictory interpretations (Cooper et al., 2011; Ames et al., 2014; Lahnakoski et al., 2014; Yeshurun et al., 2017b; Chien and Honey, 2020).

In past work, the interpretation of narratives was imposed by the researcher, resulting in uniform narrative interpretation within groups. In daily life, however, narratives can be ambiguous and are interpreted in many different ways. To what extent do the spontaneous and unguided interpretations of a complex, ambiguous, and abstract narrative covary with the degree of shared neural responses across subjects? Based on previous work, we predicted that the more similar the interpretation of an ambiguous movie across individuals, the more similar their neural responses in the DMN will be.

To test this prediction, we scanned subjects in fMRI watching a novel, abstract, ambiguous seven-minute animated movie that depicts a narrative through the movement of simple geometric shapes. This movie follows the classic work of Heider and Simmel (1944), but uses a longer, more complex social plot involving multiple characters and abstract scenes that are open to multiple interpretations. While animated shape movies have been extensively used to investigate theory of mind (Heider and Simmel, 1944; Oatley and Yuill, 1985; Berry et al., 1992; Scholl and Tremoulet, 2000), the movie in the present work is unique in its length, number of characters, social relationships, and high-level narrative arc (Appendix 1). Immediately following stimulus presentation, all subjects were asked to freely describe the stimulus. Based on textual analysis of these free recalls, we then compared recall similarity with neural similarity using both

intersubject correlation analysis (ISC) and intersubject representational similarity analysis (RSA). These analyses suggest that regions of the DMN enable the interpretation of narrative under complex, naturalistic conditions, and represent subtle, individual differences in narrative interpretation.

2.1 Methods

2.1.1 Subjects

Thirty-seven adult subjects with normal hearing and normal or corrected-to-normal vision participated in the study. One subject was excluded for falling asleep during the scan, leaving 36 subjects in the final sample (ages 18 – 35, mean 22.5 years; 19 female). The sample size was based on power analyses for reliable intersubject correlation analyses (Chen et al., 2016a; Pajula and Tohka, 2016). All experimental procedures were approved by the Princeton University Internal Review Board, and all subjects provided informed, written consent.

2.1.2 Stimuli and experimental design

Subjects were scanned in fMRI while watching a 7-min animated film (Appendix 1). The Movie depicted a short story using moving geometric shapes in the style of Heider and Simmel (1944). While there was no spoken dialogue, the animation included an original piano score that communicated mood and was congruent with events in the narrative (Figure 2.1). Immediately following stimulus presentation, while still in the scanner,

A. Experimental design

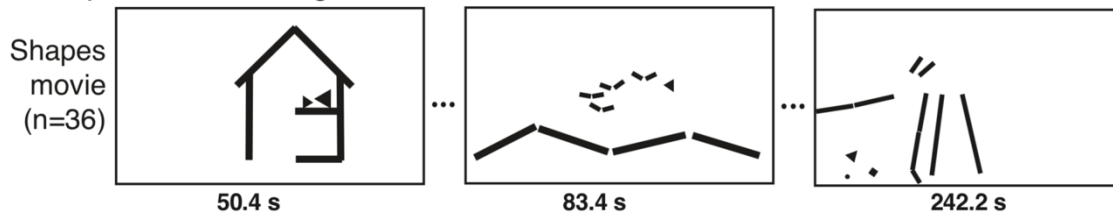


Figure 2.1 Shapes movie stimulus. While being scanned in fMRI, subjects watched a 7-min animation made of moving shapes (Movie). Three frames of the animation are shown here, for the full Movie see Appendix 1.

subjects were asked to freely recall the stimulus using their own words and in as much detail as possible. Recalls were collected during a second functional run using a customized MR-compatible recording system with online sound cancelling. Data from this functional run were not included in analyses here.

2.1.3 Stimulus presentation

Stimuli were presented using MATLAB (MathWorks) and Psychtoolbox (Kleiner et al., 2007). Video was presented by LCD projector on a rear-projection screen mounted in the back of the scanner bore and was viewed through a mirror mounted to the head coil. Audio was played through MRI-compatible insert earphones (Sensimetrics, Model S14).

2.1.4 MRI acquisition

Subjects were scanned in a 3T Magnetom scanner (Prisma, Siemens) using a 64-channel head-neck coil (Siemens). All scanning was conducted at the Princeton Neuroscience Institute Scully Center for Neuroimaging. In the functional scans, volumes were acquired

using a T2*-weighted multiband EPI pulse sequence (TR 1500 ms; TE 39 ms; voxel size 2x2x2mm; flip angle 55°; FOV 192x192 mm², multiband acceleration factor 4, no prescan normalization) with whole-brain coverage. Following functional scans, a fieldmap (mean and phase) was collected (dwell time 0.93 ms; TE diff 2.46 ms). Finally, a high-resolution anatomical image was collected using a T1-weighted MPRAGE pulse sequence (voxel size 1x1x1 mm).

2.1.5 Behavioral analysis

An independent coder blind to the aim of the study segmented the narrative into 22 separate events and then coded each subject's recall for a description of each event. A subject was rated as having recalled an event if they described any part of the event.

Free recalls were then lightly edited to remove non-stimulus related utterances (e.g. "I don't remember," "I'm done,"). The edited recalls were then assessed for similarity to each other within and across stimulus groups using Latent Semantic Analysis (LSA), a statistical method for representing the similarity of texts in semantic space (Figure 2.2A, left). In brief, LSA derives a semantic space via singular vector decomposition (SVD) on the word frequency matrix of a large corpus. Semantic similarity is defined as the cosine distance between texts (e.g. words, phrases, paragraphs or longer) in this space. Semantic similarity measured by LSA has been shown to approach human-like performance in certain ways (Landauer et al., 1998). Here, we used LSA to measure similarity of subject recalls within and across the two stimuli. The semantic space was derived from the Touchstone Applied Science Associates (TASA) college reading-level corpus with 300 factors, as implemented on lsa.colorado.edu.

Finally, to order subjects by similarity to each other for visualization purposes, we then conducted agglomerative hierarchical clustering with complete-linkage on the LSA similarity matrices.

2.1.6 MRI data analysis

2.1.6.1 Preprocessing

MRI data were preprocessed using FSL 5.0 (FMRIB, Oxford) including 3D motion correction, fieldmap correction, linear trend removal, high-pass filtering (140 Hz), and spatial smoothing with a Gaussian kernel (FWHM 4 mm). Motion correction was performed using FSL's MCFLIRT with 6 degrees of freedom, and estimates of both relative, frame-wise movement and absolute movement were extracted. Subjects with excessive head motion (>3 mm absolute movement) were removed from the sample. All data was aligned to standard 2-mm MNI space. Following preprocessing, the first 60 TRs were cropped to remove the introductory videos and transitory changes at the start of the stimulus. Voxels with low mean signal (2 std below average) were also removed. These voxels typically were located on edges of the brain or areas with typical signal loss, including fronto-orbital regions and anterior medial temporal regions. On average across subjects, 19% (std = 2.6%) of voxels were removed. Data was z -scored over time. All analyses were conducted in volume space using custom MATLAB scripts and then visualized using FSLview.

2.1.6.2 Intersubject correlation (ISC) between similarity groups

To test the hypothesis that subjects who share interpretations of the Movie show greater neural similarity, we compared the intersubject correlation (ISC) between groups of subjects who differed in how similarly they interpreted the Movie. Similarity was defined as the mean LSA similarity of each subject to every other subject (Figure 2.2A). The resulting “high interpretation similarity” group consisted of the 18 subjects with the most similar recalls to each other, while the “low interpretation similarity group” consisted of the 18 subjects with the least similar recalls. We then separately calculated ISC for each sub-group by correlating each subject’s response timecourse to the average of others in the same subgroup (N subjects in subgroup $- 1$) in the same voxel. The mean of the resulting N correlations is taken as ISC (Hasson et al., 2008; Nastase et al., 2019a). Finally, to identify areas that showed greater neural similarity among subjects who agreed on the interpretation compared to subjects with different interpretations, we directly compared ISC between the two sub-groups using t -tests (two-tailed, $\alpha = .05$) in every voxel with significant ISC, as calculated over all subjects ($N = 36$) (Regev et al., 2013).

Statistical significance of ISC was assessed using a permutation test. Each voxel’s timecourse was phase-scrambled by taking the Fast Fourier Transform of the signal, randomizing the phase of each Fourier component, and then inverting the Fourier transformation. This randomization procedure thus scrambles the phase of the signal, leaving its power spectrum intact. Using the phase-scrambled surrogate dataset, the ISC was again calculated for all voxels as described above, creating a null distribution of average correlation values for each voxel. This bootstrapping procedure was repeated

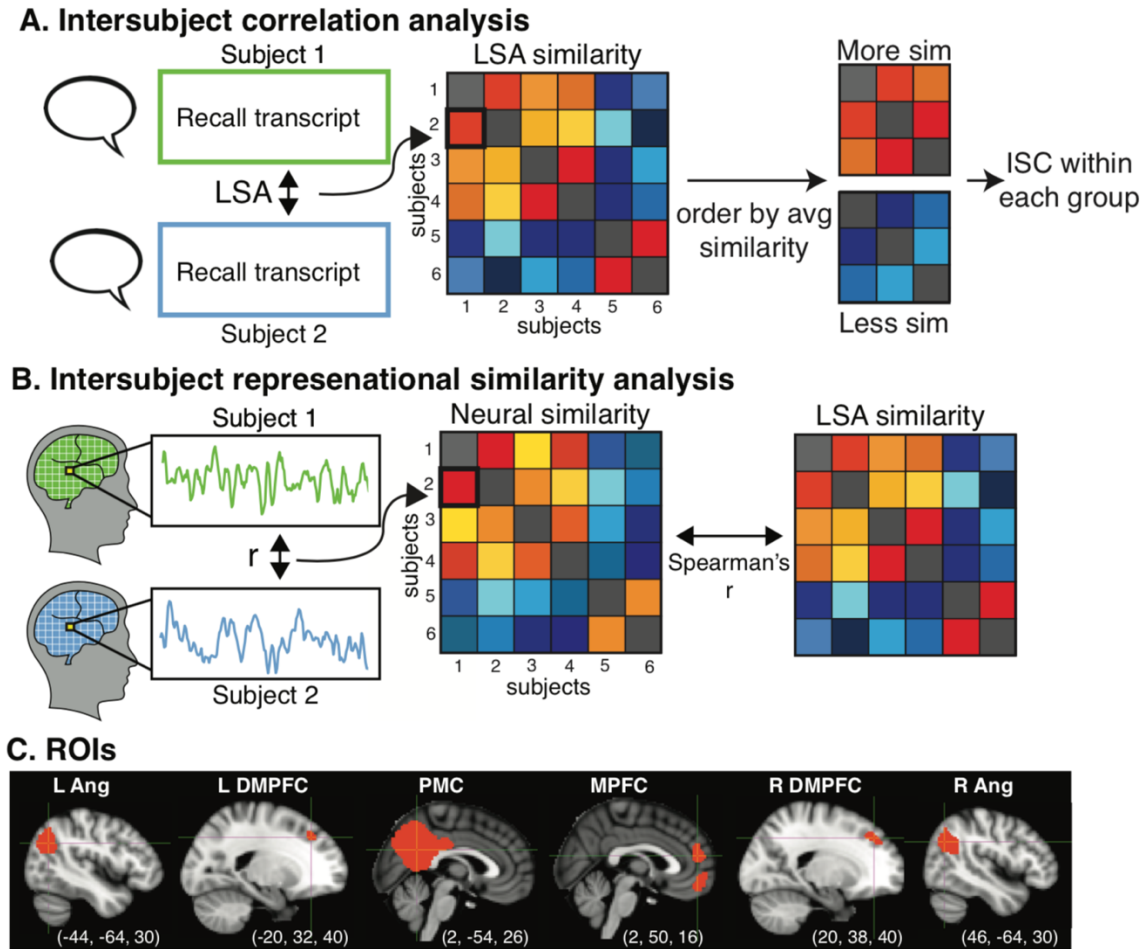


Figure 2.2 ISC and intersubject RSA procedure. **(A)** ISC analysis. Interpretation similarity between every pair of subjects was measured using Latent Semantic Analysis (LSA). Subjects were divided into equal sized groups based on average similarity to each other. Intersubject correlation (ISC) was then compared between subjects who had similar recalls and subjects who did not. **(B)** RSA. An intersubject representational similarity analysis (RSA) was conducted to identify regions of the brain where greater similarity in narrative interpretation was correlated with greater neural similarity. Neural similarity was measured by correlating each subject's response timecourse with every other subject's timecourse in every voxel. The resulting neural similarity matrix was correlated with the recall similarity matrix in all voxels. **(C)** Default mode network ROIs. ISC = intersubject correlation, LSA = latent semantic analysis, RSA = representational similarity analysis, ang = angular gyrus, DMPFC = dorsomedial prefrontal cortex, PMC = posterior medial cortex, MPFC = medial prefrontal cortex.

1000 times, producing 1000 bootstrapped correlation maps (Honey et al., 2012b; Regev et al., 2013).

To correct for multiple comparisons, the largest ISC value across the brain for each bootstrap was selected, following Regev et al. (2013). This procedure results in a null distribution of the maximum noise correlation and represents the chance level of calculating high correlation values across voxels in each bootstrap. The family-wise error rate of the measured maps was then controlled at $q = .05$ by selecting a correlation threshold (R^*) such that only 5% of the null distribution of maximum correlation values exceeded R^* . In other words, only voxels with mean correlation value (R) above the threshold derived from the boot-strapping procedure (R^*) were considered significant after correction for multiple-comparisons.

2.1.6.3 Intersubject representational similarity analysis (RSA)

To further test these findings using a more fine-grained analysis, we then identified regions of the brain where greater recall similarity between pairs of subjects predicts greater neural similarity. We conducted a voxel-wise intersubject representation similarity analysis (RSA) (Kriegeskorte et al., 2006, 2008) between LSA recall similarity and intersubject neural correlations (Figure 2.2B). Intersubject RSA follows the same logic of the classic RSA as described by Kriegeskorte et al. (2008): in RSA, a similarity matrix is built by comparing the similarity of every pair of experimental conditions (e.g. different categories of images). This similarity matrix is then compared with a neural similarity matrix, built by comparing the similarity of patterns of neural activations between each pair of experimental conditions. Larger correlations between the two

matrices in a given brain region suggest that it represents the information in the experimental condition matrix.

In the present work, rather than building similarity matrices that compare different experimental conditions, we build similarity matrices that compare similarity in interpretation of the movie across different subjects. We first constructed a recall similarity matrix using LSA cosine similarity (see Section 2.5), which represents how similarly each pair of subjects interpreted the stimulus. For each gray matter voxel in the brain, we constructed a neural similarity matrix by correlating each subject's response timecourse with every other subject's response timecourse in the same voxel. We then calculated Spearman's r between the matrix of neural similarity and the matrix of recall similarity. This analysis therefore identifies regions where subjects who have more similar interpretation also have more similar neural responses, suggesting that these areas represent idiosyncratic representations of the stimulus. Because the neural and recalls similarity matrices are symmetrical, only the lower triangles are correlated.

Following Kriegeskorte et al. (2008), statistical significance for RSA was assessed using a permutation test. For each voxel, the rows and columns of the neural similarity matrix were randomly shuffled, and the resulting shuffled matrix was correlated with the LSA similarity matrix as described above. This shuffling procedure was repeated 1000 times, resulting in a null distribution of 1000 values for the null hypothesis that there is no relationship between recall similarity and neural similarity. Following Chen et al (2016b) and Baldassano et al. (2017), the mean and standard deviation of the null distributions were used to fit a normal distribution and calculate p -

values. We corrected for multiple comparisons by controlling the False Discovery Rate (FDR) (Benjamini and Hochberg, 1995) of the RSA map using q criterion = 0.05.

2.1.6.4 ROI analysis

Because our hypotheses were focused on the DMN, we also conducted the ISC and RSA analyses on independently-defined DMN ROIs (Figure 2.2C). These ROIs were defined using functional connectivity on previously published, independent data (Chen et al., 2016b) that consists of subjects who were scanned in fMRI watching a movie. A seed ROI for posterior medial cortex was taken from a resting state-state connectivity atlas (posterior medial cluster functional ROI in “dorsal DMN” set) (Shirer et al., 2012). Following Chen et al. (2016b), the DMN ROIs were then defined by correlating the average response in the PMC ROI to every other voxel in the brain during the movie for each subject, averaging the resulting connectivity map, and thresholding at $R^* = 0.5$. Although the DMN is typically defined using resting-state data, recent work has shown that the same network is activated during temporally extended stimuli (Simony et al., 2016).

2.1.6.4 Control analyses accounting for head motion

To verify that these findings were not due to correlated subject motion, we extracted a timecourse of each subject’s estimated framewise displacement and then correlated the motion timecourses of every pair of subjects. We correlated each pair of subjects’ motion correlations with similarity of interpretation, as well as used t -tests to compare the level of correlated motion between high- and low-similarity groups. Finally, we regressed out

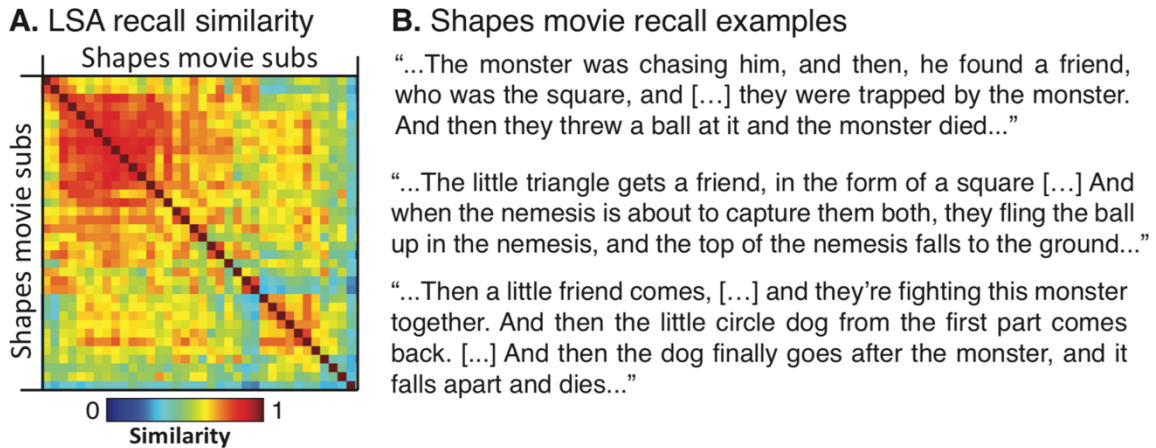


Figure 2.3 LSA results among shape movie recalls. **(A)** Recall similarity between each pair of subjects in each group was assessed using Latent Semantic Analysis (LSA). Movie subjects were varied substantially in their interpretation of the animation. **(B)** Example Shapes Movie recall excerpts.

each subject’s motion parameters (3 translation, 3 rotation) from their neural response.

The residuals from this regression were then used to repeat the above neural analyses.

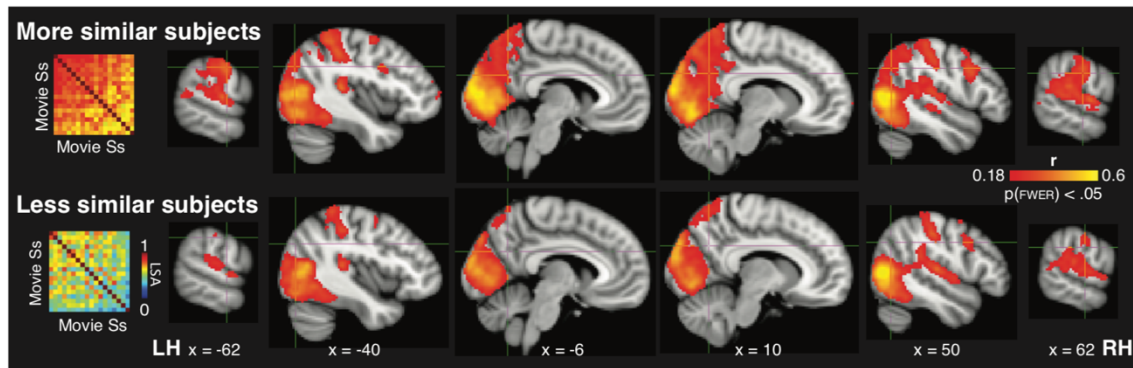
2.21 Results

We compared the behavioral and neural responses among subjects watching an ambiguous animation to identify regions of the brain that represent shared understanding of narratives.

2.2.1 Behavioral results

The average spoken recall was 177.6 sec (std = 86.6 sec) and described on average 14.4 events (std = 4.8), as rated by an independent coder. However, there was significant variance in the interpretation of the movie as measured by LSA (Landauer et al., 1998).

A. ISC in similar vs dissimilar recall groups



B. ISC is larger among subjects with similar recalls

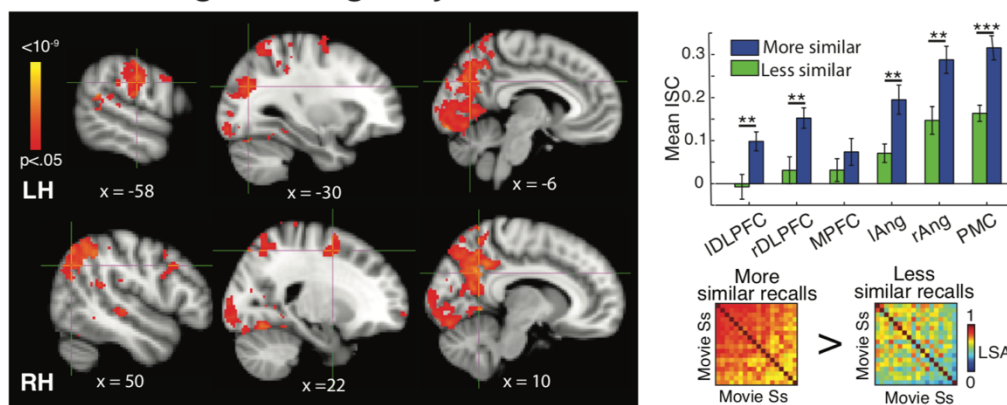


Figure 2.4 ISC among movie subjects with similar versus dissimilar recalls. **(A)** Movie subjects with more similar recalls (top row) showed widespread neural synchrony extending from early sensory areas to default mode network areas. Movie subjects with less similar recalls (bottom row) showed less widespread neural synchrony. **(B)** Movie subjects with more similar recalls showed significantly greater ISC throughout V1+, PMC, right angular gyrus, left supramarginal gyrus, and bilateral superior frontal gyrus compared to subjects who did not have similar recalls. In addition, the subjects with similar recalls had significantly greater ISC in all DMN ROIs except MPFC. DMPFC = dorsomedial prefrontal cortex, SMG = supramarginal gyrus, ang = angular gyrus, aSTS = anterior superior temporal sulcus, PMC = posterior medial cortex, dmPFC = dorsomedial prefrontal cortex. * $p < .05$, ** $p < .01$, *** $p < .001$. Error bars are SEM.

For example, subjects differed in their interpretation of a small circle as an inanimate object (e.g. a ball pushed around by the triangle child) or an animate character (e.g. a dog or pet or other animate being). This spread in behavioral outcomes was ecologically derived, meaning it was not prompted. Each subject freely came to their interpretation on

his or her own (for examples, see Figure 2.3B). This enabled us to separate subjects based on differing levels of similarity when interpreting the exact same stimulus.

Based on the average recall similarity to each other, the subjects were split into two equal sized groups ($N = 18$ each group): a “high similarity” group (mean cosine similarity = 0.745, std = 0.071, range = 0.58 – 0.89) and a “low similarity” group (mean cosine similarity = 0.528, std = 0.123, range = .26 – .81). There were no demographic differences between the two groups (age: $t(34) = .11, p > .05$; gender: $\chi^2(1) = 1.87, p > .05$; race: $p > .05$ all races, max $\chi^2 = .36$).

2.2.2 Neural results

2.2.2.1 ISC: greater shared response among subjects with shared interpretations

To test that groups of subjects with more similar interpretations of the movie showed greater neural similarity than subjects with differing interpretations, we measured neural similarity among subjects with similar versus dissimilar recalls using ISC, and then tested for differences in ISC between the two groups using t -tests.

The 18 subjects with the most similar recalls to each other showed significantly correlated neural responses throughout early visual areas, including much of occipital cortex, and auditory areas including A1+, superior temporal gyrus (STG), and middle temporal gyrus (MTG). Significant ISC was also observed in high-level regions including bilateral angular gyrus, bilateral supramarginal gyrus (SMG), bilateral inferior frontal gyrus (IFG), PMC, and anterior paracingulate cortex (PCC) ($p < .05$, FWER corrected; Figure 2.4A, top row). The 18 “low recall similarity” subjects also showed significant ISC in sensory regions, extending into linguistic areas of the superior temporal lobe

Intersubject RSA: Similarity in recall is correlated with neural similarity

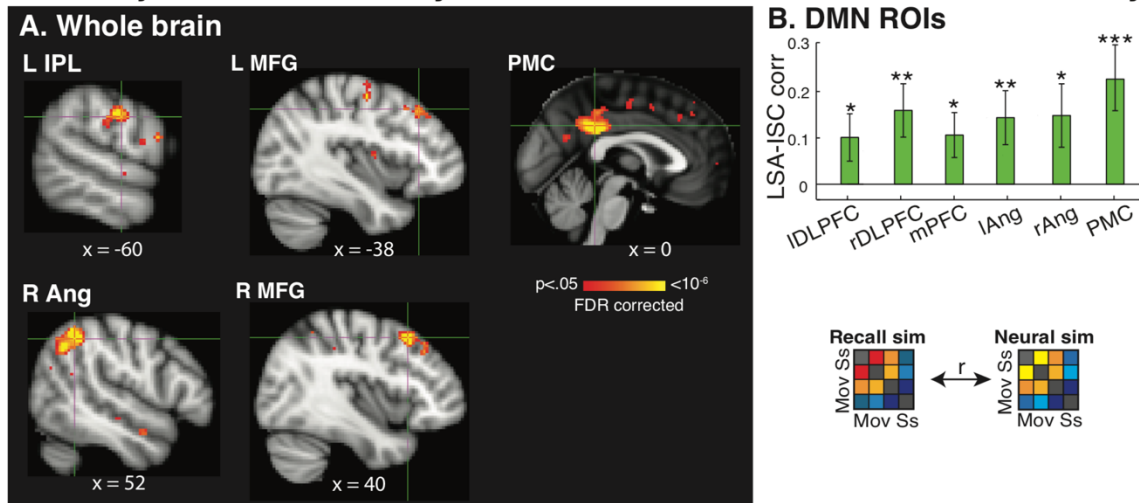


Figure 2.5 Intersubject RSA among movie subjects. **(A)** Whole brain analysis: among subjects who watched the shapes animation, neural similarity and recall similarity were significantly correlated with each other in posterior medial cortex (PMC), bilateral middle frontal gyrus (MFG), right angular gyrus, right superior temporal sulcus (STS), and left inferior parietal lobule (IPL) ($q < .05$, FDR corrected). **(B)** ROI analysis: greater recall similarity was significantly correlated with neural similarity in all DMN ROIs. DLFC = dorsolateral prefrontal cortex, mPFC = medial prefrontal cortex, ang = angular gyrus, PMC = posterior medial cortex, * $p < .05$, ** $p < .01$, *** $p < .001$. Error bars are SEM.

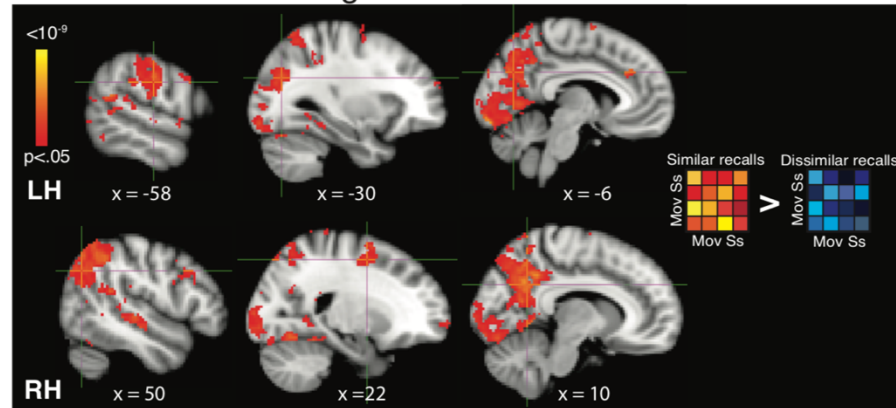
(Figure 2.4A, bottom row). However, a t-test contrasting ISC maps between the two groups revealed significantly greater ISC among the “similar” subjects throughout V1+, PMC, right angular gyrus, left SMG, and bilateral superior frontal gyrus (SFG) (Figure 2.4B, left). Moreover, the subjects with more similar recalls had significantly greater ISC in all DMN ROIs except mPFC ($p < .05$; Figure 2.4B, right) relative to the dissimilar group.

2.2.2.2 RSA: Interpretation similarity is correlated with neural similarity

To test the hypothesis that greater similarity in the interpretation of a narrative is reflected in greater neural similarity across subjects, we conducted an intersubject RSA over the

Shared responses are not due to correlated motion

A. ISC difference using motion residuals



B. Intersubject RSA using motion residuals

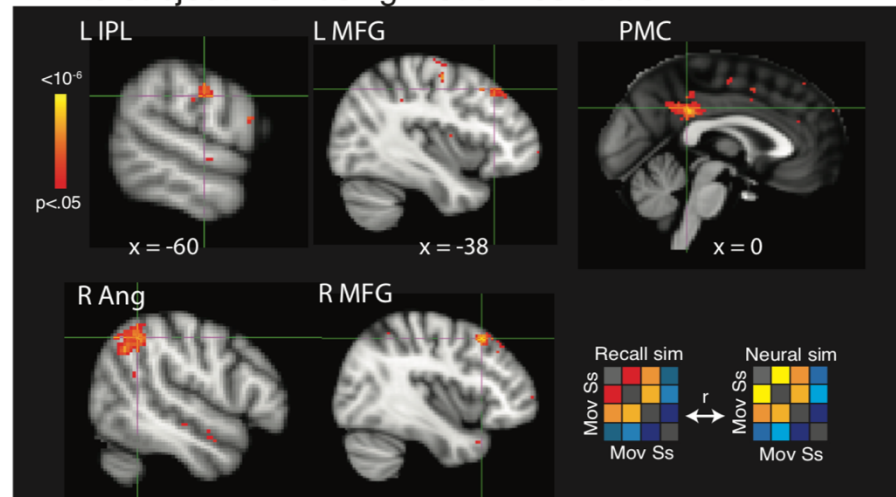


Figure 2.6. ISC and intersubject RSA on motion residuals. To confirm that our results were not confounded by correlated motion across subjects, we regressed out motion estimates from the neural responses of each subject and repeated the analyses on the residuals of the regression. We found largely the same results as in the original analysis for both ISC results (A) and intersubject RSA (B).

entire brain, directly comparing neural similarity with recall similarity in every gray matter voxel in the brain (Figure 2.2B, see Methods above for details). We found that the level of recall similarity was correlated with the level of neural similarity in PMC, right angular gyrus, right SMG, bilateral anterior STG, bilateral DMPFC, and bilateral DLPFC ($q < .05$, FDR corrected; Figure 2.5A). The same correlations were also

measured in six independently defined ROIs of the DMN (Figure 2.5B).

2.2.2.3 Shared responses cannot be explained by correlated motion

The mean frame-wise displacement during the Movie was 0.083 mm (std = 0.03), suggesting excellent participant compliance. Motion was not significantly correlated among most subjects (mean $r = 0.071$, std = 0.08), though a subset of subjects did show correlated motion (84 of 630 pairwise comparisons; all $ps < .05$, uncorrected for multiple comparisons). However, correlated motion did not vary systematically among subjects as a function of recall similarity: the correlation between motion correlations and recall similarity was $r = -0.049$ ($p > .05$). There was also no difference in the level of correlated motion between the “similar” and “dissimilar” groups used in the ISC analyses ($t(304) = 0.69$, $p > .05$). In addition, we regressed out motion estimates for each subject from their neural response and repeated the above whole-brain analyses on the residuals of this regression, identifying the same regions as above (Figure 2.6).

2.3 Discussion

Narratives form an important part of daily life, and interpreting unclear or ambiguous narratives is essential to navigating both the physical and social worlds.

This work is the first to identify high-level regions of the brain that can discriminate between idiosyncratic, spontaneous differences in the interpretation of an ambiguous narrative. We found that the more similarly two people interpreted the social events depicted in an ambiguous animation, the more similar their neural responses were

in a subset of DMN regions, including PMC, right angular gyrus, bilateral STG, and DMPFC. In addition, we observed this relationship in regions outside the DMN, including more anterior right SMG and bilateral DLPFC. These results are consistent with previous studies that directly manipulated interpretation by directing attention to different aspects of a spoken narrative (Cooper et al., 2011), changing perspective (Lahnakoski et al., 2014), or biasing with contextual information (Yeshurun et al., 2017b). Unlike these previous results, however, in the present work, we do not manipulate the interpretation of the animation into discrete experimental groups. Rather, we let subjects freely attribute intentions to the motion of simple geometric shapes, which spontaneously led to the creation of complex social narratives in subjects' minds as expressed in their post-viewing descriptions of the animation. We then show that the subtle individual differences in rich narrative interpretation are reflected in individual differences in the neural responses of high-level regions.

2.3.1 Processing in the default mode network

Many of the regions in which we identify a relationship between interpretation similarity and neural similarity overlap substantially with the DMN (Raichle et al., 2001; Buckner et al., 2008). While the DMN was originally conceptualized as a task-negative network showing decreased activity during externally-directed tasks (Raichle et al., 2001; Fox and Raichle, 2007; Buckner et al., 2008), later work observed robust DMN activity to a variety of tasks, including episodic memory (Svoboda et al., 2006; Spreng et al., 2009; Andrews-Hanna et al., 2014), working memory (Vatansever et al., 2015, 2017b), forecasting (Buckner et al., 2008; Spreng et al., 2009), semantic processing (Krieger-

Redwood et al., 2016; Vatansever et al., 2017a), and emotional processing (Barrett and Satpute, 2013). Notably, the DMN largely overlaps with the mentalizing network (Schilbach et al., 2008; Mars et al., 2012), regions of the brain that show increased activity while thinking about other minds. In measuring mentalizing in the brain, one common task contrasts shape animations that have animate motion (e.g. chasing, kicking) and non-animate motion (e.g. random motion). Greater activation was found during animate than inanimate movies in regions of DMN (Castelli et al., 2002; Vanderwal et al., 2008; Schurz et al., 2014). In contrast to these studies, we use a novel shapes animation that is unique for its length (7 mins vs the typical 10 – 30 seconds) and complex narrative arc, as well as the large number of interacting characters with different relationships (parent and child, friends, antagonists). We therefore show that these areas not only respond preferentially to animate films, but can discriminate between subtle differences in interpretations of dynamically occurring social interactions.

The DMN is also implicated in the processing of complex narratives. Across many studies, researchers have observed robust correlations among subjects in the DMN in response to the high-level features of narratives. These shared responses across subjects only occur in the DMN with temporally coherent narratives (Hasson et al., 2008; Lerner et al., 2011; Simony et al., 2016), are insensitive to low-level stimulus features such as modality or language (Honey et al., 2012b; Regev et al., 2013; Chen et al., 2016b; Zadbood et al., 2017), and are locked to narrative interpretation (Yeshurun et al., 2017a, 2017b). Based on these findings, we have previously suggested that the DMN is situated at the top of a timescale processing hierarchy, integrating high-level information over minutes or more (Hasson et al., 2008, 2015; Lerner et al., 2011). Other work has

additionally supported a role of the DMN as a high-level network that flexibly integrates information from many lower-level networks (Vatansever et al., 2015; Margulies et al., 2016; Vidaurre et al., 2017). The present findings are therefore consistent with DMN as a high-level global integrator, and additionally demonstrates that the temporal patterns within the DMN are sensitive to idiosyncratic differences in narrative interpretation.

2.3.2 Processing outside the DMN

We observed a significant relationship between recall similarity and neural similarity in regions outside of the DMN, including right posterior SMG and bilateral DLPFC centered in MFG. Both of these regions are part of the frontoparietal control network (FPCN), which is widely implicated in cognitive control and decision-making processes (Vincent et al., 2008). The FPCN is highly, albeit heterogeneously, interconnected with the DMN (Spreng et al., 2010, 2012). In particular, both SMG and DLPFC are part of a proposed FPCN subnetwork, FPCN-A, that has been shown to be more highly correlated with DMN in a variety of tasks ranging from classic cognitive control tasks, resting state, social-cognitive tasks, and movie watching (Dixon et al., 2018). Moreover, Dixon et al. found in a meta-analysis that FPCN-A is more involved in mentalizing, emotional processing, and complex social reasoning than other regions of the FPCN, which is consistent with a role for SMG and DLPFC in representing idiosyncratic interpretations of an animate shapes movie.

2.3.3 Limitations and alternate interpretations

In addition to high-level regions, we also found greater neural similarity in visual areas among subjects who shared interpretations than among subjects who had did not. This difference may arise from differences in fixation patterns as a function of interpretation, which would be consistent with a study showing differences in fixation patterns during the same audiovisual movie when subjects take different psychological perspectives. These researchers also observed significant differences in neural synchrony in early visual cortex as a result of different perspectives (Lahnakoski et al., 2014). However, future work should use eye-tracking during scanning in order to test this hypothesis.

Another possible contribution to neural differences among subjects may be due to different but systematic changes in arousal or attention in each group that track with the stimuli. For example, we observe significant correlations between neural and recall similarity in a subset of fronto-parietal control network regions, raising the possibility that cognitive control or attentional processes may contribute to our findings. To better address this possible confound, subsequent studies could take additional in-scanner physiological measures as well as utilize post-scan questionnaires to measure engagement or emotional arousal.

Finally, it is possible that differences among the high- and low-recall similarity groups may better reflect differences in memory and performance on the recall task, rather than differences in interpretation. Such effects of memory are difficult to control for in this particular experiment, as the ambiguity of the stimulus makes assessing absolute recall performance (rather than recall similarity to others) difficult. All the analyses in this study were performed during the encoding phase, and as such are related

to online processing and encoding of the narratives (for further discussion of the tight relationship between memory and online processing see Hasson et al. (2015)).

2.3.4 Intersubject RSA

Finally, this work introduces a novel analytic method, intersubject RSA, for measuring individual differences in neural responses using complex, naturalistic stimuli. This method and type of stimuli can provide important insights into the shared processing of complex social information across subjects that leads to the creation of a shared reality and facilitates social communication (Hasson et al., 2012; Hasson and Frith, 2016). Future applications of this approach could enable us to delineate the development of high-level social cognitive abilities in the DMN during childhood. This method may also enable the detection of abnormalities during complex naturalistic perception and narrative interpretation relevant to neurological disorders. For example, previous work has shown that individuals with autism spectrum disorder or schizophrenia show atypical interpretations of simpler shape-based animations (Castelli et al., 2002; Salter et al., 2008; Horan et al., 2009; Bell et al., 2010), but differences in interpretation under naturalistic conditions have not yet been linked to individual differences in neural responses. Finally, the intersubject RSA method may be used to explore neural representations of individual trait differences, such as differences in cultural background, creativity, mentalizing, or political view.

2.4 Conclusions

In conclusion, this work provides evidence that shared understanding is reflected in shared neural responses in regions of the default mode network during the processing of abstract, ambiguous narratives. The similarity of neural patterns elicited by similar interpretations of the same narrative demonstrates the remarkable sensitivity of the default mode network to subtle individual differences, as well as its strong role in mentalizing and social perception. This work invokes the role of the default mode network in representing idiosyncratic interpretation of complex social narratives.

Chapter 3

Shared default mode network responses across modality depend on shared interpretation

While a single narrative can be ambiguous in interpretation, the same narrative can also be communicated in forms as varied as words, sounds, images – even the motion of simple shapes, as seen in Chapter 2. How does the brain represent narratives across communication forms as varied as geometric shapes and spoken words, and how are these representations modulated by differences in interpretation?

Shape animations were first developed by Heider & Simmel in 1944. Their first animation was a 2.5-min movie that consisted of two triangles and a circle moving around a screen. Although the perceptual information consisted only of simple geometric shapes and their movements, the majority of subjects interpreted the geometric shapes as people, attributed personalities and mental states to them, and described their interactions in the form of a simple narrative. The Heider & Simmel results have since been extensively replicated using the original and novel animations (Oatley and Yuill, 1985; Berry et al., 1992; Scholl and Tremoulet, 2000) and is now used as a common task for measuring theory of mind (Gobbini et al., 2007; Schurz et al., 2014). Researchers have found children attribute animacy to the shapes from an early age (Berry and Springer,

1993; Csibra et al., 1999; Surian and Geraci, 2012) and that similar behavior is observed across cultures (Rimé et al., 1985). Indeed, failure to anthropomorphize shapes in these movies is associated with atypical theory of mind in schizophrenia (Horan et al., 2009; Bell et al., 2010) and autism (Abell et al., 2000; Castelli et al., 2002; Salter et al., 2008; Kana et al., 2009).

More recently, researchers have used Heider & Simmel-like shape movies to identify areas of the brain involved in theory of mind and mentalizing. Neuroimaging work has shown that animate shape movies engaged areas of the brain involved in mentalizing processes more than non-animate shape movies. These areas including bilateral temporoparietal junction (TPJ)/angular gyrus, superior temporal sulcus (STS), posterior medial cortex (PMC), medial and dorsal prefrontal cortex (PFC), fusiform gyrus, and temporal poles (Castelli et al., 2002; Gobbini et al., 2007; Schurz et al., 2014), areas which largely overlap with the mentalizing network (Frith and Frith, 2006; Mar, 2011; Mars et al., 2012) and the default mode network (DMN) (Raichle et al., 2001; Buckner et al., 2008).

Implicit in these studies is that people interpret, and therefore neurally represent, the moving shapes as animate characters or people. Previous work, however, has not directly tested to what extent moving shape animations are processed and represented in the same way as narratives communicated in more typical forms, such as spoken stories. Research has shown shared neural representations to narratives across modalities: written versus spoken word (Regev et al., 2013), live-action, audiovisual movie versus spoken word (Baldassano et al., 2017; Zadbood et al., 2017), and Russian versus English language (Honey et al., 2012a) in the DMN, but not across such divergent and abstract

stimuli. Moreover, existing research has not investigated how differences in interpretation modulates shared neural responses across modality.

Here, we will test the hypotheses that (1) the same narrative communicated via an animate shape movie or via spoken words will elicit shared neural representations in regions of the DMN, and that (2) similarity in shared responses across modality will depend on similarity in interpretation. To test these hypotheses, we used previously scanned data from subjects watching an animate shapes movie (Chapter 2) and additionally scanned two new groups of subjects. One group listened to an audio version of the shapes movie that described the shapes as animate characters interacting with each other. The second group listened to an audio version of the movie that described the shapes and their movements in purely physical terms. We directly compared the response timecourse across the Movie and the two Audios to identify areas of the brain that represented the shared content of the stimuli across forms. We then tested whether greater similarity in narrative interpretation, corresponding to greater shared understanding, was correlated with greater similarity in neural response between conditions. Our results demonstrate that regions of the DMN are at once sensitive to idiosyncratic differences in narrative interpretation and resilient to vast differences in the form of communication.

3.1 Methods

3.1.1 Subjects

Previously collected data from the 36 subjects who participated in the Movie experiment (see Chapter 2 for details) were included in the analysis here. In addition, two groups of

subjects with normal hearing and normal or corrected-to-normal vision were scanned for the Audio experiments. Eighteen subjects (ages 18 – 32, mean = 22.2 years; 4 male) were included in the Theory of Mind (ToM) Audio Group and 17 (ages 18 – 33, mean = 22.2 years; 7 male) in the Physical Description (Phys) Audio group. Two subjects scanned for the ToM group were removed from analysis due to excessive motion ($>3\text{mm}$), and one subject from the Phys groups due to an anomalous finding. Sample sizes were selected based on power analyses of intersubject analyses (Pajula and Tohka, 2016). All experimental procedures were approved by the Princeton University Internal Review Board, and all subjects provided informed, written consent.

3.1.2 Stimuli and experimental design

Stimuli and experimental design for the Movie group were previously described (Chapter 2.1.2). The two new audio groups were scanned in fMRI while listening to two different verbal descriptions of the Movie animation.

The “Theory of Mind (ToM) Audio” group listened to a description of the animation’s story, narrating the movement of the shapes as social characters (Figure 3.1, middle row; see Appendix 2 for full text). The ToM Audio story was based on the director’s interpretation of the animation and had no visual component. In the story, a small child lives in a simple house with a parent. On the first night, the child has a dream that he is flying in the sky with a flock of birds. On the second night, the child returns to the flying dream but the birds change into a monster that chases the child. On the third night, the child is joined by a new friend in the dream, and together they defeat the monster. The next day, the two friends meet in real life at the playground.

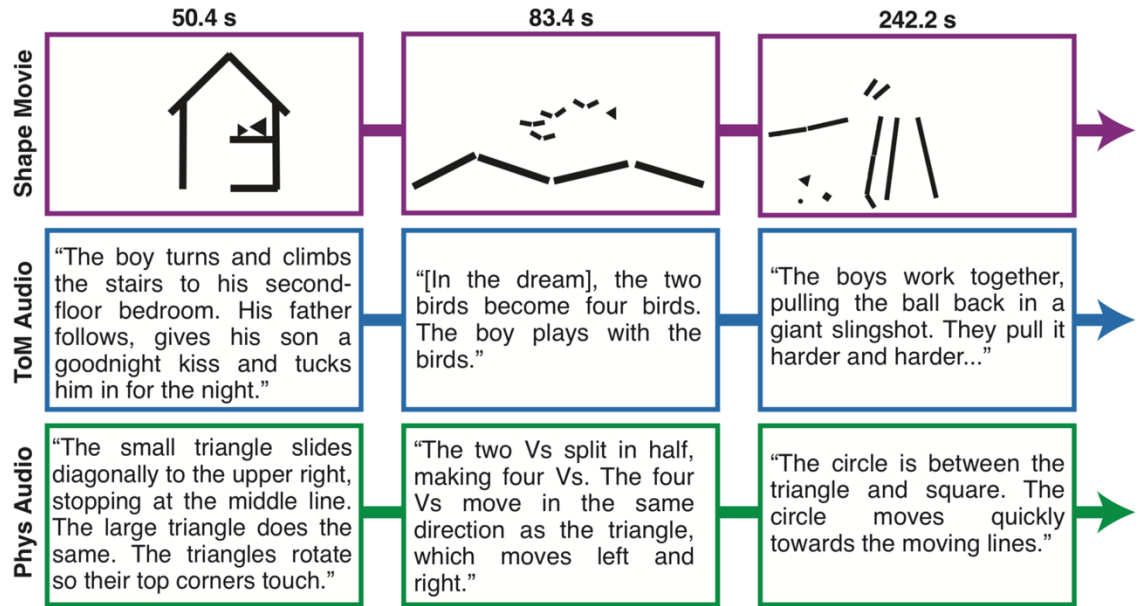


Figure 3.1 Theory of mind and physical audio stimuli. Subjects were scanned in fMRI watching a short movie consisting of moving shapes (Shape Movie, top row, $N = 36$), listening to an Theory of Mind Audio version of the movie narrative (ToM Audio, middle, $N = 18$), or listening to a physical description of the movements of the shapes in the movie (Phys Audio, bottom, $N = 17$). The two Audios were edited to be time-locked with the Shape Movie such that each segmented event began at the same time.

The “Physical (Phys) Audio” group listened to a 7-min audio that described the physical characteristics and motion of the shapes in the Movie without attributing animacy to them (Figure 3.1, bottom row; see Appendix 3 for full text). For example, a scene that is described in the ToM Audio as a child playing with a ball before being picked up by their parent is described in the Phys Audio as, “The triangle moves in a ring around the circle, rotating. A large triangle slides towards the rectangle. The small triangle slides outside the rectangle until it is adjacent to the large triangle. After bouncing, both triangles slide out of the scene.”

The two Audio versions described different aspects of the original Movie animation, but there were slight differences in the timing of different events across the

visual and auditory conditions. Thus, in order to time-lock the two stimuli, the Movie animation was segmented into 76 short events or actions (e.g. child bounces balls, parent kisses child goodnight, child and friend make a plan). The raw Audio stories were then edited to match the onset of each event in the animation, following Regev et al. (2013) and Honey et al. (2012b) (Figure 3.1). To remove transient, non-selective responses that occur at the onset of a stimulus, all scans were preceded by the same, unrelated 37-second movie clip. This clip was cropped from all analyses.

As with the Movie subjects, following stimulus presentation, subjects freely recalled the stimulus using their own words and in as much detail as possible using a customized MR-compatible recording system.

3.1.3 Stimulus presentation

Stimuli were presented using MATLAB and the same setup as previous described in Chapter 2.1.3.

3.1.4 MRI acquisition

Subjects were scanned in fMRI at the Princeton Neuroscience Institute Scully Center for Neuroimaging using the same sequences as previously described in Chapter 2.1.4.

3.1.5 Behavioral data analysis

Recall similarity among subjects in the Audio groups and between Audio and Movie Subjects were analyzed using Latent Semantic Analysis (Landauer et al., 1998), as previously described in Chapter 2.1.5.

3.1.6 MRI data analysis

3.1.6.1 Preprocessing

MRI data were preprocessed using FSL 5.0 (FMRIB, Oxford), as previously described in Chapter 2.1.6.1.

3.1.6.2 Audio correlations between stimuli

Because the Movie and Audios were temporally aligned such that the start of each event occurs at the same timepoint, the audio envelopes (audio amplitudes) may be correlated across stimuli. Following Honey et al. (2012b), we projected out the audio envelope from each subject's neural response in each condition. The audio envelope for each stimulus was calculated using a Hilbert transform and then down-sampled to the 1.5-second TR using an anti-aliasing, low-pass finite impulse response filter. The resulting envelopes were then convolved with a hemodynamic response function (Glover, 1999). The envelopes were entered into a linear regression model for each voxel in each subject in the corresponding condition. For all subsequent analyses, the BOLD response timecourse was then replaced with residuals of the regression.

3.1.6.3 Intersubject correlation (ISC) across modalities

To test the hypothesis that shared interpretation of a narrative elicits shared neural responses across narrative modality, we calculated intersubject correlations (ISC) between Movie subjects and ToM and Phys Audio subjects. ISC was calculated as in Chapter 2.1.6.3 but between Movie and Audio groups. ISC was first separately calculated between all Movie subjects and each Audio group: the time series of each Movie subject was correlated with the average timeseries of all ToM Audio subjects or all Phys Audio subjects in the same voxel in the brain. The average of these correlations is then taken as ISC.

The Movie group was subsequently divided into two equal groups based on their average similarity to each Audio. For the cross-modal analysis between the Movie subjects and the ToM Audio subjects, the Movie subjects were divided based on similarity in interpretation to that of subjects in the ToM Audio group. This resulted in a “high interpretation similarity to ToM Audio story” group and a “low interpretation similarity to the ToM Audio story” group. ISC was then separately calculated between the ToM Audio subjects and each of the Movie subgroups. The same analysis was then repeated, but between the Movie subjects and the Phys Audio subjects.

We corrected for multiple comparisons across voxels using FDR (Benjamini & Hochberg, 1995) at $q = .05$.

3.1.6.4 Intersubject representational similarity analysis (RSA)

Using the same procedure as described in Chapter 2.1.6.3, we then conducted the voxel-

wise intersubject representational similarity analysis (RSA) (Kriegeskorte et al., 2006, 2008) between the Movie group and both the ToM Audio group and the Phys Audio group.

3.1.6.5 ROI analysis

ISC and intersubject RSA were also conducted on independently defined ROIs of the default mode network (DMN). ROI definition is described in detail in Chapter 2.1.6.4.

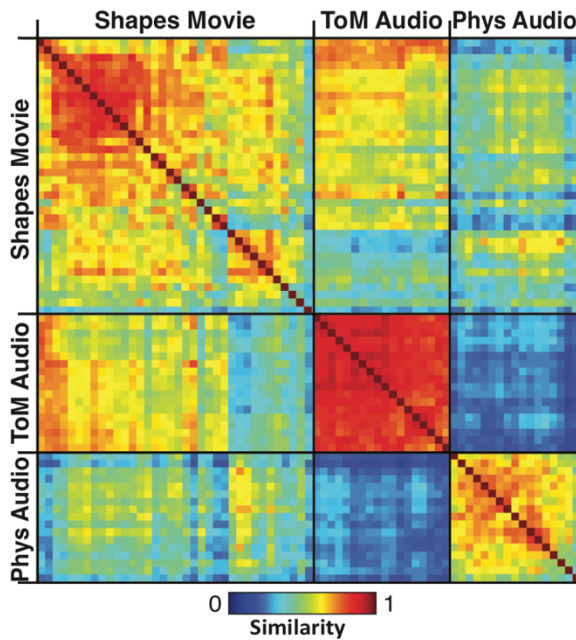
3.2 Results

We compared the behavioral and neural responses between three different groups to identify areas of the brain that represent shared understanding of ambiguous narratives across modalities. The “Movie” group was previously scanned in fMRI while they watched a novel 7-min ambiguous animation that told a complex social narrative using only the movement of simple geometric shapes. The “ToM Audio” group was scanned while listening to a verbal description of the social interactions in the movie, while the “Phys Audio” group was scanned listening to a physical description of the shapes in the movie.

3.2.1 Behavioral results

Subject recalls of the Movie, ToM Audio, and Phys Audio varied substantially in length and content (for examples of recalls, see Figure 3.2B and C). In the Movie group, the average spoken recall was 177.6 sec (std = 86.6 sec) and described on average 14.4

A. LSA recall similarity



B. ToM Audio recall examples

"There's a boy .. playing with a ball. It kicks it, and then tries to see where else he wants to kick it and so goes around the ball. He goes to the fence and kicks the ball to himself. He goes over to his dad, says a couple words to his dad, and they go home."

"He had a ball and he was playing around with the ball. And he ended up, kicking the ball at the fence, almost like soccer like. And, his father comes and wants him to come inside."

C. Phys Audio recall examples

"They bounced around a little bit. They, the circle was bouncing in between the edges of the rectangle and the triangle, the small triangle. Eventually a large triangle enters the picture and the smaller triangle moves out of the rectangular object."

"The triangle slid over the circle to enter to the rectangle. They did a little bouncing around. And the triangle circled around the circle a couple times. And, then at one point, a large triangle entered the scene. And the circle sort of disappeared."

Figure 3.2 LSA recall similarity results for movie and audio recalls. **(A)** While both Movie and Phys audio subjects showed similar variance in recalls ($t(764) = 1.63, p > .05$), the ToM Audio recalls were significantly more similar to one another than the Movie recalls were ($t(781) = 22.7, p < 10^{-6}$). Across stimuli, Movie recalls were however significantly more similar to the ToM Audio recalls than the Phys Audio recalls ($t(1258) = 17.0, p < 10^{-6}$). **(B)** and **(C)** example recall excerpts from subjects in ToM Audio and the Phys Audio conditions respectively describing the same scene.

events (std = 4.8), as rated by an independent coder. In the ToM Audio group, the average recall was 221.3 sec (std = 68.8 sec) and described an average 18.5 events (std = 3.1). In the Phys Audio group, the average recall was 145.5 sec (std = 88.5 sec) and described an average 6.4 events (std = 2.3). Recall length did not differ between the Movie and ToM Audio subjects ($t(52) = 1.82, p = .073$) or the Movie and Phys Audio subjects ($t(51) = 1.2, p = .235$). Recalls of the ToM audio group were significantly longer than the recalls of Phys Audio group ($t(31) = 2.87, p = .0071$). The number of recalled events also differed among groups: ToM Audio subjects recalled more events than Movie

subjects ($t(52) = 3.19, p = .0024$) or Phys Audio subjects ($t(33) = 12.68, p < 10^{-6}$), while Movie subjects recalled more than Phys Audio subjects ($t(51) = 6.27, p < 10^{-6}$).

The LSA (Landauer et al., 1998) results were as expected (Figure 3.2A): The mean LSA similarity within the Movie Group was 0.619 (std = 0.125, range = 0.26 – 0.89), the ToM Audio group 0.852 (std = 0.045, range = 0.70 – 0.93), and the Phys Audio group 0.638 (std = 0.94, range = 0.36 – 0.84). The recalls in the Movie group were significantly less similar to one another than the recalls in the ToM Audio group ($t(781) = 22.7, p < .001$). There was no difference in similarity within the Movie group and the Phys Audio group ($t(764) = 1.63, p > .05$). Between groups, the average LSA similarity between the Movie group and the ToM Audio group was 0.545 (std = 0.125, range = 0.26 – 0.8). Between the Movie group and the Phys Audio group, the average LSA similarity was 0.434 (std = 0.105, range = 0.18 – 0.72). The Movie recalls were significantly more similar to those in the ToM Audio group than those in the Phys Audio group ($t(1258) = 17.0, p < .001$).

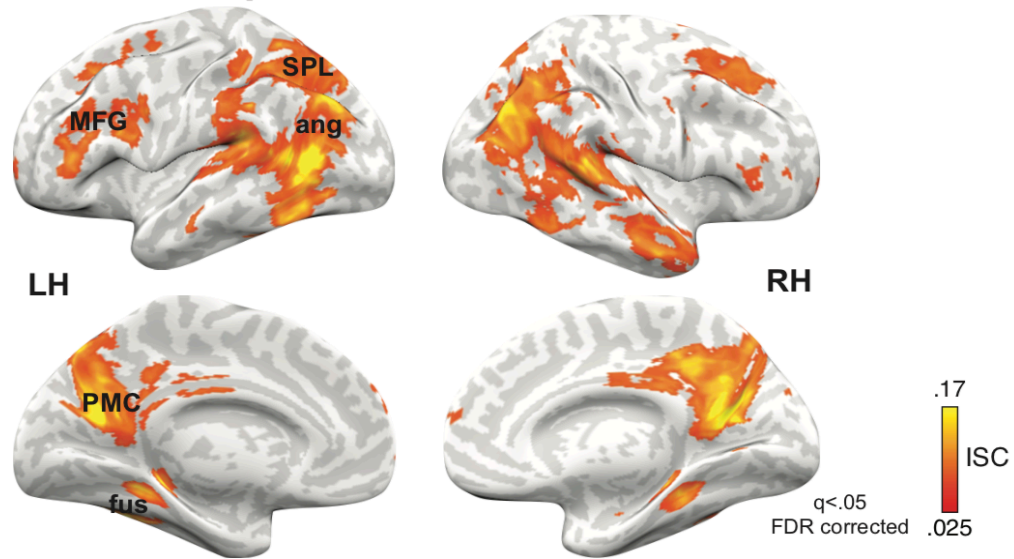
3.2.2 Neural results

3.2.2.1 ISC: Shared neural response across modality

We first conducted a cross-modal ISC analysis that correlated each Movie subject's neural response with the average Audio subject's to identify regions of the brain that process the narrative similarly, regardless of modality.

Movie and ToM Audio. We observe widespread ISC between Movie and ToM Audio subjects predominantly in regions of the DMN ($q < .05$, FDR corrected; Figure 3.3A), including PMC, bilateral angular gyrus, and bilateral MFG. There is additionally

A. ISC: Shared response between Movie and ToM Audio



B. ISC: Shared response between Movie and Phys Audio

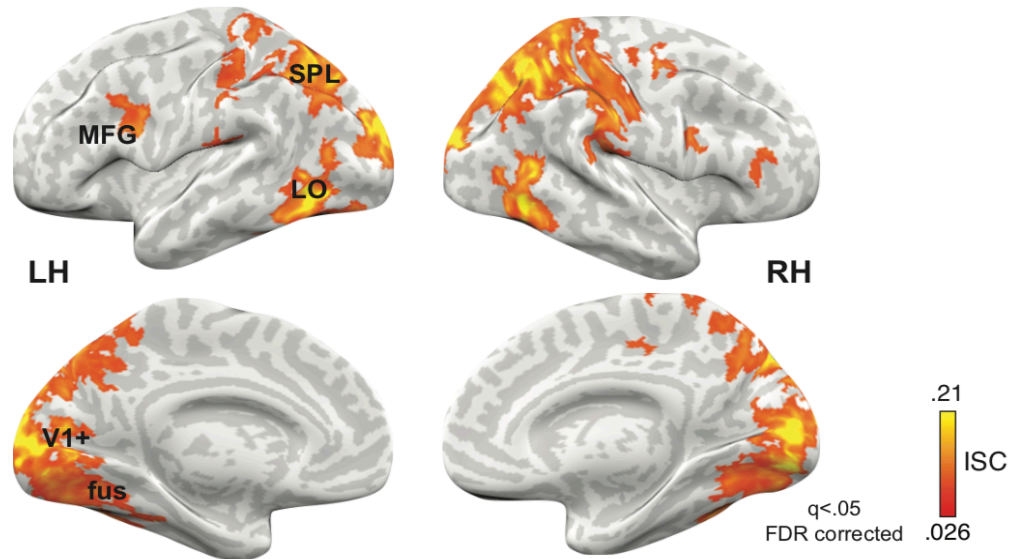


Figure 3.3 Cross-modal ISC between Shapes Movie and Audios. **(A)** Despite vast differences in the low-level stimulus properties of the Shapes Movie and the ToM Audio, the two stimuli elicited significantly correlated neural responses (ISC) across groups in several DMN regions including PMC, angular gyrus, MFG, and IFG. **(B)** In contrast, the Shapes Movie and Phys Audio showed significant ISC predominantly in visual areas, including V1+, VTC, and LO, as well as SPL. PMC = posterior medial cortex, ang =

significant ISC in temporal linguistic regions, including bilateral posterior STG, ITG and right temporal pole, as well as in bilateral SPL and bilateral middle fusiform gyrus.

Movie and Phys Audio. Movie-Phys Audio ISC is primarily observed in occipital visual cortex, including V1+ extending ventrally to fusiform gyrus and dorsally to lateral occipital complex (LOC) ($q < .05$, FDR corrected; Figure 3.3B). We also observe significant Movie-Phys Audio ISC in bilateral SPL, dorsal precuneous, and a small cluster of MFG, partially overlapping the Movie-ToM Audio ISC map.

3.2.2.2 ISC: Greater shared response among subjects with shared interpretations

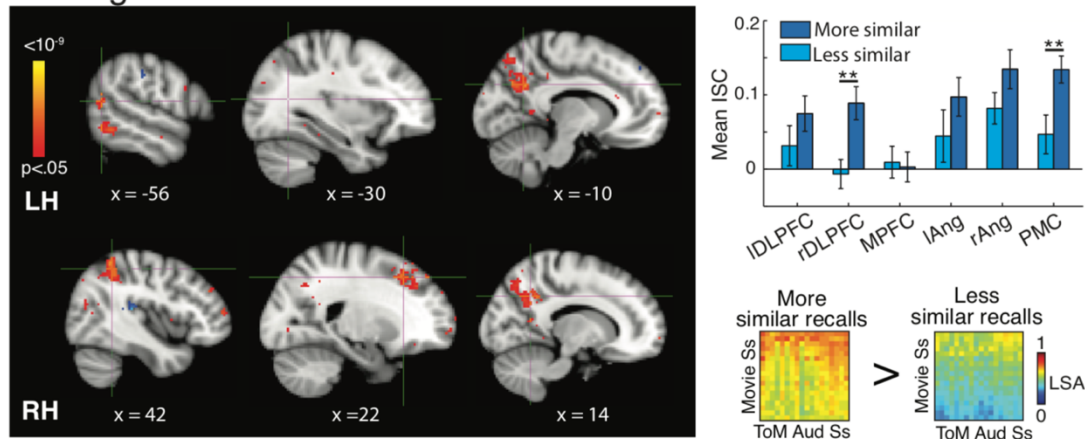
We next tested whether the extent of shared neural responses across modalities is contingent on shared interpretation of the stimulus. We therefore split Movie subjects into subgroups based on how similar their recalls were to each of the Audio groups and calculated ISC.

Movie and ToM Audio. Movie subjects who had similar interpretations to the ToM Audio subjects showed greater similarity in neural responses to the ToM Audio subjects in high-level brain regions than did Movie subjects with dissimilar interpretations to the ToM Audio subjects. This difference was observed in bilateral angular gyrus, PMC, and left MTG (Figure 3.4A). In DMN ROI analyses, the 18 subjects with more similar recalls, across modalities, had greater ISC in right dorsolateral prefrontal cortex (DLPFC) and PMC ($p < .05$; Fig 3.4A, upper right).

Movie and Phys Audio subjects. The Similar Movie-Phys ISC was greater than the Dissimilar Movie-Phys ISC in early visual areas. There were no significant differences between groups in ROI analyses (Figure 3.4B).

Difference in ISC between similar vs dissimilar groups

A. Larger ISC between Movie and ToM Audio subs with similar recalls



B. Larger ISC between Movie and Phys Audio subs with similar recalls

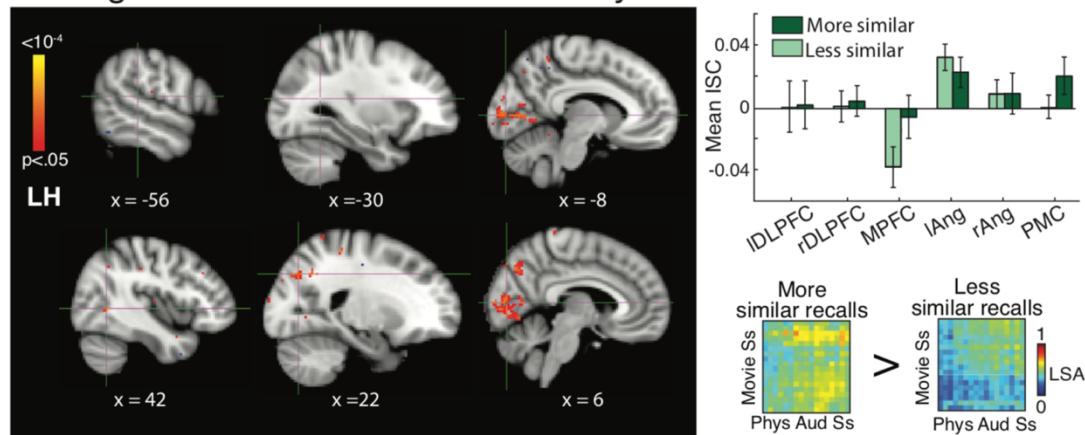


Figure 3.4 Difference in cross-modal ISC between similar and dissimilar recall groups **(A)** Across modalities, Movie and ToM Audio subjects with more similar recalls showed significantly greater ISC in bilateral angular gyrus, PMC, and left middle temporal gyrus. This pattern of results was also observed in two DMN ROIs, rDLPFC and PMC. **(B)** Movie and Phys Audio subjects with more similar recalls showed greater ISC in predominantly early visual areas. No ISC difference between similarity groups was observed in any DMN ROI. DLPFC = dorsolateral prefrontal cortex; MPFC = medial prefrontal cortex; ang = angular gyrus, PMC = posterior medial cortex. * $p < .05$, ** $p < .01$, *** $p < .001$. Error bars are SEM.

Together, these results suggest that the DMN tracks the ToM interpretation and not the non-social interpretation of the narrative.

3.2.2.3 RSA: Interpretation similarity is correlated with neural similarity across modalities

To further investigate the relationship between interpretation similarity and neural similarity, we conducted a cross-modal intersubject RSA between Movie subjects and Audio subjects.

RSA in Audio groups. By design, the interpretations of the two audio recording were explicit and unambiguous. Due to the lack of variability in interpretation across subjects (Figure 3.2A), no significant voxels were identified by comparing the variance in

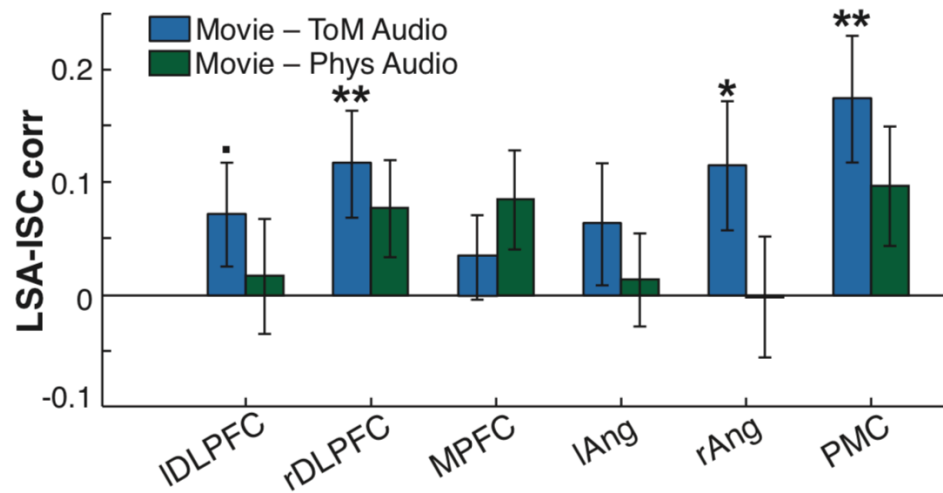


Figure 3.5 Cross-modal intersubject RSA. Across Movie and ToM Audio subjects, greater recall similarity was correlated with neural similarity in a subset of DMN ROIs: PMC and DLPFC ROIs (blue). No relationship between recall and neural similarity was observed in DMN for Movie and Phys Audio subjects (green). DLPFC = dorsolateral prefrontal cortex, mPFC = medial prefrontal cortex, ang = angular gyrus, PMC = posterior medial cortex, * $p < .05$, ** $p < .01$. Error bars are SEM.

neural and recall similarity in the ToM Audio group or the Phys Audio group in whole-brain analyses or DMN ROI analyses.

RSA across Movie and Audio groups. We additionally compared interpretation similarity and neural similarity between the Movie group and the two Audio groups. In the whole-brain analysis, no voxels passed significance testing in the Movie-ToM Audio comparison, although using a lower threshold revealed largely the same voxels as in the within-group Movie RSA. The ROI analysis did reveal significant correlations between neural and recall similarity between the two groups in right DLPFC, right angular gyrus, and PMC ($q < .05$, FDR corrected; Figure 3.5, blue bars), as well as marginally in left DLPFC ($p < .1$).

Finally, no voxels or ROIs passed thresholding for the Movie and Phys Audio comparison ($q < .05$, FDR corrected; Figure 3.5, green bars).

3.3 Discussion

Communication can be ambiguous and can be mediated by very different forms including written or spoken language, gestures, signed language and images. It can also be conveyed by moving geometric shapes, as in present work. Here, we showed that regions of the default mode network are invariant to communication modality, yet sensitive to differences in interpretation.

3.3.1 Shared processing of narratives across modality

Despite vast differences in the physical properties of moving geometric shapes and spoken words, the Movie and ToM audio elicited significantly correlated neural responses throughout the DMN, as well as in face-selective regions of the mid-fusiform gyrus. Moreover, consistent with our findings in Chapter 1, this shared response is modulated by shared interpretation: the more similar the interpretations were between someone who watched the shapes animation and someone who heard the interpretive narrative, the more similar their neural responses were in PMC, right DLPFC, and right angular gyrus ROIs.

The modality invariance of the DMN activation extends previous work showing that these regions respond similarly to the same concept presented in different modalities, including individually presented words versus images of the same item (Chee et al., 2000; Marinkovic et al., 2003; Bruffaerts et al., 2013), spoken versus written sentences, paragraphs or narratives (Spitsyna et al., 2006; Jobard et al., 2007; Lindenberg and Scheef, 2007; Regev et al., 2013), and audio-visual versus spoken narratives (Zadbood et al., 2018; Baldassano et al., 2018; Chen et al., 2016). However, the present work is the first to show that despite vast differences in stimulus properties, sparse and abstract stimuli like triangles and squares can elicit similar neural responses to explicit verbal storytelling as long as both stimuli induce similar interpretations.

3.3.2 Shared processing of visual features across modality

The shapes animation and its spoken physical description elicited significantly correlated responses in visual cortex (V1+, VO, LO) and bilateral superior parietal lobule. Shared neural activity between the movie and its physical description in visual regions and SPL is consistent with mental imagery during the listening task. While listening, subjects may have been mentally visualizing moving shapes and their spatial relations. Previous work has demonstrated early visual cortex activation during mental imagery (Kosslyn and Thompson, 2003; Slotnick et al., 2005; Cichy et al., 2012; Albers et al., 2013), while SPL has been implicated in mental rotation of shapes (Gogos et al., 2010; Prescott et al., 2010). In addition, SPL and the neighboring intraparietal sulcus are involved in processing geometric and mathematical information such as numerosity (Piazza et al., 2004; Harvey et al., 2013), geometric shape terms (Zhang et al., 2012), and mathematical visual-spatial representations more broadly (Dehaene et al., 1999). Shared, cross-modal activity in SPL may thus arise from shared geometric information over time during the shapes animation and the spoken physical description.

3.3.2 Shape animations and mentalizing

Together, these findings suggest that neural processing of animate shape movies is contingent on idiosyncratic interpretation. Researchers have long suggested that animate shape movies activate a network of mentalizing regions in the same way that making inferences about people in other theory of mind tasks do. Interestingly, these mentalizing regions largely overlap with the DMN (Schilbach et al., 2008; Mar, 2011; Mars et al., 2012), in which we observe both modality invariance and sensitivity to interpretation.

Our results thus suggest that when people attribute animacy to a triangle, it not only activates the same regions that a perceiving person would, but that the triangle is represented and processed in the same way as a person is. Indeed, activity in middle fusiform gyrus, a classically face-selective region of ventral-temporal cortex, is significantly correlated between the shapes movie and the interpretive audio suggesting that people can literally perceive and represent triangles as people.

3.4 Conclusions

This line of work simultaneously underscores the remarkable modality invariance and sensitivity to idiosyncratic interpretation in the default mode network. It attests to the potency of narrative interpretation in driving neural activity in high-order brain regions and implicates a role for early visual cortex in mental imagery.

Chapter 4

Teacher-student neural coupling during teaching and learning

Humans have a unique ability to share knowledge of the world with each other via communication. Often times, communication takes the form of sharing narratives or recalling memories in social contexts. In such cases, interlocutors have a common background and shared knowledge, which facilitates communication. However, in other contexts, such as in teaching and learning, human communication is often asymmetric and involves the transfer of highly novel, non-social information from an expert (teacher) to a novice (student). Students may lack the necessary contextual knowledge to follow the material, requiring teachers to build shared background knowledge. What neural processes enable such transfer of information across brains during naturalistic teaching and learning of complex abstract information?

Verbal communication of social narratives elicits correlated neural responses between speakers and listeners in regions overlapping with the default mode network (DMN; Stephens et al., 2010; Dikker et al., 2014; Silbert et al., 2014; Zadbood et al., 2017). This “speaker-listener coupling” is driven by shared understanding of the narrative. Speaker and listener brains are coupled in the DMN only when the speaker

uses real words in a language the listener understands, and the magnitude of DMN speaker-listener coupling is correlated with story comprehension (Stephens et al., 2010; Silbert et al., 2014). Moreover, shared DMN responses among listeners is sensitive to background context or knowledge that enables similar comprehension of interpretation of the narrative. While listening to ambiguous narratives, subjects who receive contextualizing background information have more similar neural responses in DMN than subjects who do not (van Kesteren et al., 2010; Ames et al., 2014; Chen et al., 2015a; Oren et al., 2017). Subjects who are biased by context to interpret a narrative in the same way have more similar DMN responses with each other than with subjects who are biased towards a different interpretation (Yeshurun et al., 2017b; Chien and Honey, 2020). Individuals also bring substantial prior knowledge about the expected structure of narrative events in the form of “schemas” (Schank and Abelson, 1977; Fillmore, 1985; Zacks and Tversky, 2001; Zacks et al., 2001). Researchers have shown schema-consistent narratives evoke stronger shared neural responses in high-level regions than schema-inconsistent narratives (van Kesteren et al., 2010; Oren et al., 2017; Baldassano et al., 2018) as well as during other types of schema knowledge (van Kesteren et al., 2013; Brod et al., 2015).

However, during communication of non-narrative information, such as highly-technical, abstract concepts, individuals may lack a common context or background for communicating effectively. The teacher, or expert, has abstract, complex information represented in their brain which must be transmitted to students, who may initially lack the context or schema required to understand this novel information. Here, we suggest that the process of establishing this shared knowledge via teaching is reflected in coupled

neural processes between the teacher and students in regions of the DMN. While previous fMRI studies have identified speaker-listener alignment in the communication of narratives, speaker-listener alignment has only recently been extended to the communication of non-narrative, technical information using EEG (Dikker et al., 2017 p.20; Bevilacqua et al., 2019; Davidesco et al., 2019) and fNIRS (Zheng et al., 2018; Pan et al., 2020). However, these studies have only inconsistently observed a relationship between teacher-student alignment and learning. Moreover, while EEG and fNIRS have the benefit of portability and better temporal resolution, fMRI allows for full-brain coverage with high spatial resolution. EEG in particular does not allow for measurements from regions of interest in the DMN, including posterior medial cortex and medial prefrontal cortex.

The present work extends prior research on the neural basis of real-life learning by using fMRI to investigate teacher-student alignment during teaching and learning of complex scientific material. We scanned a teacher giving an extended, 32-min lecture on a scientific topic followed by a 6-min review of the material. Students were then scanned while watching the video lessons and taking pre- and post-lesson quizzes to assess learning. Based on previous work, we predicted watching engaging video lessons would elicit shared, or correlated, neural responses among students in regions ranging from early sensory cortex to high-level regions of the default mode network (DMN). We additionally predicted that the shared neural response among students in high-level DMN regions is driven by activity in the teacher's brain, and that this coupling of the teacher and students' brains will correlate with learning outcomes. Finally, because the main

lesson serves to create a shared knowledge between teachers and students, we predicted that teacher-student coupling will be greater during the review than the lesson.

4.1 Methods

4.1.1 Subjects

One teacher (author M.N.; age: 27; female) was scanned using fMRI while giving a verbal lecture and review with accompanying slides. Forty-eight subjects with normal hearing and normal or corrected-to-normal vision participated in the experiment as students. The data from four subjects were excluded from analysis for excessive motion during scanning ($>3\text{mm}$), two for falling asleep during scanning, and two for significant pre-existing knowledge of the lesson material (scored $\geq 80\%$ on the pre-test). This left 20 subjects (ages 19–34, mean = 22.2 years; 13 female) who participated in the Intact Learning condition and 20 subjects (ages 18–26, mean=19.65 years; 10 female) who participated in the Scrambled Lesson control condition. All experimental procedures were approved by the Princeton University Internal Review Board, and all subjects provided written, informed consent.

4.1.2 Experimental design

4.1.2.1 Stimuli

The teacher developed a 32-min Lesson and a 6-min Review of the lesson, both with slides (Figure 4.1A). In order to engage the subjects with the lesson, we selected fMRI as the lesson topic, covering an introduction to fMRI, fMRI physics, action potentials, and

fMRI task design

A. Teacher



B. Stimuli



- What is fMRI?
- MR physics
- Action potentials
- BOLD signal

C. Students (Intact Learning)



D. Students (Scrambled Control)



Figure 4.1 Teaching and learning fMRI task design. **(A)** A teacher was scanned in fMRI giving a 32-min lesson and a 7-min review with slides, repeating each five times in the scanner while her voice were recorded. **(B)** The lecture and review introduced fMRI concepts. **(C)** Students in the Intact Learning condition took a Pre-test quiz and watched the Lesson in the scanner on the first day of the study. On the following day, they were scanned watching the Review and then completed a Post-test quiz. **(D)** Subjects in the Scrambled condition took the same Pre- and Post-test quiz, but watched a scrambled version of the Lesson in the scanner. They did not watch the Review.

the biological basis for the blood-oxygen-level-dependent (BOLD) signal (Figure 4.1B).

The teacher was then scanned using fMRI while giving the Lesson and Review. The teacher repeated the same Lesson and Review five times across three different scanning sessions (Lesson duration: mean = 32:20 min, range = 31:29–34:29 min; Review duration: mean = 5:58 min, range = 5:54–6:13 min). During teaching, slides were presented using PsychoPy 2 (Peirce et al., 2019) and presented via an LCD projector on a rear-projection screen mounted in the back of the scanner bore. The teacher viewed slides through a mirror mounted on the head coil and advance slides using a button box. The teacher's speech was recorded using an MRI-compatible microphone with online sound canceling (FOMRI III; Optoacoustics Ltd).

Following the teacher scans, the audio recordings were further denoised using Adobe Audition, transcribed and timestamped, and then aligned with slide presentation. The highest-quality recordings were then selected to present to students. The Lesson

presented to students was 32:19 min, and the Review was 6:13 min (see Appendices 4 and 5 for links to videos and full transcripts). To minimize transient, non-selective responses that occur at the abrupt onset of a stimuli, the lectures were preceded by the same unrelated 37s movie clip followed by a 6 second fixation cross. Videos ended with another 6 seconds of fixation cross. The unrelated movie clip and fixation periods were cropped from subsequent analyses.

Finally, the teacher developed a 25-question, multiple choice quiz to assess learning (Appendix 6). The same quiz was administered before and after learning (Pre-test, Post-test).

4.1.2.2 Student sessions: Intact Learning

Student subjects completed the main experiment on two consecutive days. In the Intact Learning condition, on Day 1, subjects ($N = 20$) took the multiple choice quiz (Pre-test) assessing their pre-existing knowledge of the lesson topic. They were then scanned in fMRI watching the 32-min Intact Lesson. On Day 2, subjects watched the 6-min Review, completed additional scans that were not analyzed here, and then answered the same 25-question quiz (Post-test). Stimuli were presented and viewed in an identical setup as during the teacher sessions. Audio was played through MRI-compatible insert earphones (Sensimetrix models S14 and S15). At the start of each scanning session, a short unrelated audio clip of the teacher was played during a scan to calibrate audio volume. Following scanning, students were asked to rate how closely they attended to the video on a scale of 1–5 (1 = not at all attentive, 5 = extremely attentive).

4.1.2.3 *Student sessions: Scrambled Lesson control*

To test that neural coupling between the teacher and students reflected shared processing of lesson content, rather than shared perception of low-level stimulus features, a second sample of subjects ($N = 20$) was scanned while watching a Scrambled Lesson control. In this control, the 32-min Intact Lesson was scrambled at the sentence level, while the text on the slides was scrambled at the word level. Subjects in the Scrambled Lesson control also completed the experiment on two consecutive days: on Day 1, students took the Pre-test and were scanned watching the Scrambled Lesson control and then made the same ratings as Intact Learning subjects; on Day 2, students took the Post-test outside of the scanner.

4.1.3 MRI acquisition

Subjects were scanned in a 3T Siemens Magnetom Skyra scanner at the Princeton Neuroscience Institute using a 64-channel head/neck coil (Siemens). During functional scans, volumes were acquired using a T2*-weighted gradient-echo echo-planar pulse sequence (TR 2000ms; TE 28ms; voxel size 3x3x3mm with 38 slices; flip angle 80°; FOV 192x192 mm²; matrix size 64x64; slice orientation axial; anterior-to-posterior phase encoding; interleaved slice acquisition; iPAT GRAPPA 2) with whole-brain coverage. For both Intact and Scrambled Lesson, total scan duration was 966 volumes (32:12 min). For the Review, total scan duration was 207 volumes (6:54 min). A high-resolution anatomical image was collected using a T1-weighted MPRAGE pulse sequence (TR

2300ms; TE 2.98ms; TI 900ms; voxel size 1x1x1mm with 176 slices; flip angle 9°; FOV 256x256 mm²; slice orientation axial; no fat suppression; total scan duration 5:20 min).

4.1.4 MRI data analysis

4.1.4.1 Preprocessing

MRI data were preprocessed using FSL 5.0 and FEAT 6.0 (FMRIB, Oxford; (Jenkinson et al., 2012) including motion correction, linear trend removal, high-pass filtering (cutoff: 140s or ~ 0.007 Hz), and spatial smoothing with a 6-mm FWHM Gaussian kernel. Motion correction was performed using FSL's MCFLIRT with 6 degrees of freedom. Subjects with excessive head motion (>3 mm absolute displacement) were discarded from the analysis. Functional data were registered to high-resolution structural images and then to 3-mm MNI standard space using FSL's FLIRT with 12 degrees of freedom. Preprocessed data were z-scored over time. All analyses were conducted in volume space using custom MATLAB and Python scripts and visualized using Nilearn (<https://nilearn.github.io>; (Abraham et al., 2014) and NeuroElf (<https://neuroelf.net>).

4.1.4.2 Temporal interpolation procedure for Teacher scans

Scans of the teacher giving the Lesson and Review were separately averaged to provide a more reliable measure of the teacher's neural responses. The teacher was extensively trained to reproduce the same Lesson and Review with high fidelity in timing and word use. To account for natural and minor variation in timing, the recorded Lessons and Reviews were timestamped at the approximate sentence level (137 segments, mean

length = 13.78 sec, SD = 4.11 sec). The teacher's neural response was then linearly interpolated within each timestamped sentence in order to align stimulus content over time relative to the actual recordings presented to the student subjects. The interpolated teacher neural responses were then averaged and used in subsequent analyses.

4.1.4.3 Unscrambling procedure for Scrambled Lesson control

The order of sentences in the Scrambled Lesson were randomized, so to compute coupled neural responses between the Teacher and Student Scrambled Lesson, the Scrambled Lesson response timeseries were reordered to match the temporal order of the Intact Lesson. Following Lerner et al. (2011), sentence-level timestamps were shifted by 6 seconds (3 TRs) to account for the hemodynamic response function, and then the response timeseries were segmented by sentence and reordered. Sentences less than six seconds long were removed from analysis. The resulting reordered timeseries was used in the Teacher-Student "Unscrambled" Lesson analyses described below. The full, original scrambled timeseries was used in Student-Student Scrambled analyses.

4.1.4.4 Intersubject correlation (ISC)

Shared neural responses among subjects were measured using intersubject correlation (ISC; Hasson et al., 2004; Nastase et al., 2019). Four ISC analyses were conducted: (1) among students in the Intact Lesson and Review (Student-Student), (2) among the scrambled students in the Scrambled Lesson control (Student-Student Scrambled), (3) between the teacher and the students in the Intact Lesson and Review (Teacher-Student),

and (4) between the teacher and the students in the Scrambled Lesson control (Teacher-Student Unscrambled). ISC was calculated at each voxel within an anatomically-defined gray matter mask. For within-group analyses among students (analyses 1 and 2), ISC was calculated using the leave-one-out approach: each student's response time course was correlated to the average response time course of the remaining ($N - 1$) students at each voxel. ISC was separately calculated among the students in the Intact Learning condition for both Lesson and Review and among the students in the Scrambled Control condition for the Scrambled Lesson only. The same analysis was repeated for the between-group Student and Teacher analyses (analyses 3 and 4). In these analyses, ISC was calculated by correlating each student's response time series to the average Teacher response time series.

Statistical significance of ISC was assessed using a bootstrap hypothesis test following Chen et al. (2016) and Nastase et al. (2019). In each iteration of the bootstrap, we randomly sampled $N = 20$ subject ISC values with replacement then computed the mean of the bootstrap sample. We repeated this procedure 10,000 times, producing a bootstrapped distribution around the mean ISC value across subjects. We then subtracted the observed mean ISC value from the bootstrap distribution to create a null distribution, and used this null distribution to calculate p -values. We corrected for multiple comparisons across voxels by controlling the false discovery rate (FDR) (Benjamini and Hochberg, 1995) at $q = .05$.

4.1.4.5 ISC in Lesson versus Review

We predicted that Teacher-Student coupling during the Review will be greater than during the Lesson, as the initial Lesson serves to build a shared knowledge background. To test this prediction, we directly compared Teacher-Student and Student-Student ISCs in the Intact Lesson and Review conditions using paired *t*-tests (two-tailed) in every region of a 61-region parcellation comprising regions of the brain that respond reliably to naturalistic audiovisual stimuli (Regev et al., 2018). Statistical significance was assessed using a sign test: in each permutation, the sign of the difference in the two conditions was randomly flipped, and the *t*-value was calculated under the null hypothesis of no systematic difference between conditions. This procedure was repeated 10,000 times to produce a permutation-based null distribution, which was used to calculate *p*-values. We corrected for multiple comparisons using FDR ($q = .05$).

4.1.4.6 Correlation with learning outcomes

To test the prediction that the level of shared neural responses will be correlated with learning, we first calculated a normalized measure of learning outcome: normalized score = $(\text{Posttest} - \text{Pretest}) * \text{mean}(\text{Posttest} + \text{Pretest})$. This measure incorporates both improvement in score and total score, balancing the amount a student learns with an absolute measure of how much they know. Two students may improve by the same amount, but a student who has a higher total score, indicating greater understanding of the material, will have a higher normalized score.

We correlated both Student-Student ISC and Teacher-Student ISC in the Intact Lesson and Review with the normalized score in each of the 61 parcels. Statistical significance of the ISC–score correlations was assessed using a permutation test in each parcel. In each permutation, the ISC values were randomly shuffled across subjects and correlated with the normalized scores. This procedure was repeated 10,000 times, producing a permutation-based null distribution (under the null hypothesis of no systematic relationship between ISC values and learning scores across subjects). We again corrected for multiple comparisons across parcels using FDR ($q = .05$).

4.1.4.7 Lagged correlation

Previous research has suggested that during verbal communication, the listener’s neural response follows the speaker’s neural response with a several second lag in certain brain areas. We therefore tested for a lag between student and teacher neural responses during the Intact Lesson in two parcels, posterior cingulate cortex (PCC) and V1+, by shifting the Students’ timeseries by -10 to +10 TRs relative to the teacher’s timeseries and calculating ISC at each lag. The same analysis was repeated for Student-Student ISC in the Intact Lesson by shifting each subject relative to the average of others.

4.2 Results

4.2.1 Behavioral results

Subjects in the Intact Learning condition improved significantly from the Pre-test to the Post-test Quiz ($t(19) = 10.35, p < .001$), scoring on average 48.8% (SD = 12.2%) on the

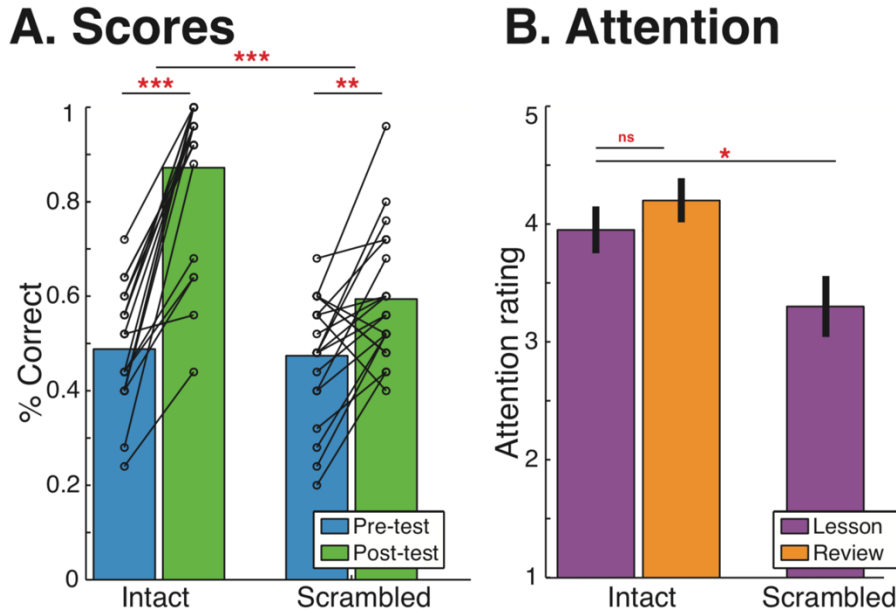


Figure 4.2 Behavior: Learning scores and attention. **(A)** Students in both the Intact and Scrambled conditions significantly improved their scores from Pre-test to Post-test (Intact Learning: $t(19) = 10.35, p < .00001$; Scrambled Lesson control: $t(19) = 3.47, p = .0026$), but Students learned significantly more in the Intact condition ($t(38) = 5.21, p < .00001$). **(B)** On average, students were very attentive during the videos, but students in the Intact Lesson were more attentive than students in the Scrambled Lesson ($t(38) = 2.27, p = .029$). *** $p < .00001$, ** $p < .01$, * $p < .05$

Pre-test and 87.2% (SD = 17.5%) on the Post-test (Figure 4.2A). All subjects in the Intact Learning condition improved their score. Subjects in the Scrambled Lesson control condition also significant improved their scores (Pre-test: mean = 47.4%, SD = 13.3%; Post-test: 59.4%, SD = 14.1%; $t(19) = 3.47, p = .003$), but by significantly less than the subjects in the Intact Learning condition ($t(38) = 5.21, p < .001$). Several subjects in the Scrambled Lesson control did not improve their scores. The two groups did not differ significantly in Pre-test scores ($t(38) = .345, p = .73$), indicating a similar baseline of prior knowledge.

Subjects in both conditions were attentive throughout the video lessons (Figure 4.2B). On a five-point scale (1 = not at all attentive, 5 = extremely attentive), subjects in

the Intact Learning condition rated their attention during the Lesson on average to be 3.95 (SD = .76) and during the Review to be 4.20 (SD = .70). There was no significant difference between attention during the Lesson and the Review ($t(19) = 1.56, p = .14$). During the Scrambled Lesson control, subjects rated their attention as 3.30 (SD = 1.03), which is significantly lower than the subjects in the Intact Lesson ($t(38) = 2.27, p = .029$).

4.2.2 fMRI results

4.2.2.1 *Aligned neural responses among students during learning*

We first identified regions of the brain that are significantly correlated responses across students during the Lesson and Review in the Intact Learning condition. Consistent with previous work using audiovisual narrative stimuli, we observe widespread significant ISC in during both the Lesson and Review ($q < .05$, FDR corrected voxel-wise; Figure 4.3A) throughout much of visual cortex, auditory and linguistic regions from early auditory cortex (A1+) to superior temporal gyrus (STG) and middle temporal gyrus (MTG), and extending into higher-order DMN regions including posterior medial cortex (PMC), bilateral dorsolateral prefrontal cortex (DLPFC), and right medial prefrontal cortex (mPFC). However, unlike previous work on narrative processing (e.g. Hasson et al., 2008; Lerner et al., 2011; Yeshurun et al., 2017; Nguyen et al., 2019), we also observe strong ISC in bilateral superior parietal lobule (SPL). There were no significant differences between Student-Student ISC during the Lesson versus Review.

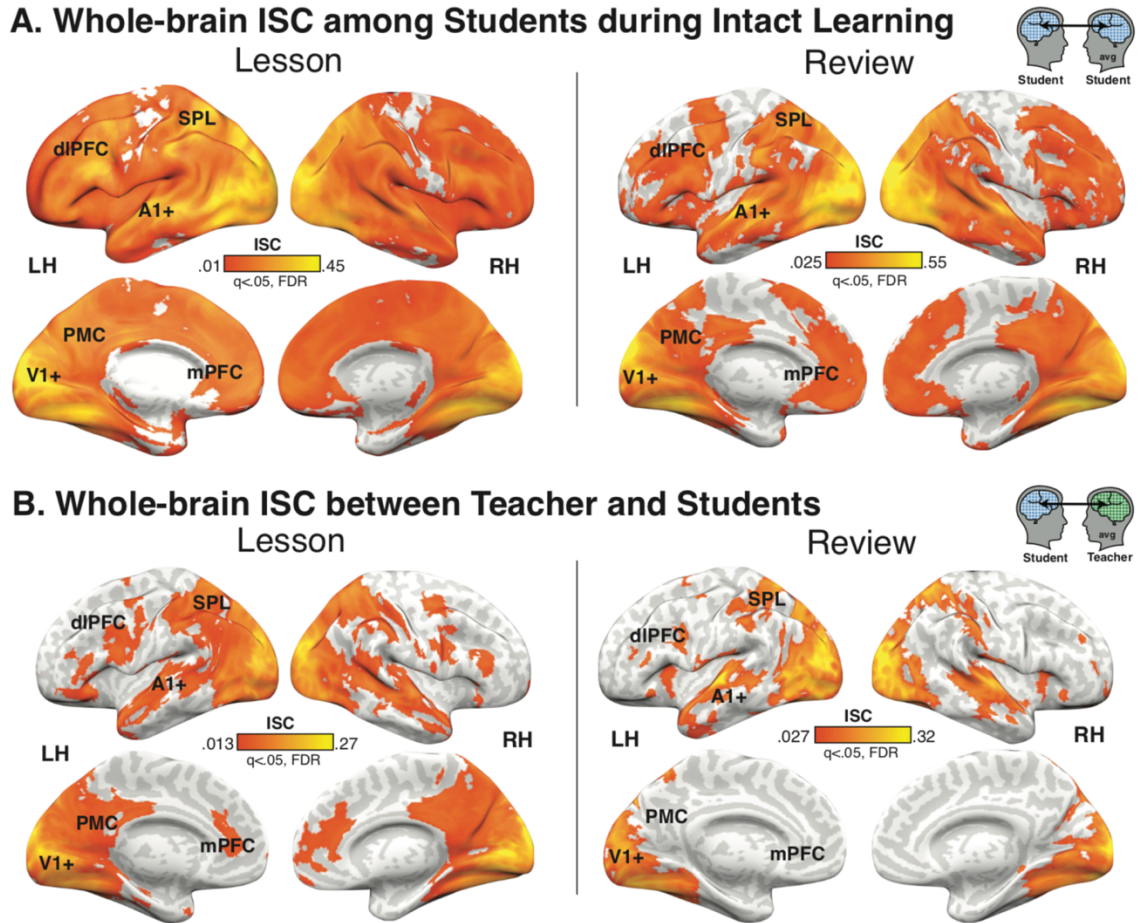


Figure 4.3 Student-Student and Teacher-Student ISC during teaching and learning. **(A)** Students in the Intact Learning condition have significant ISC throughout the cortex in both the Lesson and Review, with the strongest ISC in visual and auditory cortex, superior and middle temporal gyrus, bilateral SPL, bilateral DLPFC, and right mPFC. **(B)** Student neural responses are coupled to the teacher's response in similar regions. dIPFC = dorsolateral prefrontal cortex, SPL = superior parietal lobule, PMC = posterior medial cortex, mPFC = medial prefrontal cortex (non-parametric bootstrap hypothesis test; $q < .05$, FDR corrected).

4.2.2.2 Teacher-student neural coupling during teaching and learning

We next calculated ISC between the Teacher and Student neural responses in the Intact Learning condition to identify regions of the brain that are coupled during teaching and learning (Figure 4.3B). Similar to speaker-listener coupling during narrative storytelling, significant Teacher-Student ISC was observed in early sensory cortices, linguistic and

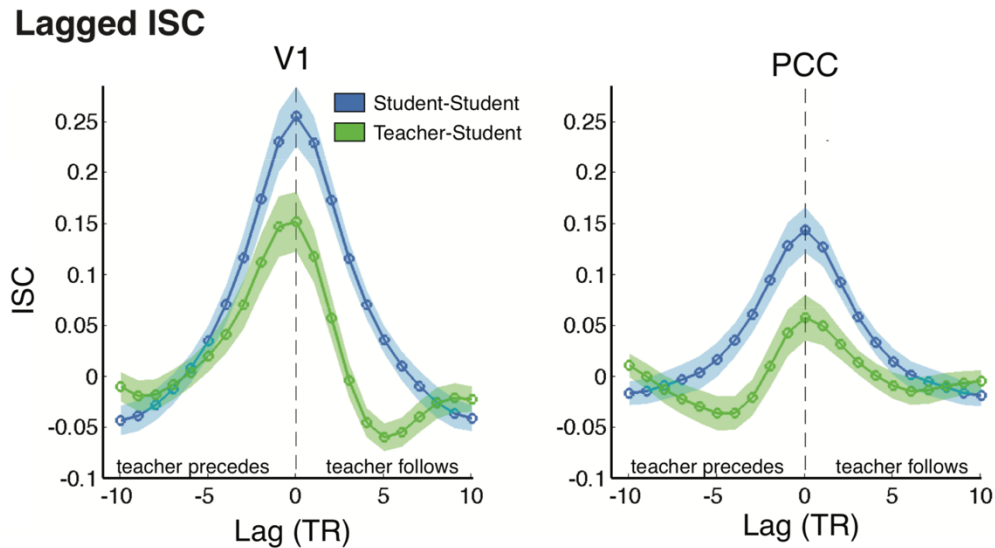


Figure 4.4. Lagged ISC for Student-Student and Teacher-Student coupling. **(A)** There was no lag between the Student and Teacher neural response during the Lesson in V1, consistent with previous results. **(B)** However, unlike previous results, there also was no lag between the Teacher and the Student in higher-order areas such as PCC. Indeed, the Teacher-Student lagged correlation (blue) closely matched the Student-Student lagged correlation (green). PCC = posterior cingulate cortex, TR = 1.5 sec

extralinguistic regions including STG, MTG, and temporal pole, and DMN regions including PMC, dlPFC, and mPFC. Unlike coupling during narrative storytelling, we additionally observe significant ISC in bilateral SPL and notably weak ISC in bilateral TPJ. There were no significant differences between Teacher-Student ISC during the Intact Lesson versus Review.

The lagged correlation analysis does not indicate a lag between the Teacher and Students' neural responses (Figure 4.4). In both V1 and PCC, Student-Student and Teacher-Student ISC peaks at lag = 0, suggesting that fluctuations in neural response in the Students and Teacher was locked to slide progression as the lectures unfold over time.

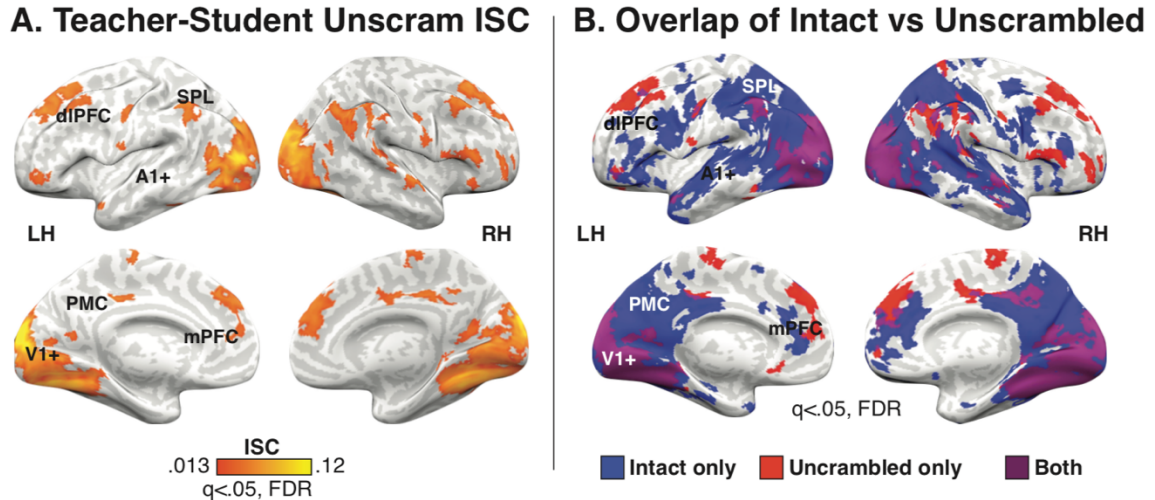


Figure 4.5 Teacher-Student ISC in Intact versus Unscrambled Lesson. **(A)** ISC between the teacher and the students in the Scrambled Lesson control. The Student neural responses were “unscrambled” to reorder events in the original sequences prior to correlating with the teacher. Only sentences longer than 6 seconds were included in the analysis. **(B)** ISC in both the Teacher-Student Intact Lesson and Teacher-Student Unscrambled Lesson conditions were significant in visual cortex (purple). However, only the Intact condition showed significant ISC in high-level DMN regions, including PMC, MPFC, and DLPFC (blue). Surprisingly, the Scrambled condition also showed significant ISC in dorsal PFC (red). dIPFC = dorsolateral prefrontal cortex, SPL = superior parietal lobule, PMC = posterior medial cortex, mPFC = medial prefrontal cortex ($q < .05$, FDR corrected).

4.2.2.3 Teacher-student ISC emerges in DMN only during intact learning

Shared neural responses between the Teacher and Students could be driven by shared sensory input rather than shared high-level understanding of the lesson content as the Teacher and Students view and hear the same visual and auditory stimuli. To dissociate high-level content-related shared responses from low-level sensory-related shared responses, we scanned an second group of students who watched a temporally scrambled version of the Lesson. This manipulation preserves the low-level visual and auditory properties of the stimulus while disrupting conceptual understanding and reducing learning (Figure 4.2A). The neural response of students in the Scrambled Lesson control

were “unscrambled” to reorder events in the same order as the original Intact Lesson and then correlated with the Teacher’s neural response (Figure 4.5A).

Consistent with the hypothesis that Teacher-Student coupling in high-order areas reflects shared understanding, not just shared stimulus processing, we observed significant Teacher-Student ISC in DMN regions only in the Intact Learning conditions (Figure 4.5B). These regions include PMC, bilateral angular gyrus, bilateral temporal poles, mPFC, and lateral PFC. Teacher-Student Unscrambled ISC largely overlapped with the ISC in the Intact Lesson in visual areas, extending from V1+ to ventral-temporal cortex and dorsal occipital regions. In addition, however, there was significant Teacher-Student Unscrambled ISC in attention-related regions of dorsal PFC that did not overlap with the Intact Lesson ISC map.

4.2.2.4 Teacher-Student ISC correlates with learning

If Teacher-Student coupling reflects shared understanding of novel information, we also expect that coupling in high-level regions, but not low-level sensory regions, will be related to learning outcomes. We therefore correlated both Student-Student and Teacher-Student ISC with a measure of learning outcomes (improvement in score weighted by total score) across each of 61 areas from an independent parcellation (Regev et al., 2018). The analysis revealed seven regions with a significant correlation between Teacher-Student ISC during the lesson and learning outcomes (Figure 4.6, FDR corrected). Five

Teacher-Student ISC is correlated with learning

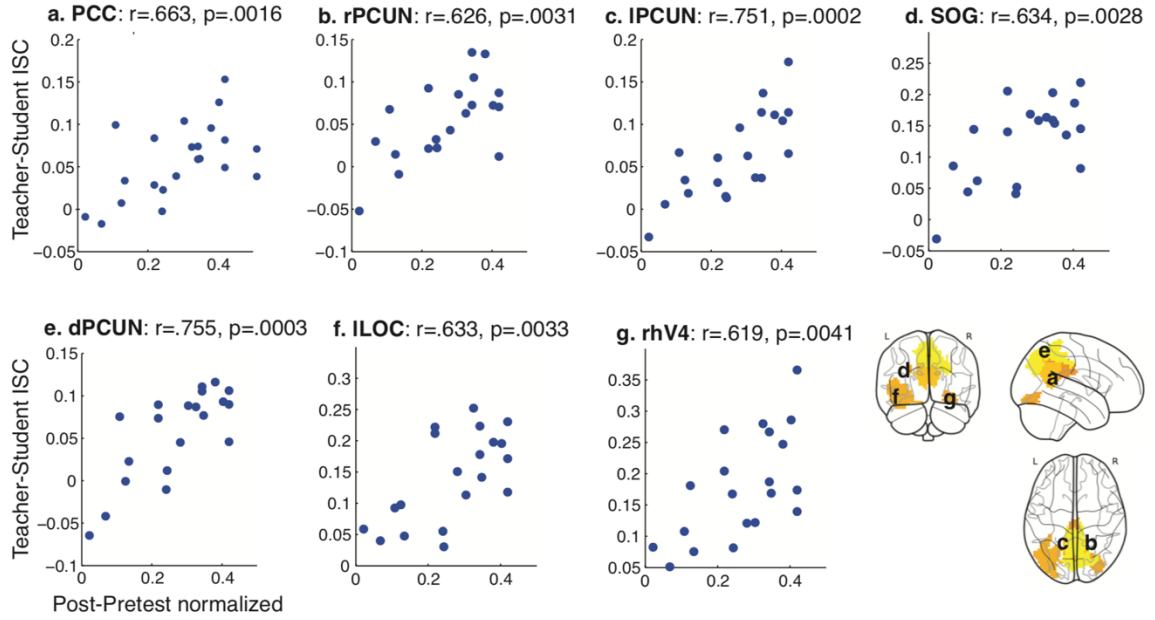


Figure 4.6 Teacher-Student ISC correlates with learning outcomes. Teacher-student ISC is correlated with normalized improvement in quiz score (Post-Pre/ (1/avg(Post+Pre))) in seven parcels, five in posterior medial cortex and two in high-level visual cortex ($q < .05$, FDR corrected). PCC = posterior cingulate cortex, rPCUN = right precuneus, IPCUN = left precuneus, SOG = superior occipital gyurs, dPCUN = dorsal precuneus, ILOC = left lateral occipital complex, rhV4 = right human V4. ROIs from 61 ROI parcellation (Regev et al., 2018). Inset shows significant parcels on glass brain visualization.

regions were located primarily in posterior medial cortex: posterior cingulate cortex (PCC; $r = .663$, $p = .002$), right precuneus (rPCUN; $r = .626$, $p = .003$), left precuneus (IPCUN; $r = .751$, $p < .001$), superior occipital gyrus (SOG; $r = .634$, $p = .003$), and dorsal precuneus (dPCUN; $r = .755$, $p < .001$). Two regions are in high-order visual cortex: left lateral occipital complex (ILOC; $r = .633$, $p = .003$) and right hV4 (rhV4; $r = .619$, $p = .004$). Notably, these areas largely overlap with regions that only show significant Teacher-Student coupling in the Intact Learning condition.

While the correlation between Student-Student ISC and improvement in quiz score showed a similar pattern of effects, no region passed FDR correction. There was no relationship between learning and either Student-Student or Teacher-Student ISC during the review.

4.3 Discussion

Teaching and learning are key processes by which humans build cultural and technical knowledge, enabling the effective transmission of novel and sometimes highly abstracted information (Hasson et al., 2012; Hasson and Frith, 2016). Here, we report the first fMRI study to identify high-level regions of the brain that are coupled between teachers and students during naturalistic teaching and learning of novel information. These regions primarily overlap with the DMN, and teacher-student coupling in these regions predict learning outcomes.

We observe teacher-student coupling in widespread regions of the brain, extending from visual and auditory areas to high-level areas including DMN. While teacher-student alignment in early sensory regions is likely driven by shared stimulus features, our results suggest that alignment in higher-order regions is driven by shared understanding of lesson content. First, when learning was disrupted by temporally scrambling the lesson, teacher-student coupling was only observed in early visual and linguistic regions, consistent with shared processing of stimulus features (visual input, individual phrases or sentences) on short timescales of a few seconds (Hasson et al., 2008; Lerner et al., 2011; Chen et al., 2015a). Teacher-student alignment in DMN and

bilateral SPL only emerged during intact learning, consistent with these regions supporting integration of information over longer timescales of minutes (Hasson et al., 2008; Lerner et al., 2011; Honey et al., 2012a; Stephens et al., 2013) and high-level context-dependent understanding (Honey et al., 2012b; Regev et al., 2013; Ames et al., 2014; Yeshurun et al., 2017b; Nguyen et al., 2019). Second, teacher-student coupling in sub-regions of PMC during the intact lesson were significantly correlated with learning outcomes, suggesting that students who learn the most are the ones who are most coupled with the teacher. Notably, PMC was also among the regions that only showed teacher-student coupling during intact learning.

Our findings are consistent with several recent teacher-student hyperscanning studies: in EEG, teacher-student synchrony in the alpha band (8-12 Hz) is correlated to delayed learning outcomes (Davidesco et al., 2019), while fNIRS studies have found an association in inferior frontal regions and superior temporal regions that is related to learning (Pan et al., 2018, 2019; Zheng et al., 2018). While EEG and fNIRS have the benefit of portability and better temporal resolution, fMRI allows for full-brain coverage with high spatial resolution, including measurements from regions of interest in the DMN that are not accessible by other imaging modalities. Moreover, we employed an extended lecture (32 min) on a single scientific topic, more closely mimicking the typical college-level lecture than the short lessons on unrelated topics that have previously been used. Finally, we recorded the teacher's brain as she delivered the lessons over multiple sessions, increasing the measurement reliability of the teacher's neural responses and allowing us to relate both teacher-student and student-student neural coupling to a measure of learning that takes into account both improvement in score and total score.

Teacher-student coupling was observed in many of the same regions that show coupled responses between speaker and listener during narrative storytelling (Stephens et al., 2010; Dikker et al., 2014; Silbert et al., 2014). These regions include linguistic regions STS and STG extending to the temporal poles, as well as DMN regions including PMC, mPFC, and dlPFC. However, unlike during narrative storytelling, and in accordance with the scientific nature of the lecture, we also observed teacher-student alignment in bilateral SPL, which has been implicated in mental rotation, geometry, and processing of mathematical symbols (Gogos et al., 2010; Prescott et al., 2010; Zhang et al., 2012; Harvey et al., 2013), as well as the manipulation and rearrangement of information in working memory (Koenigs et al., 2009). We do not observe teacher-student coupling in bilateral TPJ, a region of the brain that is strongly implicated in mentalizing and theory of mind (Saxe and Kanwisher, 2003; Schilbach et al., 2008; Mars et al., 2012), suggesting that aspects of speaker-listener coupling are content specific.

Overall, these findings suggest that neural coupling between teachers and students can be used as an index of learning. While the biological processes that give rise to neural alignment has not yet been elucidated, we suggest that neural alignment reflects shared representation of semantic knowledge (Vodrahalli et al., 2018; Cetron et al., 2019; Nastase et al., 2019b). Teaching is therefore a process of building or creating this shared semantic knowledge and aligning representations between students (novices) and teachers (experts). For some concepts (e.g. “net magnetization” or “action potential”), students may be starting from scratch, requiring teachers to build up new concepts from their existing semantic or conceptual knowledge. In other cases, however, teaching may require shifting existing concepts to new contexts. For example, the word “BOLD” in

everyday life describes a fearless or daring person or idea. In fMRI, however, “BOLD” refers to the blood-oxygen-level-dependent response. Teachers (or experts) may be able to flexibly shift between conceptual representations of the word “BOLD” while students (or novices) will need to more laboriously shift representations from “fearless” to “fMRI.” Alignment of conceptual knowledge and shared understanding of terms is then reflected in shared neural responses between teachers and students. Joint attention also likely has a role in synchronizing neural responses among individuals (Golumbic et al., 2013; Dikker et al., 2014; Cohen et al., 2018; Regev et al., 2018).

4.3.1 Limitations and future directions

Unlike previous studies (Stephens et al., 2010; Silbert et al., 2014; Zadbood et al., 2017; Davidesco et al., 2019), we did not observe a reliable temporal lag between the teacher’s and the listener’s neural responses. The time-locked visual stimulus viewed by both the teacher and the students may have exerted a strong bottom-up, visual signal that was absent in narrative storytelling studies that used auditory-only stimuli. In addition, the low temporal resolution of fMRI and the temporal interpolation procedure for averaging teacher sessions may obscure a temporal lag. Indeed, a recent EEG study with both visual and auditory processing identified a short, 200-ms lag between the teacher and student neural responses (Davidesco et al., 2019).

We also did not observe significant differences in the level of neural coupling between the first Lesson video and the second Review video. Based on previous research suggesting that context increases neural coupling, we had hypothesized that the initial

lesson would serve to create a shared knowledge base. This shared background would then be reflected in greater student-student and teacher-student alignment during the review than during the lesson. The lack of significant differences between the two scans could be due to countervailing differences between the two stimuli. For example, this result may reflect the nature of teaching: a teacher gradually builds common ground with students by connecting novel ideas to basic concepts. Using this “scaffolding” approach, the teacher should effectively guide the student’s learning process to minimize gaps in understanding, thus reducing the difference in coupling between learning and review.

The role of attention in communication and eliciting shared neural responses in an area of active research (Schmälzle et al., 2015; Cohen and Parra, 2016; Cohen et al., 2018; Regev et al., 2018). Indeed, some have suggested that joint attention to stimulus features during processing of naturalistic stimuli may underlie coupled neural responses (Cohen et al., 2018; Dikker et al., 2019). Lower Teacher-Student Unscrambled coupling in high-level regions, including DMN, may reflect lower attention rather than disrupted learning. However, we also observe that Teacher-Student coupling in sub-regions of the DMN is correlated with learning outcomes. Thus while shared attention is necessary for shared understanding, both likely play a role in inducing neural coupling.

Future studies are needed to further explore the building and shifting of shared representations of concepts over time. While we do not observe differences in teacher-student alignment between lesson and review, studies over longer time periods (e.g. over a semester of classes) or using higher temporal resolution imaging coupled with more frequent measures of learning may be able to detect changes in neural alignment reflecting changes in conceptual alignment. Such work may dissect the content of

learning by taking advantage of semantic embeddings (e.g., Mikolov et al., 2013).

Finally, future studies are needed to further explore the effects of joint attention during interactive learning and teaching using hyperscanning. While the present work explores unidirectional communication and learning, which is increasingly common in the era of online learning, much of learning occurs interactively and in person. During in-person teaching, teachers may be able to adjust their style and explanation following real-time student feedback, enabling better teacher-student alignment and hence learning in contrast to no-feedback conditions.

4.4 Conclusions

In conclusion, this work provides evidence that shared understanding of technical, non-narrative information during teaching and learning is reflected in the coupled neural responses between teachers and students in the DMN and other high-level regions. These findings speak to the flexibility and importance of the DMN in integrating novel information over long time periods and provides insight into how our brains build a shared understanding via communication.

Chapter 5

General discussion

The studies in this dissertation identified shared and idiosyncratic neural representations in the default mode network (DMN) during the processing of narrative and non-narrative communication. In Chapter 2, we introduced a novel method, intersubject representational similarity analysis, and showed that similarity in idiosyncratic interpretation of an abstract, ambiguous animation across individuals is correlated with neural similarity in the DMN. In Chapter 3, we extended these findings to show that the same narrative presented in different forms elicited similar neural responses in the DMN as a function of shared interpretation, despite vast differences in low-level visual and acoustic properties of the stimuli. Finally, in Chapter 4, we expanded beyond shared neural responses among receivers of narrative information to coupled neural responses between senders and receivers in a non-narrative learning context. We showed that a teacher's brain drives shared neural responses among students in the DMN, and that this coupling emerged as a product of shared understanding of novel information.

Together, these findings suggest that activity in the DMN is shaped by communication and that it is at once highly sensitive to individual differences in understanding and relatively invariant to the form of communication. While much

research on the DMN has focused on internally oriented or non-stimulus induced thought, we suggest that findings in this dissertation point to a role of the DMN in integrating internally- and externally-oriented cognition during communication over long timescales.

5.1 A new view of the default mode network

Converging evidence shows that the DMN is, on the one hand, robustly driven by exogenous naturalistic stimuli (narratives, technical information, visual and auditory modalities), eliciting significantly synchronized neural activity across individuals that is locked to the content on the stimulus (Hasson et al., 2004; Honey et al., 2012b; Regev et al., 2013; Lerner et al., 2014; Nastase et al., 2019a; Nguyen et al., 2019, 2020a). On the other hand, activity in the DMN is modulated by internal representations (context, background, memory, individual differences) that may vary substantially across individuals (Ames et al., 2014; Yeshurun et al., 2017b; Finn et al., 2018, 2019; Nguyen et al., 2019, 2020a). We therefore suggest that the DMN is not a default or intrinsic network that is only active when individuals are not engaged with the external world. Instead, we propose that the DMN is an active and dynamic network that integrates incoming extrinsic information with prior internal information represented in memory to form rich, context-dependent models of the situation as they dynamically unfold over time. By taking into account an individual's personal background and societal context, activity in the DMN reflects both shared and idiosyncratic components of making sense of the external world. Through communication, we suggest that the DMN provides a space for social others to shape the self, which in turn can enable individuals to shape others. The

unique interplay between the extrinsic and intrinsic forces provides a mechanism for negotiating a shared neural code to facilitate communication via shared language.

5.1.2 Individual differences and the DMN: integrating internal representations with ongoing information

Let's take an excerpt from the short story, "Pretty Mouth and Green My Eyes," by J.D. Salinger. In the story, the main character, Arthur, calls his friend and asks, "Have you seen my wife?" Research on timescale processing (see Chapter 1.2) showed that the DMN integrates information over many seconds to minutes and longer into the past (Hasson et al., 2008; Lerner et al., 2011). Responses in the DMN during the current, external stimulus ("wife") is therefore dependent on a coherent stimulus past ("Have you seen my ..."). In other words, the internally represented *stimulus past* is integrated with the external *stimulus present* in the DMN (Figure 5.1). Work later showed that the DMN also integrates information *outside of the stimulus* to the surrounding context (Cooper et al., 2011; van Kesteren et al., 2013; Ames et al., 2014; Lahnakoski et al., 2014; Chen et al., 2015b, 2015a; Oren et al., 2017; Yeshurun et al., 2017b; Chien and Honey, 2020). From the example above with Arthur and his friend, interpretation of the question "Have you seen my wife?" differs if the reader thinks Arthurs' friend is sleeping with his wife ("cheating" context) versus if they think Arthur has a faithful wife but is paranoid ("paranoia" context) (Figure 5.1). This difference in interpretation is reflected in differences in neural response in the DMN (Yeshurun et al., 2017b).

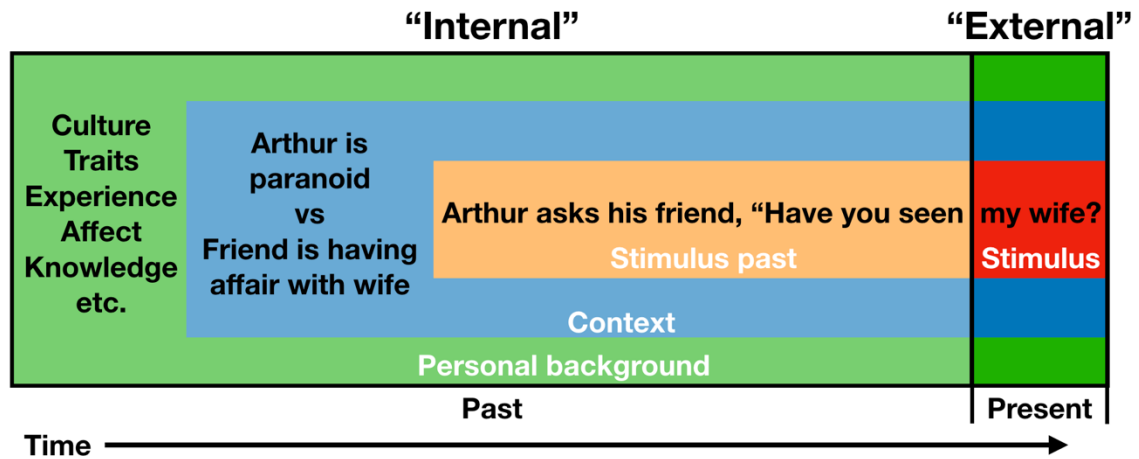


Figure 5.1 DMN integrates internal and external information over long timescales. Processing the stimulus present (“my wife”) depends on the recent stimulus past, (“Arthur asks his friend, ‘Have you seen ...’”), but also on the surrounding context prior to the stimulus (Arthur is paranoid versus Arthur’s friend is sleeping with Arthur’s wife) and the reader’s own personal background. Activity in the DMN is therefore contingent the integration of external stimuli with internal representations of the past and personal characteristics.

We now suggest that the DMN integrates even further back in time to capture internal, individual differences such as trait-level personality characteristics, cultural background, pre-existing knowledge, and political view, and that the interplay of these internal characteristics is reflected in individual differences in the interpretation and neural processing of naturalistic stimuli in the DMN. These internal representations or characteristics are not manipulated by the experimenter, but arise from the larger social context and environments in which individuals are embedded (Figure 5.1).

In this dissertation, we show for the first time that individual differences in the interpretation of narratives (Chapters 2 and 3) is reflected in the differences in the similarity of neural activity: people who interpret a narrative more similarly, even across modalities, have more similar neural responses in the DMN (Nguyen et al., 2019). These

findings are consistent with a few very recent studies that have also shown idiosyncratic responses in the DMN as function of shared understanding. Saalasti et al. (2019) closely replicated the Chapter 2 results using a spoken narrative, while others have shown a relationship between similarity in DMN activity and similarity of emotional responses during spoken narratives (Smirnov et al., 2019), similarity in humor perception (Jääskeläinen et al., 2016), similarity in sociosexual desire during erotic movies (Chen et al., 2019), similarity in affective response (Chang et al., 2018), and similarity in moral decision-making choices (Tei et al., 2019; van Baar et al., 2019). Emerging work also suggests that functional connectivity patterns between individuals in heteromodal cortex, including frontoparietal control network (FCPN), attention networks, and default mode network, can be used to decode individuals at rest and during movie stimulation (Vanderwal et al., 2017).

While these studies and the work described Chapters 2 and 3 did not collect other behavioral or personality trait data, we suggest that differences in interpretation and neural responses may arise from individual differences such as cultural background, creativity, mentalizing, pre-existing knowledge, cognitive style, affective state, belief systems, and more. Going back to the example with Arthur and his missing wife, the reaction or processing of this narrative depends on the context (“cheating” versus “paranoia”), but also on the listener’s individual experiences and traits such as their previous experience with infidelity, their preference for monogamous or non-monogamous relationships, their willingness to trust others, their level of paranoia, and cultural or societal norms about marriage. Indeed, recent studies have begun exploring individual characteristics that may give rise to differences in interpretation and neural

processing in the DMN. These include trait level paranoia (Finn et al., 2018), working memory (Finn et al., 2019), analytic versus holistic perspective (Bacha-Trams et al., 2018), and socio-sexual desire (Chen et al., 2019). Group-level differences in neural responses between cultures have also been observed during processing of culturally relevant stimuli, showing that the similarity of responses among individuals in a culture is related to the individual's degree of identification with the culture (Kaplan et al., 2016; Levy et al., 2016). Several researchers have also observed correlations between neural similarity in regions of the DMN and measures of mental health in clinical populations, including depression (Guo et al., 2015; Gruskin et al., 2019) and schizophrenia (Yang et al., 2019).

5.1.2.1 What about the hippocampus?

The role of the hippocampus in sustaining the DMN's ability to integrate over long timescales and recruit prior information stored in memory remains an area of active research. The DMN and medial temporal lobe/hippocampus are highly anatomically and functionally coupled (Vincent et al., 2008; Aggleton, 2012). While hippocampal amnesiacs show profound impairments in episodic memory (Scoville and Milner, 1957) and perform poorly on typical lab tasks involving memorization of random associations between isolated stimuli (Squire and Shimamura, 1986; Spiers et al., 2001), patients can often sustain short coherent conversations (Scoville and Milner, 1957) or recall coherent narratives (Baddeley and Wilson, 2002). Consistent with behavioral evidence, neuroimaging work suggests that DMN integration over at least 10–30 second periods does not depend on hippocampus. Chen et al. (2015a) found that hippocampus activity, as

well as hippocampus-DMN functional connectivity, is less reliable across individuals during coherent processing narratives that rely on recent memory (previous minutes) versus distant memory (previous day). Similarly, van Kesteren et al. (2010) observed hippocampal–vmPFC coupling only during processing of a non-coherent narrative. Together these two studies suggest that hippocampus is not recruited for coherent processing of intact stimuli. Most intriguingly, a hippocampal amnesiac patient was recently found to have intact narrative processing in DMN (Zuo et al., 2020). The patient showed normal ISC in the DMN while listening to an intact narrative. Moreover, like neurotypical controls, the patient’s DMN activity was disrupted by temporally scrambling the narrative, suggesting that patient’s DMN successfully integrated information over windows of at least 10-30 seconds. Further research however is needed to better characterize the role of hippocampus on longer timescales.

5.1.3 Communication and DMN: integrating self-input with input from other minds

Shared activity in the DMN is driven by both internal and external sources of information. Shared activity, however, is ultimately driven by other brains: one brain’s internal source of information (stories, memories, ideas) are externalized into physical signals such as words, sounds, movies, TV shows, and lectures. This external signal is then detected by other brains, eliciting neural activity that is driven by, or coupled to, activity in the first brain. The receiver-receiver model of shared responses during processing of external stimuli can thus be expanded to a dynamical sender-receiver model of communication indexed by coupled neural responses between senders and receivers.

Such “brain-to-brain coupling,” observed primarily in DMN regions (Stephens et al., 2010; Silbert et al., 2014), has been argued to underlie successful communication and the formation of a shared social environment (Hasson et al., 2012; Hasson and Frith, 2016). Here, we build on this work to suggest that DMN integrates self-input (internal representations of context, personal characteristics, affect, etc.) with input from other minds (stories, ideas, speech, written text, etc.). The integration of these two streams of information is reflected, first, in the shared activity between senders and receivers, and, second, in divergence of receivers from senders as a function of individual differences in shared understanding.

While most of the work demonstrating sender-receiver coupling during communication has focused on narratives (Stephens et al., 2010; Kuhlén et al., 2012; Silbert et al., 2014; Heidlmayr et al., 2019; Smirnov et al., 2019), in Chapter 4 we extend this work to teaching and learning of non-narrative, technical information. We observe teacher-student coupling throughout the DMN, and that teacher-student coupling in posterior medial cortex is correlated with learning outcomes (Nguyen et al., 2020a). These results suggest that senders and receivers show coupled neural activity during non-narrative communication, but that the extent of coupling is dependent on shared understanding of the information. Several recent studies have observed a similar relationship between learning and teacher-student or student-student coupling using EEG (Cohen et al., 2018; Davidesco et al., 2019) and fNIRS (Pan et al., 2018, 2019; Zheng et al., 2018).

These findings are important for two key reasons: First, they show that neural coupling in the DMN is not dependent on the type of information communicated. Some

have argued that shared responses among listeners and between speaker and listeners in the DMN during narratives may reflect the need to keep track of animate characters and their mental states, thereby recruiting the mentalizing network which overlaps with the DMN (Mars et al., 2012; Schilbach et al., 2013). During teaching of technical information, however, there were no humans or animate characters in the lecture.

Coupled neural responses therefore have to reflect the long timescale and semantic aspect of the stimuli, rather than the presence of animate characters. Second, we suggest that teaching and learning is a prime example of using communication to develop a shared context or background among interlocutors. Unlike describing a personal story, memory, or movie, in which predictable events or event structure can enhance coupling (Dikker et al., 2014; Baldassano et al., 2017; Zadbood et al., 2017; Heidlmayr et al., 2019), students may have very little shared background or context with the teacher to facilitate communication. For example, the word “BOLD” in everyday life describes a fearless or daring person or idea. In fMRI, however, “BOLD” refers to the blood-oxygen-level-dependent response. Teachers (or experts) may be able to flexibly shift between conceptual representations of the word “BOLD” while students (or novices) will need to more laboriously shift representations from “fearless” to “fMRI.” The aim of teaching then is to create a shared context that either shifts or creates semantic representations of concepts and enabling communication of novel ideas.

Relatively little work has investigated how internal representations and individual differences affect sender-receiver neural coupling. Several studies, like the findings in Chapter 4, have focused on coupling as a function of shared understanding, showing that speaker-listener coupling is modulated by how well a listener understands the speaker

(Stephens et al., 2010; Silbert et al., 2014; Davidesco et al., 2019; Pan et al., 2019). A few studies have explored the role of context in modulating speaker-listener coupling. Baldassano et al. (2017) and Zadbood et al. (unpublished results) show that listeners with prior knowledge of a narrative anticipate the speaker's activity earlier than listeners without prior knowledge. Speaker-listener coupling must therefore integrate internal memory representations with processing of incoming speech. In a study that used context to manipulate interpretation of a speaker's utterances, Heidlmayr et al. (2019) scanned speakers telling ambiguous stories and then scanned listeners listening to these stories with or without disambiguating context. The authors found that speaker-listener coupling was greater in the context than no-context condition in dorsolateral prefrontal cortex. Across conditions, speaker-listener coupling in DMN regions was correlated with comprehension. This novel finding suggests that speaker-listener coupling is dependent on the listener understanding the speaker in the way that the speaker intends. Finally, individual differences in emotional alignment of listeners with speakers is associated with individual differences in similarity of neural response with speakers (Smirnov et al., 2019). Individual differences in likeability ratings are also associated with differences in neural coupling, with listeners who like the speaker more showing greater neural alignment with the speaker (Bevilacqua et al., 2019).

5.2 Importance of naturalistic paradigms

The work in this dissertation highlights the importance of naturalistic paradigms, which aim to mimic the complexity and dynamism of the world outside the laboratory.

Traditionally, neuroimaging work uses tightly controlled but artificial stimuli and experimental design, aiming to contrast different conditions and localize discrete regions of the brain specialized for specific cognitive tasks. This reductionist approach subdivides cognition into separable, discrete processes such as working memory, long-term memory, perception, language comprehension, speech production, task-related activity, intrinsic activity, and other component processes (Ben-Yakov et al., 2012; Hasson and Honey, 2012; Hasson et al., 2012; Schilbach et al., 2013; García and Ibáñez, 2014; Liu and Pelowski, 2014; Kasai et al., 2015; Nastase et al., 2020).

In everyday life though, these processes are not generally separable. In natural speech, for example, words gain meaning in the context of a sentence, which in turn gain meaning in the context of sentences, paragraphs, dialogues and discourses. An entire dialogue occurs between multiple individuals with their own mental states and personal backgrounds who alternate between speaking and listening. The online processing of an individual word (language comprehension, semantics), then, requires processing it in the context of words, sentences, and paragraphs preceding the word (memory, language), attending to the speaker (attention, perception), making inferences about the speaker's mindset (mentalizing), integrating with the larger personal context (schemas, situation models, previous knowledge, beliefs), and then formulating an appropriate response (memory, language comprehension, language production, mentalizing, and more). By employing naturalistic stimuli and paradigms, such as watching an animation or following a lecture, neuroscientists sidestep these kinds of divisions between types of cognition and aim to better understand how the brain integrates information to support real-life behaviors. Researchers have thus shown that language comprehension and

production are coupled, recruiting overlapping and shared neural processes (Stephens et al., 2010; Silbert et al., 2014); that perception and memory are intrinsically linked in neural circuits at different timescales (Hasson et al., 2015; Chen et al., 2016b; Lee et al., 2020); that perception, construction/mental imagery, communication, and memory are linked in a “communication circle” that recruits the same regions of the brain (Zadbood et al., 2017); and that context, individual differences, and past knowledge are linked to online stimulus processing (Ames et al., 2015; Yeshurun et al., 2017b; Finn et al., 2018, 2019; Nguyen et al., 2019, 2020a; Chien and Honey, 2020), as shown in this dissertation.

5.3 Future directions and open questions

The work in this dissertation suggest that individuals who share understanding with each other or of a stimulus show shared neural responses in regions of the DMN. An open question is how these high-level, concept-driven shared neural responses emerge or change over time, both during development and during social interactions and learning.

5.3.1 Brain-to-brain coupling over time

In early development, infants show DMN-like networks in intrinsic connectivity at less than 12 months (Gao et al., 2015) and even pre-term (Fransson et al., 2007). The DMN develops and differentiates itself from other networks throughout childhood into early adulthood (Supekar et al., 2010; Sherman et al., 2014). Linking the development of the DMN to the development of narrative processing (construction and comprehension) and sense-making would shed important insight into how these neural systems mature and support adult-level cognition. A few studies have shown that shared neural responses in

the DMN are present from early to middle childhood (Moraczewski et al., 2018, 2020) and develop through childhood and adolescence (Moraczewski et al., 2018, 2020; Lerner et al., 2019) to become more adult-like. Two studies linked similarity in neural responses in DMN to development of mathematical and verbal ability (Cantlon and Li, 2013) and mental state inference during narratives (Moraczewski et al., 2020). Further studies could explore how changes in shared responses through development are linked to narrative processing (construction and comprehension), language development, and learning of abstract topics.

Another intriguing area of inquiry is investigating how social interaction shapes representations in DMN over time. In a striking study, Parkinson et al. (2018) scanned a cohort of business school students and found similarity in neural responses in DMN was correlated with relationship closeness: people who have closer relationships have more similar neural responses. Do individuals become friends with people who are similar to them to start with? Or do individuals become more similar to the people they spend time with as social interactions shape and change the way they perceive and interact with the world? The answer is most likely a combination of both, but longitudinal work tracking a cohort over time is needed to untangle these relations. Measuring neural alignment before and after conflict resolution may also be interesting. Contact theory (Allport, 1954), for example, posits that interaction with out-group members is an effective means for long-term bias reduction. Measures of neural synchrony before and after such an intervention may provide novel insights into how interaction shapes and changes group identification and measures of bias. Finally, building on work showing neural synchrony between children and their caregivers (Reindl et al., 2018; Leong et al., 2019; Miller et al., 2019;

Nguyen et al., 2020b; Piazza et al., 2020; Santamaria et al., 2020), further longitudinal studies may also shed insight into how interactions in early childhood shape neural development.

5.3.2 Intersubject representational similarity analysis and individual differences

Intersubject representational similarity analysis, introduced in Chapter 2, opens new avenues for investigating individual differences in naturalistic contexts. In brief, intersubject RSA operates by computing a pairwise similarity matrix between every pair of subjects in an experiment. In Chapter 2, this similarity matrix is computed by comparing semantic similarity among free recalls of a narrative, but it could equally be constructed from correlating responses from personality assessments, cognitive tasks, or other individual measures. The similarity matrix is then compared to a pairwise neural similarity matrix, testing the hypothesis that similarity in behavioral measures are related to similarity in neural response in high-level brain regions. In this way, a measure of shared neural response can be used to identify regions of the brain that are sensitive to individual differences (Finn et al., 2019).

As discussed above (section 5.1.2), a handful of recent studies have used intersubject RSA to identify neural correlates of individual trait differences and behaviors. Future work should continue this line of work, while also incorporating individual differences in connectivity during the processing of naturalistic stimuli. We suggest, however, that advances in the study of individual differences in naturalistic processing will require recruiting large, diverse, and cross-cultural study populations.

Several large fMRI datasets, such as the Human Connectome Project and the Adolescent Brain Cognitive Development project can be leveraged for this purpose (see Finn et al., 2019 for example). However, like psychology, neuroimaging studies typically recruit from populations that are WEIRD – Western, Educated, Industrial, Rich, and Democratic (Henrich et al., 2010). Recruiting larger, more diverse populations could take advantage of new initiatives, such as the Psychological Science Accelerator (psysciacc.org) which recruits researchers across the globe to conduct the same experiment and aggregate data. Future studies should also leverage more portable, lower-cost neuroimaging methods (Burns et al., 2019; Dikker et al., 2019) that can more easily be used in collecting large datasets in diverse environments.

5.3.3 Towards the neuroscience of social interaction

The work in this dissertation focused on unidirectional communication: receivers watched or listened to a pre-recorded stimulus. In Chapter 4, because the stimulus was recorded by a sender (the teacher) in the scanner, we can link neural responses between receivers and senders during unidirectional communication. While much communication does occur unidirectionally (e.g. watching videos, reading instructions, etc.) such work does not capture the interactivity of in-person, social communication. Future work uncovering brain-to-brain coupling during social communication will necessitate shifting towards simultaneously scanning multiple individuals as they interact with one another in real-time, a method called “hyperscanning” (Montague, 2002). Hyperscanning has been conducted using a number of neuroimaging methods including EEG (Babiloni et al., 2006; Lindenberger et al., 2009; Naeem et al., 2012; Yun et al., 2012; Müller et al., 2013;

Konvalinka et al., 2014), MEG (Baess et al., 2012; Zhou et al., 2016), fMRI (Montague, 2002; Tomlin et al., 2006; Fliessbach et al., 2007; Saito et al., 2010; Tanabe et al., 2012; Stolk et al., 2014; Bilek et al., 2015; Koike et al., 2016), and fNIRS (Funane et al., 2011; Cui et al., 2012; Jiang et al., 2012, 2015; Osaka et al., 2014; Cheng et al., 2015; Liu et al., 2016; Piazza et al., 2020; Santamaria et al., 2020) and employing a number of experimental paradigms. As hyperscanning measures brain activity in two or more individuals simultaneously, it is uniquely situated for studying brain-to-brain coupling during dialogue. However, while there is growing interest in “two-brain neuroscience” (Dumas, 2011; Hasson et al., 2012; Menenti et al., 2012; Schilbach et al., 2013; Babiloni and Astolfi, 2014; García and Ibáñez, 2014; Liu and Pelowski, 2014; Kasai et al., 2015; Stolk et al., 2016), there is surprisingly limited hyperscanning research on naturalistic, bidirectional communication.

A few exceptions have used hyperscanning to uncover the neural signatures of bidirectional flow of information during free dialogue and other complex tasks requiring communication and cooperation. Jiang et al. (2012), for example, used fNIRS to simultaneously scan dyads engaged in free dialogue, finding that face-to-face communication elicits greater brain-to-brain coupling than back-to-back communication. Non-verbal signals may therefore play a role in enhancing communication and neural synchrony. A later study identified conversational “leaders” that emerge naturally in group discussions and used granger causality to show that the leader’s neural activity more strongly predicted the neural activity of followers than the reverse (Jiang et al., 2015). Bidirectional brain-to-brain coupling has also been observed between adults and infants as young as 9 months (Leong et al., 2019; Nguyen et al., 2020b; Piazza et al.,

2020; Santamaria et al., 2020). During structured play (singing, reading, and playing with a toy), adult and infant brains mutually influence each other, suggesting a role for both adults and children in directing coupled neural responses (Piazza et al., 2020). Other work has focused on groups of individuals completing complex, real-life tasks that require substantial goal-directed coordination among group members (Astolfi et al., 2011; Dodel et al., 2011; Stevens and Galloway, 2016; Fishburn et al., 2018), finding increased neural synchrony during periods of the task that required heightened coordination and communication. Several studies have also explored teacher-student coupling using hyperscanning in real-life (Dikker et al., 2017; Bevilacqua et al., 2019) and simulated classrooms (Zheng et al., 2018; Davidesco et al., 2019; Pan et al., 2019), though only the latter studies identified a relationship between teacher-student synchrony and learning outcomes and some restricted students from asking questions during the teaching phase (Bevilacqua et al., 2019; Davidesco et al., 2019).

These hyperscanning studies have demonstrated the feasibility of simultaneously measuring neural activity among interacting individuals in naturalistic contexts. In general, they show that brain-to-brain synchrony is stronger during interactive or coordination tasks than during periods of non-interaction. We suggest, however, that future work needs to go beyond measuring greater synchronization during certain tasks than others to include more rigorous controls and more theoretically-driven hypotheses. To this end, we propose that future hyperscanning work should have four key features: (1) bidirectional interactions among individuals that lead to emergent behaviors; (2) more rigorous control conditions; (3) stronger ties between brain-to-brain coupling and behaviors; and (4) more theoretically driven hypotheses.

1. *Interactions among individuals.* First, hyperscanning should study social interactions that result in behaviors that could not be observed in individuals by themselves or interacting with a computer. If subjects are performing the same task together versus apart, it is difficult to disentangle which aspects of shared neural activity is due to social interaction versus simply doing the same thing.
2. *Rigorous control conditions.* Rigorously showing brain-to-brain coupling is related to communication rather than shared processing of a common stimulus requires more careful controls that preserve aspects of the task except for interactive transmission of information. In single-scanner studies of communication, this control has been achieved by using narratives in a foreign language that the listener doesn't understand (Stephens et al., 2010; Silbert et al., 2014), by repeating meaningless syllables (Stephens et al., 2010; Suda et al., 2010; Takei et al., 2014), or by temporally scrambling the stimulus (Nguyen et al., 2020b). This critical control ensures that any brain-to-brain coupling identified during a dialogue is due to mutual transmission of information, not merely shared processing of low-level visual or acoustic properties.
3. *Tying brain-to-brain coupling to behavior.* Just as important as appropriate controls is showing how brain-to-brain coupling relates to measured behaviors, allowing for stronger inferences about psychological, cognitive, and neural processes during social interactions. Some hyperscanning studies already strongly relate brain-to-brain coupling to behavior. For example, Jiang et al. (2015) showed that emergent discussion leaders were more effective at inducing neural

coupling than non-leaders and that this coupling was correlated to independent ratings of leadership and communication skills.

4. *Theoretically driven hypotheses.* Finally, now that hyperscanning and measuring cross-brain similarity has been shown to be a viable method, experimental questions need to shift from “Can neural synchrony during a task be measured with hyperscanning?” to more theoretically-driven questions. Future studies could ask whether greater brain-to-brain coupling results in better coordination or cooperation in complex tasks (e.g. does a team with strong inter-brain coupling solve problems more efficiently, have fewer inter-personal problems, or coordinate tasks better?); in more effective communicating or teaching (e.g. are charismatic speakers more effective in changing people’s minds and inducing neural synchrony?); or in more empathetic or compassionate interactions (e.g. does inter-group contact change brain-to-brain coupling with an out-group member?).

5.4 Concluding remarks

In sum, this dissertation presents evidence that the default mode network is sensitive to individual differences in narrative interpretation, is relatively invariant to narrative modality, and is driven by activity in other brains during non-narrative learning. We integrate these findings with the larger literature on neural processing of naturalistic stimuli and propose a new view of the default mode network as an interface between internal representations and external stimuli.

References

- AbdulSabur NY, Xu Y, Liu S, Chow HM, Baxter M, Carson J, Braun AR (2014) Neural correlates and network connectivity underlying narrative production and comprehension: A combined fMRI and PET study. *Cortex* 57:107–127.
- Abell F, Happé F, Frith U (2000) Do triangles play tricks? Attribution of mental states to animated shapes in normal and abnormal development. *Cogn Dev* 15:1–16.
- Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, Gramfort A, Thirion B, Varoquaux G (2014) Machine learning for neuroimaging with scikit-learn. *Front Neuroinformatics* 8.
- Aggleton JP (2012) Multiple anatomical systems embedded within the primate medial temporal lobe: implications for hippocampal function. *Neurosci Biobehav Rev* 36:1579–1596.
- Albers AM, Kok P, Toni I, Dijkerman HC, de Lange FP (2013) Shared Representations for Working Memory and Mental Imagery in Early Visual Cortex. *Curr Biol* 23:1427–1431.
- Allport GW (1954) *The nature of prejudice*. Oxford, England: Addison-Wesley.
- Ames DL, Honey CJ, Chow MA, Todorov A, Hasson U (2014) Contextual Alignment of Cognitive and Neural Dynamics. *J Cogn Neurosci* 27:655–664.
- Anders S, Heinzle J, Weiskopf N, Ethofer T, Haynes J-D (2011) Flow of affective information between communicating brains. *NeuroImage* 54:439–446.
- Andrews-Hanna JR, Reidler JS, Huang C, Buckner RL (2010) Evidence for the Default Network's Role in Spontaneous Cognition. *J Neurophysiol* 104:322–335.
- Andrews-Hanna JR, Saxe R, Yarkoni T (2014) Contributions of episodic retrieval and mentalizing to autobiographical thought: Evidence from functional neuroimaging, resting-state connectivity, and fMRI meta-analyses. *NeuroImage* 91:324–335.
- Anticevic A, Cole MW, Murray JD, Corlett PR, Wang X-J, Krystal JH (2012) The Role of Default Network Deactivation in Cognition and Disease. *Trends Cogn Sci* 16:584–592.

- Ardila A, Bernal B, Rosselli M (2016) How Localized are Language Brain Areas? A Review of Brodmann Areas Involvement in Oral Language. *Arch Clin Neuropsychol* 31:112–122.
- Astolfi L, Toppi J, Borghini G, Vecchiato G, Isabella R, Fallani FDV, Cincotti F, Salinari S, Mattia D, He B, Caltagirone C, Babiloni F (2011) Study of the functional hyperconnectivity between couples of pilots during flight simulation: An EEG hyperscanning study. In: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp 2338–2341.
- Babiloni F, Astolfi L (2014) Social neuroscience and hyperscanning techniques: Past, present and future. *Neurosci Biobehav Rev* 44:76–93.
- Babiloni F, Cincotti F, Mattia D, Mattiocco M, Fallani FDV, Tocci A, Bianchi L, Marciani MG, Astolfi L (2006) Hypermethods for EEG hyperscanning. In: 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2006. EMBS '06, pp 3666–3669.
- Bacha-Trams M, Alexandrov YI, Broman E, Glerean E, Kauppila M, Kauttonen J, Ryyppö E, Sams M, Jääskeläinen IP (2018) A drama movie activates brains of holistic and analytical thinkers differentially. *Soc Cogn Affect Neurosci* 13:1293–1304.
- Bacha-Trams M, Glerean E, Dunbar R, Lahnakoski JM, Ryyppö E, Sams M, Jääskeläinen IP (2017) Differential inter-subject correlation of brain activity when kinship is a variable in moral dilemma. *Sci Rep* 7:1–16.
- Baddeley A, Wilson BA (2002) Prose recall and amnesia: implications for the structure of working memory. *Neuropsychologia* 40:1737–1743.
- Baess P, Zhdanov A, Mandel A, Parkkonen L, Hirvenkari L, Mäkelä JP, Jousmäki V, Hari R (2012) MEG dual scanning: a procedure to study real-time auditory interaction between two persons. *Front Hum Neurosci* 6.
- Baldassano C, Chen J, Zadbood A, Pillow JW, Hasson U, Norman KA (2017) Discovering Event Structure in Continuous Narrative Perception and Memory. *Neuron* 95:709–721.e5.
- Baldassano C, Hasson U, Norman KA (2018) Representation of Real-World Event Schemas during Narrative Perception. *J Neurosci* 38:9689–9699.
- Barrett LF, Satpute AB (2013) Large-scale brain networks in affective and social neuroscience: towards an integrative functional architecture of the brain. *Curr Opin Neurobiol* 23:361–372.

- Bašnáková J, Weber K, Petersson KM, van Berkum J, Hagoort P (2014) Beyond the Language Given: The Neural Correlates of Inferring Speaker Meaning. *Cereb Cortex* 24:2572–2578.
- Bell MD, Fiszdon JM, Greig TC, Wexler BE (2010) Social attribution test — multiple choice (SAT-MC) in schizophrenia: Comparison with community sample and relationship to neurocognitive, social cognitive and symptom measures. *Schizophr Res* 122:164–171.
- Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Ser B Methodol* 57:289–300.
- Ben-Yakov A, Honey CJ, Lerner Y, Hasson U (2012) Loss of reliable temporal structure in event-related averaging of naturalistic stimuli. *NeuroImage* 63:501–506.
- Berry DS, Misovich SJ, Kean KJ, Baron RM (1992) Effects of Disruption of Structure and Motion on Perceptions of Social Causality. *Pers Soc Psychol Bull* 18:237–244.
- Berry DS, Springer K (1993) Structure, Motion, and Preschoolers' Perceptions of Social Causality. *Ecol Psychol* 5:273–283.
- Bevilacqua D, Davidesco I, Wan L, Chaloner K, Rowland J, Ding M, Poeppel D, Dikker S (2019) Brain-to-Brain Synchrony and Learning Outcomes Vary by Student–Teacher Dynamics: Evidence from a Real-world Classroom Electroencephalography Study. *J Cogn Neurosci* 31:401–411.
- Bilek E, Ruf M, Schäfer A, Akdeniz C, Calhoun VD, Schmahl C, Demanuele C, Tost H, Kirsch P, Meyer-Lindenberg A (2015) Information flow between interacting human brains: Identification, validation, and relationship to social expertise. *Proc Natl Acad Sci* 112:5207–5212.
- Binder JR, Desai RH, Graves WW, Conant LL (2009) Where Is the Semantic System? A Critical Review and Meta-Analysis of 120 Functional Neuroimaging Studies. *Cereb Cortex* 19:2767–2796.
- Brod G, Lindenberger U, Werkle-Bergner M, Shing YL (2015) Differences in the neural signature of remembering schema-congruent and schema-incongruent events. *NeuroImage* 117:358–366.
- Bruffaerts R, Dupont P, Peeters R, De Deyne S, Storms G, Vandenberghe R (2013) Similarity of fMRI Activity Patterns in Left Perirhinal Cortex Reflects Semantic Similarity between Words. *J Neurosci* 33:18597–18607.
- Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The Brain's Default Network: Anatomy, Function, and Relevance to Disease. *Ann N Y Acad Sci* 1124:1–38.

- Buckner RL, Carroll DC (2007) Self-projection and the brain. *Trends Cogn Sci* 11:49–57.
- Buckner RL, DiNicola LM (2019) The brain’s default network: updated anatomy, physiology and evolving insights. *Nat Rev Neurosci* 20:593–608.
- Burns SM, Barnes LN, McCulloh IA, Dagher MM, Falk EB, Storey JD, Lieberman MD (2019) Making social neuroscience less WEIRD: Using fNIRS to measure neural signatures of persuasive influence in a Middle East participant sample. *J Pers Soc Psychol* 116:e1–e11.
- Cantlon JF, Li R (2013) Neural Activity during Natural Viewing of Sesame Street Statistically Predicts Test Scores in Early Childhood. *PLOS Biol* 11:e1001462.
- Castelli F, Frith C, Happé F, Frith U (2002) Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain* 125:1839–1849.
- Cetron JS, Connolly AC, Diamond SG, May VV, Haxby JV, Kraemer DJM (2019) Decoding individual differences in STEM learning from functional MRI data. *Nat Commun* 10:2027.
- Chang CHC, Lazaridi C, Yeshurun Y, Norman KA, Hasson U (2020) Relating the past with the present: Information integration and segregation during ongoing narrative processing. *bioRxiv*:2020.01.16.908731.
- Chang LJ, Jolly E, Cheong JH, Rapuano K, Greenstein N, Chen P-HA, Manning JR (2018) Endogenous variation in ventromedial prefrontal cortex state dynamics during naturalistic viewing reflects affective experience. *bioRxiv*:487892.
- Chee MW, Weekes B, Lee KM, Soon CS, Schreiber A, Hoon JJ, Chee M (2000) Overlap and dissociation of semantic processing of Chinese characters, English words, and pictures: evidence from fMRI. *NeuroImage* 12:392–403.
- Chen G, Shin Y-W, Taylor PA, Glen DR, Reynolds RC, Israel RB, Cox RW (2016a) Untangling the relatedness among correlations, part I: Nonparametric approaches to inter-subject correlation analysis at the group level. *NeuroImage* 142:248–259.
- Chen J, Honey CJ, Simony E, Arcaro MJ, Norman KA, Hasson U (2015a) Accessing Real-Life Episodic Information from Minutes versus Hours Earlier Modulates Hippocampal and High-Order Cortical Dynamics. *Cereb Cortex* 26:3428–3441.
- Chen J, Leong YC, Honey CJ, Yong CH, Norman KA, Hasson U (2016b) Shared memories reveal shared structure in neural activity across individuals. *Nat Neurosci* 20:115–125.

- Chen P-H (Cameron), Chen J, Yeshurun Y, Hasson U, Haxby J, Ramadge PJ (2015b) A Reduced-Dimension fMRI Shared Response Model. In: *Advances in Neural Information Processing Systems 28* (Cortes C, Lawrence ND, Lee DD, Sugiyama M, Garnett R, eds), pp 460–468. Curran Associates, Inc.
- Chen P-HA, Jolly E, Cheong JH, Chang LJ (2019) Inter-subject representational similarity analysis reveals individual variations in affective experience when watching erotic movies. *bioRxiv*:726570.
- Cheng X, Li X, Hu Y (2015) Synchronous brain activity during cooperative exchange depends on gender of partner: A fNIRS-based hyperscanning study. *Hum Brain Mapp* 36:2039–2048.
- Chien H-YS, Honey CJ (2020) Constructing and Forgetting Temporal Context in the Human Cerebral Cortex. *Neuron* 106:1-12.
- Christoff K, Irving ZC, Fox KCR, Spreng RN, Andrews-Hanna JR (2016) Mind-wandering as spontaneous thought: a dynamic framework. *Nat Rev Neurosci* 17:718–731.
- Christoff K, Keramatian K, Gordon AM, Smith R, Mädlér B (2009) Prefrontal organization of cognitive control according to levels of abstraction. *Brain Res* 1286:94–105.
- Cichy RM, Heinze J, Haynes J-D (2012) Imagery and Perception Share Cortical Representations of Content and Location. *Cereb Cortex* 22:372–380.
- Cohen SS, Madsen J, Touchan G, Robles D, Lima SFA, Henin S, Parra LC (2018) Neural engagement with online educational videos predicts learning performance for individual students. *Neurobiol Learn Mem* 155:60–64.
- Cohen SS, Parra LC (2016) Memorable Audiovisual Narratives Synchronize Sensory and Supramodal Neural Responses. *eNeuro* 3.
- Cooper EA, Hasson U, Small SL (2011) Interpretation-mediated changes in neural activity during language comprehension. *NeuroImage* 55:1314–1323.
- Corbetta M, Shulman GL (2002) Control of goal-directed and stimulus-driven attention in the brain. *Nat Rev Neurosci* 3:201–215.
- Costabile KA (2016) Narrative Construction, Social Perceptions, and the Situation Model. *Pers Soc Psychol Bull* 42:589–602.
- Csibra G, Gergely G, Bíró S, Koos O, Brockbank M (1999) Goal attribution without agency cues: the perception of ‘pure reason’ in infancy. *Cognition* 72:237–267.

- Cui X, Bryant DM, Reiss AL (2012) NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage* 59:2430–2437.
- Davidesco I, Laurent E, Valk H, West T, Dikker S, Milne C, Poeppel D (2019) Brain-to-brain synchrony between students and teachers predicts learning outcomes. *bioRxiv*:644047.
- Dehaene S, Spelke E, Pinel P, Stanescu R, Tsivkin S (1999) Sources of Mathematical Thinking: Behavioral and Brain-Imaging Evidence. *Science* 284:970–974.
- Dehghani M, Boghrati R, Man K, Hoover J, Gimbel SI, Vaswani A, Zevin JD, Immordino-Yang MH, Gordon AS, Damasio A, Kaplan JT (2017) Decoding the neural representation of story meanings across languages. *Hum Brain Mapp* 38:6096–6106.
- Deniz F, Nunez-Elizalde AO, Huth AG, Gallant JL (2019) The Representation of Semantic Information Across Human Cerebral Cortex During Listening Versus Reading Is Invariant to Stimulus Modality. *J Neurosci* 39:7722–7736.
- Dikker S, Michalareas G, Oostrik M, Serafimaki A, Kahraman HM, Struiksma ME, Poeppel D (2019) Crowdsourcing neuroscience: inter-brain coupling during face-to-face interactions outside the laboratory. *bioRxiv*:822320.
- Dikker S, Silbert LJ, Hasson U, Zevin JD (2014) On the Same Wavelength: Predictable Language Enhances Speaker-Listener Brain-to-Brain Synchrony in Posterior Superior Temporal Gyrus. *J Neurosci* 34:6267–6272.
- Dikker S, Wan L, Davidesco I, Kaggen L, Oostrik M, McClintock J, Rowland J, Michalareas G, Van Bavel JJ, Ding M, Poeppel D (2017) Brain-to-Brain Synchrony Tracks Real-World Dynamic Group Interactions in the Classroom. *Curr Biol* 27:1375–1380.
- Ding N, Melloni L, Zhang H, Tian X, Poeppel D (2016) Cortical tracking of hierarchical linguistic structures in connected speech. *Nat Neurosci* 19:158–164.
- Dixon ML, Vega ADL, Mills C, Andrews-Hanna J, Spreng RN, Cole MW, Christoff K (2018) Heterogeneity within the frontoparietal control network and its relationship to the default and dorsal attention networks. *Proc Natl Acad Sci*:201715766.
- Dodel S, Cohn J, Mersmann J, Luu P, Forsythe C, Jirsa V (2011) Brain Signatures of Team Performance. In: *Foundations of Augmented Cognition. Directing the Future of Adaptive Systems* (Schmorrow DD, Fidopiastis CM, eds), pp 288–297 *Lecture Notes in Computer Science*. Springer Berlin Heidelberg.
- Dumas G (2011) Towards a two-body neuroscience. *Commun Integr Biol* 4:349–352.

- Farbood MM, Heeger DJ, Marcus G, Hasson U, Lerner Y (2015) The neural processing of hierarchical structure in music and speech at different timescales. *Front Neurosci* 9.
- Fillmore CJ (1985) Frames and the Semantics of Understanding. *Quad Semantica* 6:222–254.
- Finn E, Glerean E, Khojandi AY, Nielson D, Molfese PJ, Handwerker DA, Bandettini P (2019) Idiosynchrony: From shared responses to individual differences during naturalistic neuroimaging. *PsyArXiv*.
- Finn ES, Corlett PR, Chen G, Bandettini PA, Constable RT (2018) Trait paranoia shapes inter-subject synchrony in brain activity during an ambiguous social narrative. *Nat Commun* 9:1–13.
- Fishburn FA, Murty VP, Hlutkowsky CO, MacGillivray CE, Bemis LM, Murphy ME, Huppert TJ, Perlman SB (2018) Putting our heads together: interpersonal neural synchronization as a biological mechanism for shared intentionality. *Soc Cogn Affect Neurosci* 13:841–849.
- Fliessbach K, Weber B, Trautner P, Dohmen T, Sunde U, Elger CE, Falk A (2007) Social Comparison Affects Reward-Related Brain Activity in the Human Ventral Striatum. *Science* 318:1305–1308.
- Fox KCR, Spreng RN, Ellamil M, Andrews-Hanna JR, Christoff K (2015) The wandering brain: Meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes. *NeuroImage* 111:611–621.
- Fox MD, Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci* 8:700–711.
- Fox MD, Snyder AZ, Vincent JL, Corbetta M, Essen DCV, Raichle ME (2005) The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc Natl Acad Sci* 102:9673–9678.
- Fransson P (2005) Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Hum Brain Mapp* 26:15–29.
- Fransson P, Skiold B, Horsch S, Nordell A, Blennow M, Lagercrantz H, Aden U (2007) Resting-state networks in the infant brain. *Proc Natl Acad Sci* 104:15531–15536.
- Friederici AD, Gierhan SM (2013) The language network. *Curr Opin Neurobiol* 23:250–254.
- Frith CD, Frith U (2006) The Neural Basis of Mentalizing. *Neuron* 50:531–534.

- Funane T, Kiguchi M, Atsumori H, Sato H, Kubota K, Koizumi H (2011) Synchronous activity of two people's prefrontal cortices during a cooperative task measured by simultaneous near-infrared spectroscopy. *J Biomed Opt* 16:077011-077011-10.
- Gao W, Alcauter S, Smith JK, Gilmore JH, Lin W (2015) Development of human brain cortical network architecture during infancy. *Brain Struct Funct* 220:1173-1186.
- García AM, Ibáñez A (2014) Two-Person Neuroscience and Naturalistic Social Communication: The Role of Language and Linguistic Variables in Brain-Coupling Research.
- Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* 9:416-429.
- Gobbini MI, Koralek AC, Bryan RE, Montgomery KJ, Haxby JV (2007) Two Takes on the Social Brain: A Comparison of Theory of Mind Tasks. *J Cogn Neurosci* 19:1803-1814.
- Gogos A, Gavrilescu M, Davison S, Searle K, Adams J, Rossell SL, Bell R, Davis SR, Egan GF (2010) Greater superior than inferior parietal lobule activation with increasing rotation angle during mental rotation: An fMRI study. *Neuropsychologia* 48:529-535.
- Golland Y, Bentin S, Gelbard H, Benjamini Y, Heller R, Nir Y, Hasson U, Malach R (2006) Extrinsic and Intrinsic Systems in the Posterior Cortex of the Human Brain Revealed during Natural Sensory Stimulation. *Cereb Cortex* 17:766-777.
- Greicius MD, Krasnow B, Reiss AL, Menon V (2003) Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proc Natl Acad Sci* 100:253-258.
- Gruskin DC, Rosenberg MD, Holmes AJ (2019) Relationships between depressive symptoms and brain responses during emotional movie viewing emerge in adolescence. *NeuroImage*:116217.
- Guo CC, Nguyen VT, Hyett MP, Parker GB, Breakspear MJ (2015) Out-of-sync: disrupted neural activity in emotional circuitry during film viewing in melancholic depression. *Sci Rep* 5:1-12.
- Gusnard DA, Raichle ME (2001) Searching for a baseline: functional imaging and the resting human brain. *Nat Rev Neurosci* 2:685-694.
- Harvey BM, Klein BP, Petridou N, Dumoulin SO (2013) Topographic Representation of Numerosity in the Human Parietal Cortex. *Science* 341:1123-1126.
- Hassabis D, Maguire EA (2007) Deconstructing episodic memory with construction. *Trends Cogn Sci* 11:299-306.

- Hassabis D, Maguire EA (2009) The construction system of the brain. *Philos Trans R Soc B Biol Sci* 364:1263–1271.
- Hasson U, Chen J, Honey CJ (2015) Hierarchical process memory: memory as an integral component of information processing. *Trends Cogn Sci* 19:304–313.
- Hasson U, Frith CD (2016) Mirroring and beyond: coupled dynamics as a generalized framework for modelling social interactions. *Philos Trans R Soc B Biol Sci* 371:20150366.
- Hasson U, Ghazanfar AA, Galantucci B, Garrod S, Keysers C (2012) Brain-to-brain coupling: a mechanism for creating and sharing a social world. *Trends Cogn Sci* 16:114–121.
- Hasson U, Honey CJ (2012) Future trends in Neuroimaging: Neural processes as expressed within real-life contexts. *NeuroImage* 62:1272–1278.
- Hasson U, Malach R, Heeger DJ (2010) Reliability of cortical activity during natural stimulation. *Trends Cogn Sci* 14:40.
- Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Intersubject Synchronization of Cortical Activity During Natural Vision. *Science* 303:1634–1640.
- Hasson U, Yang E, Vallines I, Heeger DJ, Rubin N (2008) A Hierarchy of Temporal Receptive Windows in Human Cortex. *J Neurosci* 28:2539–2550.
- Heider F, Simmel M (1944) An experimental study of apparent behavior. *Am J Psychol* 57:243–259.
- Heidlmayr K, Weber K, Takashima A, Hagoort P (2019) No title, no theme: The joined neural space between speakers and listeners during production and comprehension of multi-sentence discourse. *bioRxiv*:2019.12.25.884775.
- Henrich J, Heine SJ, Norenzayan A (2010) Most people are not WEIRD. *Nature* 466:29–29.
- Hirsch J, Adam Noah J, Zhang X, Dravida S, Ono Y (2018) A cross-brain neural mechanism for human-to-human verbal communication. *Soc Cogn Affect Neurosci* 13:907–920.
- Honey CJ, Thesen T, Donner TH, Silbert LJ, Carlson CE, Devinsky O, Doyle WK, Rubin N, Heeger DJ, Hasson U (2012a) Slow Cortical Dynamics and the Accumulation of Information over Long Timescales. *Neuron* 76:423–434.
- Honey CJ, Thompson CR, Lerner Y, Hasson U (2012b) Not Lost in Translation: Neural Responses Shared Across Languages. *J Neurosci* 32:15277–15283.

- Horan WP, Nuechterlein KH, Wynn JK, Lee J, Castelli F, Green MF (2009) Disturbances in the spontaneous attribution of social meaning in schizophrenia. *Psychol Med* 39:635–643.
- Iacoboni M, Lieberman MD, Knowlton BJ, Molnar-Szakacs I, Moritz M, Throop CJ, Fiske AP (2004) Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. *NeuroImage* 21:1167–1173.
- Jääskeläinen IP, Pajula J, Tohka J, Lee H-J, Kuo W-J, Lin F-H (2016) Brain hemodynamic activity during viewing and re-viewing of comedy movies explained by experienced humor. *Sci Rep* 6:27741.
- Jaaskelainen PI, Koskentalo K, Balk HM, Autti T, Kauramäki J, Pomren C, Sams M (2008) Inter-Subject Synchronization of Prefrontal Cortex Hemodynamic Activity During Natural Viewing. *Open Neuroimaging J* 2:14–19.
- Jansma JM, Ramsey NF, Zwart JA de, Gelderen P van, Duyn JH (2007) fMRI study of effort and information processing in a working memory task. *Hum Brain Mapp* 28:431–440.
- Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM (2012) FSL. *NeuroImage* 62:782–790.
- Jiang J, Chen C, Dai B, Shi G, Ding G, Liu L, Lu C (2015) Leader emergence through interpersonal neural synchronization. *Proc Natl Acad Sci* 112:4274–4279.
- Jiang J, Dai B, Peng D, Zhu C, Liu L, Lu C (2012) Neural Synchronization during Face-to-Face Communication. *J Neurosci* 32:16064–16069.
- Jobard G, Vigneau M, Mazoyer B, Tzourio-Mazoyer N (2007) Impact of modality and linguistic complexity during reading and listening tasks. *NeuroImage* 34:784–800.
- Kana RK, Keller TA, Cherkassky VL, Minshew NJ, Just MA (2009) Atypical frontal-posterior synchronization of Theory of Mind regions in autism during mental state attribution. *Soc Neurosci* 4:135–152.
- Kaplan JT, Gimbel SI, Dehghani M, Immordino-Yang MH, Sagae K, Wong JD, Tipper CM, Damasio H, Gordon AS, Damasio A (2016) Processing Narratives Concerning Protected Values: A Cross-Cultural Investigation of Neural Correlates. *Cereb Cortex*:bhw325.
- Kasai K, Fukuda M, Yahata N, Morita K, Fujii N (2015) The future of real-world neuroscience: Imaging techniques to assess active brains in social environments. *Neurosci Res* 90:65–71.

- Kiebel SJ, Daunizeau J, Friston KJ (2008) A Hierarchy of Time-Scales and the Brain. *PLOS Comput Biol* 4:e1000209.
- Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C, others (2007) What's new in Psychtoolbox-3. *Perception* 36:1.
- Koenigs M, Barbey AK, Postle BR, Grafman J (2009) Superior Parietal Cortex Is Critical for the Manipulation of Information in Working Memory. *J Neurosci* 29:14980–14986.
- Koike T, Tanabe HC, Okazaki S, Nakagawa E, Sasaki AT, Shimada K, Sugawara SK, Takahashi HK, Yoshihara K, Bosch-Bayard J, Sadato N (2016) Neural substrates of shared attention as social memory: A hyperscanning functional magnetic resonance imaging study. *NeuroImage* 125:401–412.
- Konvalinka I, Bauer M, Stahlhut C, Hansen LK, Roepstorff A, Frith CD (2014) Frontal alpha oscillations distinguish leaders from followers: Multivariate decoding of mutually interacting brains. *NeuroImage* 94:79–88.
- Kosslyn SM, Thompson WL (2003) When is early visual cortex activated during visual mental imagery? *Psychol Bull* 129:723–746.
- Koster-Hale J, Saxe R, Dungan J, Young LL (2013) Decoding moral judgments from neural representations of intentions. *Proc Natl Acad Sci* 110:5648–5653.
- Krieger-Redwood K, Jefferies E, Karapanagiotidis T, Seymour R, Nunes A, Ang JWA, Majernikova V, Mollo G, Smallwood J (2016) Down but not out in posterior cingulate cortex: Deactivation yet functional coupling with prefrontal cortex during demanding semantic cognition. *NeuroImage* 141:366–377.
- Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping. *Proc Natl Acad Sci U S A* 103:3863–3868.
- Kriegeskorte N, Mur M, Bandettini P (2008) Representational similarity analysis – connecting the branches of systems neuroscience. *Front Syst Neurosci*.
- Kucyi A, Esterman M, Riley CS, Valera EM (2016) Spontaneous default network activity reflects behavioral variability independent of mind-wandering. *Proc Natl Acad Sci* 113:13899–13904.
- Kuhlen AK, Allefeld C, Haynes J-D (2012) Content-specific coordination of listeners' to speakers' EEG during communication. *Front Hum Neurosci* 6.
- Lahnakoski JM, Glerean E, Jääskeläinen IP, Hyönä J, Hari R, Sams M, Nummenmaa L (2014) Synchronous brain activity across individuals underlies shared psychological perspectives. *NeuroImage* 100:316–324.

- Landauer TK, Foltz PW, Laham D (1998) An introduction to latent semantic analysis. *Discourse Process* 25:259–284.
- Lee H, Bellana B, Chen J (2020) What can narratives tell us about the neural bases of human memory? *PsyArXiv*..
- Leong V, Noreika V, Clackson K, Georgieva S, Brightman L, Nutbrown R, Fujita S, Neale D, Wass S (2019) Mother-infant interpersonal neural connectivity predicts infants' social learning. *PsyArXiv*.
- Lerner Y, Honey CJ, Katkov M, Hasson U (2014) Temporal scaling of neural responses to compressed and dilated natural speech. *J Neurophysiol* 111:2433–2444.
- Lerner Y, Honey CJ, Silbert LJ, Hasson U (2011) Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. *J Neurosci Off J Soc Neurosci* 31:2906–2915.
- Lerner Y, Scherf KS, Katkov M, Hasson U, Behrmann M (2019) Age-Related Changes in Neural Networks Supporting Complex Visual and Social Processing in Adolescence. *bioRxiv*:650887.
- Levy J, Goldstein A, Inlus M, Masalha S, Zagoory-Sharon O, Feldman R (2016) Adolescents growing up amidst intractable conflict attenuate brain response to pain of outgroup. *Proc Natl Acad Sci* 113:13696–13701.
- Lindenberg R, Scheef L (2007) Supramodal language comprehension: Role of the left temporal lobe for listening and reading. *Neuropsychologia* 45:2407–2415.
- Lindenberger U, Li S-C, Gruber W, Müller V (2009) Brains swinging in concert: cortical phase synchronization while playing guitar. *BMC Neurosci* 10:22.
- Liu N, Mok C, Witt EE, Pradhan AH, Chen JE, Reiss AL (2016) NIRS-Based Hyperscanning Reveals Inter-brain Neural Synchronization during Cooperative Jenga Game with Face-to-Face Communication. *Front Hum Neurosci* 10.
- Liu T, Pelowski M (2014) A new research trend in social neuroscience: Towards an interactive-brain neuroscience. *PsyCh J* 3:177–188.
- Liu Y, Piazza EA, Simony E, Shewokis PA, Onaral B, Hasson U, Ayaz H (2017) Measuring speaker–listener neural coupling with functional near infrared spectroscopy. *Sci Rep* 7:1–13.
- Mar RA (2011) The Neural Bases of Social Cognition and Story Comprehension. *Annu Rev Psychol* 62:103–134.
- Margulies DS, Ghosh SS, Goulas A, Falkiewicz M, Huntenburg JM, Langs G, Bezgin G, Eickhoff SB, Castellanos FX, Petrides M, Jefferies E, Smallwood J (2016)

- Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc Natl Acad Sci* 113:12574–12579.
- Marinkovic K, Dhond RP, Dale AM, Glessner M, Carr V, Halgren E (2003) Spatiotemporal Dynamics of Modality-Specific and Supramodal Word Processing. *Neuron* 38:487–497.
- Mars RB, Neubert F-X, Noonan MP, Sallet J, Toni I, Rushworth MFS (2012) On the relationship between the “default mode network” and the “social brain.” *Front Hum Neurosci*.
- Mason MF, Norton MI, Horn JDV, Wegner DM, Grafton ST, Macrae CN (2007) Wandering Minds: The Default Network and Stimulus-Independent Thought. *Science* 315:393–395.
- McKiernan KA, D’Angelo BR, Kaufman JN, Binder JR (2006) Interrupting the “stream of consciousness”: an fMRI investigation. *Neuroimage* 29:1185–1191.
- McKiernan KA, Kaufman JN, Kucera-Thompson J, Binder JR (2003) A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *J Cogn Neurosci* 15:394–408.
- Menenti L, Pickering MJ, Garrod SC (2012) Toward a neural basis of interactive alignment in conversation. *Front Hum Neurosci* 6.
- Mikolov T, Sutskever I, Chen K, Corrado GS, Dean J (2013) Distributed Representations of Words and Phrases and their Compositionality. In: *Advances in Neural Information Processing Systems 26* (Burges CJC, Bottou L, Welling M, Ghahramani Z, Weinberger KQ, eds), pp 3111–3119.
- Mildner JN, Tamir DI (2019) Spontaneous Thought as an Unconstrained Memory Process. *Trends Neurosci* 42:763–777.
- Miller JG, Vrtička P, Cui X, Shrestha S, Hosseini SMH, Baker JM, Reiss AL (2019) Inter-brain synchrony in mother-child dyads during cooperation: An fNIRS hyperscanning study. *Neuropsychologia* 124:117–124.
- Montague P (2002) Hyperscanning: Simultaneous fMRI during Linked Social Interactions. *NeuroImage* 16:1159–1164.
- Moraczewski D, Chen G, Redcay E (2018) Inter-subject synchrony as an index of functional specialization in early childhood. *Sci Rep* 8:1–12.
- Moraczewski D, Nketia J, Redcay E (2020) Cortical temporal hierarchy is immature in middle childhood. *bioRxiv*:678326.
- Müller V, Sängler J, Lindenberger U (2013) Intra- and Inter-Brain Synchronization during Musical Improvisation on the Guitar. *PLOS ONE* 8:e73852.

- Murray JD, Bernacchia A, Freedman DJ, Romo R, Wallis JD, Cai X, Padoa-Schioppa C, Pasternak T, Seo H, Lee D, Wang X-J (2014) A hierarchy of intrinsic timescales across primate cortex. *Nat Neurosci* 17:1661–1663.
- Naeem M, Prasad G, Watson DR, Kelso JAS (2012) Electrophysiological signatures of intentional social coordination in the 10–12 Hz range. *NeuroImage* 59:1795–1803.
- Nastase SA, Gazzola V, Hasson U, Keysers C (2019a) Measuring shared responses across subjects using intersubject correlation. *Soc Cogn Affect Neurosci* 14:667–685.
- Nastase SA, Goldstein A, Hasson U (2020) Keep it real: rethinking the primacy of experimental control in cognitive neuroscience. *bioRxiv*.
- Nastase SA, Liu Y-F, Hillman H, Norman KA, Hasson U (2019b) Leveraging shared connectivity to aggregate heterogeneous datasets into a common response space. *bioRxiv*:741975.
- Nguyen M, Vanderwal T, Hasson U (2019) Shared understanding of narratives is correlated with shared neural responses. *NeuroImage* 184:161–170.
- Nguyen ML, Chang A, Micciche E, Meshulam M, Nastase SA (2020a) Teacher-student neural coupling during teaching and learning. *bioRxiv*.
- Nguyen T, Schleihau H, Kayhan E, Matthes D, Vrtička P, Hoehl S (2020b) The effects of interaction quality on neural synchrony during mother-child problem solving. *Cortex* 124:235–249.
- Oatley K, Yuill N (1985) Perception of personal and interpersonal action in a cartoon film. *Br J Soc Psychol* 24:115–124.
- Oren N, Shapira-Lichter I, Lerner Y, Tarrasch R, Hendler T, Giladi N, Ash EL (2017) Schema benefit vs. proactive interference: Contradicting behavioral outcomes and coexisting neural patterns. *NeuroImage* 158:271–281.
- Osaka N, Minamoto T, Yaoi K, Azuma M, Osaka M (2014) Neural Synchronization During Cooperated Humming: A Hyperscanning Study Using fNIRS. *Procedia - Soc Behav Sci* 126:241–243.
- Pajula J, Tohka J (2016) How Many Is Enough? Effect of Sample Size in Inter-Subject Correlation Analysis of fMRI. *Comput Intell Neurosci* 2016:1–10.
- Pan Y, Dikker S, Goldstein P, Zhu Y, Yang C, Hu Y (2019) Instructor-learner brain coupling discriminates between instructional approaches and predicts learning. *bioRxiv*:704239.

- Pan Y, Novembre G, Song B, Li X, Hu Y (2018) Interpersonal synchronization of inferior frontal cortices tracks social interactive learning of a song. *NeuroImage* 183:280–290.
- Parkinson C, Kleinbaum AM, Wheatley T (2018) Similar neural responses predict friendship. *Nat Commun* 9:1–14.
- Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, Kastman E, Lindeløv JK (2019) PsychoPy2: Experiments in behavior made easy. *Behav Res Methods* 51:195–203.
- Piazza EA, Hasenfratz L, Hasson U, Lew-Williams C (2020) Infant and Adult Brains Are Coupled to the Dynamics of Natural Communication. *Psychol Sci*:12.
- Piazza M, Izard V, Pinel P, Le Bihan D, Dehaene S (2004) Tuning Curves for Approximate Numerosity in the Human Intraparietal Sulcus. *Neuron* 44:547–555.
- Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012) Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* 59:2142–2154.
- Prescott J, Gavrilescu M, Cunnington R, O’Boyle MW, Egan GF (2010) Enhanced brain connectivity in math-gifted adolescents: An fMRI study using mental rotation. *Cogn Neurosci* 1:277–288.
- Price CJ (2010) The anatomy of language: a review of 100 fMRI studies published in 2009. *Ann N Y Acad Sci* 1191:62–88.
- Raichle ME (2015) The Brain’s Default Mode Network. *Annu Rev Neurosci* 38:433–447.
- Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL (2001) A default mode of brain function. *Proc Natl Acad Sci* 98:676–682.
- Regev M, Honey CJ, Simony E, Hasson U (2013) Selective and Invariant Neural Responses to Spoken and Written Narratives. *J Neurosci* 33:15978–15988.
- Regev M, Simony E, Lee K, Tan KM, Chen J, Hasson U (2018) Propagation of Information Along the Cortical Hierarchy as a Function of Attention While Reading and Listening to Stories. *Cereb Cortex*.
- Reindl V, Gerloff C, Scharke W, Konrad K (2018) Brain-to-brain synchrony in parent-child dyads and the relationship with emotion regulation revealed by fNIRS-based hyperscanning. *NeuroImage* 178:493–502.
- Rimé B, Boulanger B, Laubin P, Richir M, Stroobants K (1985) The perception of interpersonal emotions originated by patterns of movement. *Motiv Emot* 9:241–260.

- Saalasti S, Alho J, Bar M, Glerean E, Honkela T, Kauppila M, Sams M, Jääskeläinen IP (2019) Inferior parietal lobule and early visual areas support elicitation of individualized meanings during narrative listening. *Brain Behav* 9:e01288.
- Saito DN, Tanabe HC, Izuma K, Hayashi MJ, Morito Y, Komeda H, Uchiyama H, Kosaka H, Okazawa H, Fujibayashi Y, Sadato N (2010) “Stay Tuned”: Inter-Individual Neural Synchronization During Mutual Gaze and Joint Attention. *Front Integr Neurosci* 4.
- Salter G, Seigal A, Claxton M, Lawrence K, Skuse D (2008) Can autistic children read the mind of an animated triangle? *Autism* 12:349–371.
- Santamaria L, Noreika V, Georgieva S, Clackson K, Wass S, Leong V (2020) Emotional valence modulates the topology of the parent-infant inter-brain network. *NeuroImage* 207:116341.
- Saxe R, Kanwisher N (2003) People thinking about thinking people The role of the temporo-parietal junction in “theory of mind.” *NeuroImage* 19:1835–1842.
- Schacter DL, Addis DR, Buckner RL (2007) Remembering the past to imagine the future: the prospective brain. *Nat Rev Neurosci* 8:657–661.
- Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo X-N, Holmes AJ, Eickhoff SB, Yeo BTT (2018) Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex N Y N 1991* 28:3095–3114.
- Schank RC, Abelson RP (1977) Scripts, plans, goals and understanding: An inquiry into human knowledge structures. Oxford, England: Lawrence Erlbaum.
- Schilbach L, Eickhoff SB, Rotarska-Jagiela A, Fink GR, Vogeley K (2008) Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the “default system” of the brain. *Conscious Cogn* 17:457–467.
- Schilbach L, Timmermans B, Reddy V, Costall A, Bente G, Schlicht T, Vogeley K (2013) Toward a second-person neuroscience. *Behav Brain Sci* 36:393–414.
- Schippers MB, Roebroek A, Renken R, Nanetti L, Keysers C (2010) Mapping the information flow from one brain to another during gestural communication. *Proc Natl Acad Sci* 107:9388–9393.
- Schmälzle R, Häcker FEK, Honey CJ, Hasson U (2015) Engaged listeners: shared neural processing of powerful political speeches. *Soc Cogn Affect Neurosci* 10:1137–1143.
- Scholl BJ, Tremoulet PD (2000) Perceptual causality and animacy. *Trends Cogn Sci* 4:299–309.

- Schurz M, Radua J, Aichhorn M, Richlan F, Perner J (2014) Fractionating theory of mind: A meta-analysis of functional brain imaging studies. *Neurosci Biobehav Rev* 42:9–34.
- Scoville WB, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* 20:11–21.
- Sherman LE, Rudie JD, Pfeifer JH, Masten CL, McNealy K, Dapretto M (2014) Development of the Default Mode and Central Executive Networks across early adolescence: A longitudinal study. *Dev Cogn Neurosci* 10:148–159.
- Shirer WR, Ryali S, Rykhlevskaia E, Menon V, Greicius MD (2012) Decoding Subject-Driven Cognitive States with Whole-Brain Connectivity Patterns. *Cereb Cortex* 22:158–165.
- Shulman GI, Fiez JA, Corbetta M, Buckner RL, Miezin FM, Raichle ME, Petersen SE (1997) Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J Cogn Neurosci* 9:648–663.
- Silbert LJ, Honey CJ, Simony E, Poeppel D, Hasson U (2014) Coupled neural systems underlie the production and comprehension of naturalistic narrative speech. *Proc Natl Acad Sci* 111:E4687–E4696.
- Simony E, Honey CJ, Chen J, Lositsky O, Yeshurun Y, Wiesel A, Hasson U (2016) Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat Commun* 7:12141.
- Singh KD, Fawcett IP (2008) Transient and linearly graded deactivation of the human default-mode network by a visual detection task. *NeuroImage* 41:100–112.
- Slotnick SD, Thompson WL, Kosslyn SM (2005) Visual Mental Imagery Induces Retinotopically Organized Activation of Early Visual Areas. *Cereb Cortex* 15:1570–1583.
- Smirnov D, Saarimäki H, Glerean E, Hari R, Sams M, Nummenmaa L (2019) Emotions amplify speaker–listener neural alignment. *Hum Brain Mapp* 40:4777–4788.
- Spiers HJ, Maguire EA, Burgess N (2001) Hippocampal Amnesia. *Neurocase* 7:357–382.
- Spitsyna G, Warren JE, Scott SK, Turkheimer FE, Wise RJS (2006) Converging Language Streams in the Human Temporal Lobe. *J Neurosci* 26:7328–7336.
- Spreng RN, Mar RA, Kim ASN (2009) The Common Neural Basis of Autobiographical Memory, Prospection, Navigation, Theory of Mind, and the Default Mode: A Quantitative Meta-analysis. *J Cogn Neurosci* 21:489–510.

- Spreng RN, Sepulcre J, Turner GR, Stevens WD, Schacter DL (2012) Intrinsic Architecture Underlying the Relations among the Default, Dorsal Attention, and Frontoparietal Control Networks of the Human Brain. *J Cogn Neurosci* 25:74–86.
- Spreng RN, Stevens WD, Chamberlain JP, Gilmore AW, Schacter DL (2010) Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *NeuroImage* 53:303–317.
- Spunt RP, Meyer ML, Lieberman MD (2015) The Default Mode of Human Brain Function Primes the Intentional Stance. *J Cogn Neurosci* 27:1116–1124.
- Squire LR, Shimamura AP (1986) Characterizing Amnesic Patients for Neurobehavioral Study. *Behav Neurosci* 100:866–877.
- Stephens GJ, Honey CJ, Hasson U (2013) A place for time: the spatiotemporal structure of neural dynamics during natural audition. *J Neurophysiol* 110:2019–2026.
- Stephens GJ, Silbert LJ, Hasson U (2010) Speaker-listener neural coupling underlies successful communication. *Proc Natl Acad Sci* 107:14425–14430.
- Stevens RH, Galloway TL (2016) Modeling the neurodynamic organizations and interactions of teams. *Soc Neurosci* 11:123–139.
- Stolk A, Noordzij ML, Verhagen L, Volman I, Schoffelen J-M, Oostenveld R, Hagoort P, Toni I (2014) Cerebral coherence between communicators marks the emergence of meaning. *Proc Natl Acad Sci* 111:18183–18188.
- Stolk A, Verhagen L, Toni I (2016) Conceptual Alignment: How Brains Achieve Mutual Understanding. *Trends Cogn Sci* 20:180–191.
- Suda M, Takei Y, Aoyama Y, Narita K, Sato T, Fukuda M, Mikuni M (2010) Frontopolar activation during face-to-face conversation: An in situ study using near-infrared spectroscopy. *Neuropsychologia* 48:441–447.
- Supekar K, Uddin LQ, Prater K, Amin H, Greicius MD, Menon V (2010) Development of functional and structural connectivity within the default mode network in young children. *NeuroImage* 52:290–301.
- Surian L, Geraci A (2012) Where will the triangle look for it? Attributing false beliefs to a geometric shape at 17 months. *Br J Dev Psychol* 30:30–44.
- Svoboda E, McKinnon MC, Levine B (2006) The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia* 44:2189–2208.
- Takei Y, Suda M, Aoyama Y, Sakurai N, Tagawa M, Motegi T, Yamaguchi M, Narita K, Fukuda M (2014) Near-infrared spectroscopic study of frontopolar activation during face-to-face conversation in major depressive disorder and bipolar disorder. *J Psychiatr Res* 57:74–83.

- Tanabe HC, Kosaka H, Saito DN, Koike T, Hayashi MJ, Izuma K, Komeda H, Ishitobi M, Omori M, Munesue T, Okazawa H, Wada Y, Sadato N (2012) Hard to “tune in”: neural mechanisms of live face-to-face interaction with high-functioning autistic spectrum disorder. *Front Hum Neurosci* 6.
- Tei S, Kauppi J-P, Fujino J, Jankowski KF, Kawada R, Murai T, Takahashi H (2019) Inter-subject correlation of temporoparietal junction activity is associated with conflict patterns during flexible decision-making. *Neurosci Res* 144:67–70.
- Tikka P, Kauttonen J, Hlushchuk Y (2018) Narrative comprehension beyond language: Common brain networks activated by a movie and its script. *PLoS ONE* 13.
- Tomlin D, Kayali MA, King-Casas B, Anen C, Camerer CF, Quartz SR, Montague PR (2006) Agent-specific responses in the cingulate cortex during economic exchanges. *Science* 312:1047–1050.
- Uchiyama HT, Saito DN, Tanabe HC, Harada T, Seki A, Ohno K, Koeda T, Sadato N (2012) Distinction between the literal and intended meanings of sentences: A functional magnetic resonance imaging study of metaphor and sarcasm. *Cortex* 48:563–583.
- van Baar JM, Chang LJ, Sanfey AG (2019) The computational and neural substrates of moral strategies in social decision-making. *Nat Commun* 10:1–14.
- van Kesteren MTR, Beul SF, Takashima A, Henson RN, Ruiter DJ, Fernández G (2013) Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: from congruent to incongruent. *Neuropsychologia* 51:2352–2359.
- van Kesteren MTR van, Fernández G, Norris DG, Hermans EJ (2010) Persistent schema-dependent hippocampal-neocortical connectivity during memory encoding and postencoding rest in humans. *Proc Natl Acad Sci* 107:7550–7555.
- Vanderwal T, Eilbott J, Finn ES, Craddock RC, Turnbull A, Castellanos FX (2017) Individual differences in functional connectivity during naturalistic viewing conditions. *NeuroImage* 157:521–530.
- Vanderwal T, Hunyadi E, Grupe DW, Connors CM, Schultz RT (2008) Self, mother and abstract other: An fMRI study of reflective social processing. *NeuroImage* 41:1437–1446.
- Vatansever D, Bzdok D, Wang H-T, Mollo G, Sormaz M, Murphy C, Karapanagiotidis T, Smallwood J, Jefferies E (2017a) Varieties of semantic cognition revealed through simultaneous decomposition of intrinsic brain connectivity and behaviour. *NeuroImage* 158:1–11.

- Vatansever D, Manktelow A, Sahakian BJ, Menon DK, Stamatakis EA (2018) Default Mode Network Engagement Beyond Self-Referential Internal Mentation. *Brain Connect* 8:245–253.
- Vatansever D, Manktelow AE, Sahakian BJ, Menon DK, Stamatakis EA (2017b) Angular default mode network connectivity across working memory load. *Hum Brain Mapp* 38:41–52.
- Vatansever D, Menon DK, Manktelow AE, Sahakian BJ, Stamatakis EA (2015) Default Mode Dynamics for Global Functional Integration. *J Neurosci* 35:15254–15262.
- Vidaurre D, Smith SM, Woolrich MW (2017) Brain network dynamics are hierarchically organized in time. *Proc Natl Acad Sci* 114:12827–12832.
- Vincent JL, Kahn I, Snyder AZ, Raichle ME, Buckner RL (2008) Evidence for a Frontoparietal Control System Revealed by Intrinsic Functional Connectivity. *J Neurophysiol* 100:3328–3342.
- Vodrahalli K, Chen P-H, Liang Y, Baldassano C, Chen J, Yong E, Honey C, Hasson U, Ramadge P, Norman KA, Arora S (2018) Mapping between fMRI responses to movies and their natural language annotations. *NeuroImage* 180:223–231.
- Wilson SM, Bautista A, McCarron A (2018) Convergence of spoken and written language processing in the superior temporal sulcus. *NeuroImage* 171:62–74.
- Wilson SM, Molnar-Szakacs I, Iacoboni M (2008) Beyond Superior Temporal Cortex: Intersubject Correlations in Narrative Speech Comprehension. *Cereb Cortex* 18:230–242.
- Yang Z, Wu J, Xu L, Deng Z, Tang Y, Gao J, Hu Y, Zhang Y, Qin S, Li C, Wang J (2019) Individualized psychiatric imaging based on inter-subject neural synchronization in movie watching. *NeuroImage*:116227.
- Yeo BT, Krienen FM, Sepulcre J, Mert R, Sabuncu, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL (2011) The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 106:1125–1165.
- Yeshurun Y, Nguyen M, Hasson U (2017a) Amplification of local changes along the timescale processing hierarchy. *Proc Natl Acad Sci* 114:9475–9480.
- Yeshurun Y, Swanson S, Simony E, Chen J, Lazaridi C, Honey CJ, Hasson U (2017b) Same story, different story: the neural representation of interpretive frameworks. *Psychol Sci* 28:307–319.
- Yun K, Watanabe K, Shimojo S (2012) Interpersonal body and neural synchronization as a marker of implicit social interaction. *Sci Rep* 2.

- Zacks JM, Braver TS, Sheridan MA, Donaldson DI, Snyder AZ, Ollinger JM, Buckner RL, Raichle ME (2001) Human brain activity time-locked to perceptual event boundaries. *Nat Neurosci* 4:651–655.
- Zacks JM, Tversky B (2001) Event structure in perception and conception. *Psychol Bull* 127:3–21.
- Zadbood A, Chen J, Leong YC, Norman KA, Hasson U (2017) How We Transmit Memories to Other Brains: Constructing Shared Neural Representations Via Communication. *Cereb Cortex*:1–13.
- Zhang H, Chen C, Zhou X (2012) Neural correlates of numbers and mathematical terms. *NeuroImage* 60:230–240.
- Zheng L, Chen C, Liu W, Long Y, Zhao H, Bai X, Zhang Z, Han Z, Liu L, Guo T, Chen B, Ding G, Lu C (2018) Enhancement of teaching outcome through neural prediction of the students' knowledge state. *Hum Brain Mapp* 39:3046–3057.
- Zhou G, Bourguignon M, Parkkonen L, Hari R (2016) Neural signatures of hand kinematics in leaders vs. followers: A dual-MEG study. *NeuroImage* 125:731–738.
- Zion Golumbic EM, Ding N, Bickel S, Lakatos P, Schevon CA, McKhann GM, Goodman RR, Emerson R, Mehta AD, Simon JZ, Poeppel D, Schroeder CE (2013) Mechanisms underlying selective neuronal tracking of attended speech at a “cocktail party.” *Neuron* 77:980–991.
- Zuo X, Honey CJ, Barense MD, Crombie D, Norman KA, Hasson U, Chen J (2020) Temporal integration of narrative information in a hippocampal amnesic patient. *NeuroImage* 213:116658.

Appendices

Appendix 1

Animate shapes movie

The full shapes animation, titled “When Heider met Simmel,” is available for download from the Supplementary Data in:

Nguyen, Vanderwal, & Hasson (2019). Shared understanding of narratives is correlated with shared neural responses. *NeuroImage*, doi:

<https://doi.org/10.1016/j.neuroimage.2018.09.010>

The animation was directed by Tamara Vanderwal and created by visual artist Tobias Hoffman. The original score was written by Jodi S. van der Woude.

Appendix 2

Theory of Mind audio story

The Theory of Mind audio story was written by author Mai Nguyen in consultation with animation director Tamara Vanderwal. The audios were recorded by voice actor Max Rosmarin.

The full transcript of the story follows below:

“A young boy, carrying a ball, goes to an empty playground to play by himself. The boy kicks the ball and then dashes after it. He lines up his shot and kicks the ball again. The boy catches up to the ball and bounces it off the playground’s fences several times, passing it to himself. As he circles the ball, lining up his next shot, his father shows up and calls for him. The boy takes one final kick at the ball before leaving the playground to meet his waiting father. The boy and his father exchange a few words and leave.

“The boy and his father live in a modest, two story home with a pointed roof. After arriving home, they have a short chat, and the boy turns and climbs the stairs to his second floor bedroom. His father follows, gives his son a goodnight kiss and tucks him in for the night. While the boy falls asleep, his father walks back downstairs to his bedroom on the first floor, climbs into bed, and quickly falls asleep. Upstairs, the boy begins to dream. The ceiling opens up and he floats into the sky, along with a huge bird. High above mountains, the boy can fly! The giant bird splits into two smaller birds, which soar across the sky with the boy. The two birds become four birds. The boy plays with the birds, but then the flock gangs up on him and start chasing him. He tries to fly away but

cannot escape! Thankfully, the dream ends, and the birds disappear. The boy sleeps peacefully for the rest of the night.

“In the morning, the boy’s father climbs the stairs to his son’s bedroom and wakes him up. The boy follows his father back down the stairs to the first floor of their home. They have a brief conversation before the boy heads out for the day. The day passes. The father waits for his son to return home. When he does, the two talk briefly and the boy climbs into bed as does his father downstairs. Both fall right asleep. Shortly after falling asleep, the boy starts to have the bird dream again. The four birds from the previous night are back. They surround him, preventing him from escaping no matter which direction he turns. The birds transform into a giant monster with two long arms and two wobbly legs. The monster swats at the boy! He misses, and tries again. Frightened, the boy flies up and over the monster, evading its grasp. The monster scrunches into a little ball. It pounces, lunging at the boy! After the boy dodges this attack the nightmare finally ends. The boy sleeps peacefully for the rest of the night.

“In the morning, the boy’s dad wakes up, walks to the stairs, and shouts for his son to come down. The boy jumps up, and goes down the stairs. After a brief good morning, the boy again heads out for the day, and the day passes. At the end of the day, the boy returns home, greets his father quickly and then goes straight upstairs to his bedroom. The boy climbs into bed, as does his father downstairs. Both eventually fall asleep. For the third night in a row, the boy begins to dream. The towering monster from the previous dream appears again, slowly following the boy and reaching for him with one arm. The boy stays out of his reach and hovers uncertainly around the monster’s feet. Suddenly, he is joined by another boy. The boys speak frantically as the monster reaches.

The boys fly over the monster, which is still clawing at them. The boys make a desperate plan, and the new boy speeds away. He quickly returns with a ball. The boys put the ball in a slingshot. At the last minute, they let it go, shooting the ball into the monster, which breaks into a heap. Excited, the boy swoops over the heap until he notices that the other boy is gone. He goes looking for his new friend, but can't find him. He checks under the rubble for his friend. He finds the ball, but not the other boy. His friend is nowhere to be seen. Disappointed, the boy abandons the ball and slowly leaves the dream. He sleeps the rest of the night.

“In the morning, the boy's father wakes up and calls for him from the bottom of the stairs. The boy wakes up, gets out of bed, and walks downstairs. Father and son exchange a few words, and the boy heads out. The boy walks to the playground, where his ball is still sitting at the far end of the yard. The boy enters the playground, closing the gate after himself. He forlornly walks to the single ball. He slowly and unenthusiastic kicks the ball halfway down the yard. As he kicks the ball again, his mysterious friend from the dream appears and sneaks towards the playground. He checks that his friend is there. He quietly opens the gate. He sneaks up on the boy, who still doesn't see him. He taps the boy on the shoulder. The boy whirls around! The two friends talk back and forth. The boy races to get his ball as his friend gets in position. The boy kicks the ball hard to his friend, who quickly passes it back, playing together in real life just as they had in the dream.”

Appendix 3

Physical audio story

The Physical audio story was written by author Mai Nguyen in consultation with animation director Tamara Vanderwal. The audios were recorded by voice actor Max Rosmarin.

The full transcript of the story follows below:

“A small triangle overlapping a smaller circle slides into a rectangle. The triangle touches the circle, which glides away. The triangle slides to it and the circle moves away. The triangle comes into contact with the circle again, and the circle ricochets between the side of the rectangle and the triangle several times. The triangle moves in a ring around the circle, rotating. A large triangle slides towards the rectangle. The small triangle slides outside the rectangle until it is adjacent to the large triangle. After bouncing, both triangles slide out of the scene.

“The rectangle re-assembles into a square like structure with a triangular top and a line in the middle. The triangles slide into the structure through a gap in the bottom. The triangles rotate ninety degrees. The small triangle slides diagonally to the upper right, stopping at the middle line. The large triangle does the same. The triangles rotate so their top corners touch, and then the small triangle rotates so the top corner points up. The large triangle slides diagonally down and rotates in place. Both triangles move slightly, up and down, up and down. The top of the structure rotates open, and the small triangle moves slowly upwards while rotating. The structure turns into a large V. The branches of

the V rotate up and down. The V splits into two smaller Vs, and the triangle moves left and right with the Vs. Again, the two Vs split in half, making four Vs. The four Vs move in the same direction as the triangle, which moves left and right. The Vs reforms the original square like structure. The big triangle appears, and the triangles move slightly up and down.

“The big triangle moves diagonally up to the middle line. Adjacent to the little triangle, it moves up and down once. The large triangle moves diagonally down. The small triangle does the same. The small triangle slides down, out of the structure. The large triangle disappears, and the entire structure rotates three hundred sixty degrees. The large triangle appears at the bottom. The small triangle moves into the structure, moves up and down next to the large triangle once, and then slides diagonally up to the middle of the structure. Both triangles rotate ninety degrees, and then bounce slowly in place. The structure’s lower pieces form four V’s that again move up and down. The small triangle moves upward. The Vs are positioned around the triangle, moving with it. The Vs become rigid and form separate bars. Other pieces form four longer lines. The long line on the left has hinges and moves towards the triangle. The triangle moves to the upper right while the lines move down. The lines then move vertically, extending towards the triangle. The moving lines reform the square like structure, and the small triangle moves down to the middle line.

“The large triangle appears, moves left, and then rotates to point diagonally up. It bounces. The little triangle rotates ninety degrees, and slides diagonally down. Both bounce, and the little triangle exits the structure. The large triangle disappears, and the structure rotates 360 degrees. The large triangle appears. The small triangle slides into the

structure, bounces adjacent to the large triangle, and glides diagonally upwards towards the middle line. It rotates until it points up, as does the large triangle below. Both bob slightly up and down. The top of the square like structure opens, and the small triangle moves upwards. The square forms six lines. Two lines have moving parts. The lines move sideways all together, and a long line with hinges swings out towards the triangle. The triangle moves downwards. A small square appears from the left. Two of the lines rotate to point at them. The square and the triangle move counterclockwise around the screen. The square and triangle move up and down. The square leaves and re-enters again with a small circle. The circle is between the triangle and square. The circle moves quickly towards the moving lines. When the circle touches the lines, they collapse into pieces. The triangle moves above the pieces, making a small loop, and then glides off the screen. The pieces are stationary. The triangle reappears from the bottom left of the screen. The top of the triangle touches one of the pieces, which rotates up and back down. The motion is repeated on another piece, and the triangle touches the circle underneath it. The circle and triangle, moving in sync, move upwards and make a full revolution. The triangle moves off screen while the circle and the pieces are stationary. The pieces reform the square like structure.

“The small triangle moves slowly into it from the top. The large triangle appears in the bottom, moves left, and bounces once. The small triangle rotates ninety degrees, glides diagonally down, and bounces next to the large triangle. It slides out of the structure. The structure reconfigures into a rectangle with a gap in the upper left side and a circle in the bottom left. The small triangle slides through the gap. It touches the angled bar, which rotates to close the gap. The triangle slides towards the circle until they are

touching. The circle moves upward and stops. The triangle moves towards the circle, and the circle moves away. A square appears on the left and moves towards the rectangle. The square stops. It touches the bar, which rotates to make a gap. It moves to the triangle's position. The shapes touch and then bounce adjacent to each other. The triangle moves to the circle and the circle moves to the square. The circle moves repeatedly between the square and the triangle.”

Appendix 4

fMRI video lesson

The fMRI video lesson was developed by Mai Nguyen and can be downloaded from <https://youtu.be/RTPqJy-80-M>.

The full transcript of the story follows below:

“Hi! Today, we’re going to be discussing a common form of neuroimaging called fMRI.

We’re going to discuss some fMRI basics, as well as talking about the MRI scanner itself. We’ll then move on to discuss fMRI physics, and we’ll end on the biology of fMRI, or how we go from neurons and neural activity to fMRI images. So let’s get started with some fMRI basics. fMRI stands for functional magnetic resonance imaging, and at the most basic level, it’s simply a method for indirectly measuring brain activity. Now what do I mean by indirectly? Well more precisely, what fMRI does is measure changes in net magnetization, which depend in part on the level of blood oxygenation in your brain, which itself is correlated with neural activity. Now I know this all sounds very indirect and kind of complicated sounding, but don’t worry, but we’re going to go through each of these parts step by step and by the end of the lesson, you should have a pretty good idea of what fMRI is and how it works. So let’s get started.

“As I said before, fMRI is a method that neuroscientists like me use to measure brain activity while people are doing different types of tasks. We like using fMRI for doing this for a couple major reasons. One of the main ones is that fMRI is pretty non-invasive. During an fMRI scan, all you have to do is lie inside a big tube for a while. We

don't use, uh, we don't inject any radioactive dyes or use X-rays or even do brain surgery, so we can basically do fMRI on anyone. A second major advantage is that it has high-spatial resolution. Now Spatial resolution refers to how sharp or detailed of an image we can take. High spatial resolution means that we can take very sharp and detail image, while low spatial resolution means that we can only take blurry images. fMRI has relatively high spatial resolution on the order of one to three millimeters. Many other commonly used methods for neuroimaging are quite a bit worse with resolution on the order of one half to two centimeters or more. Now of course, there are a few downsides to fMRI as well. The main one is that fMRI has low temporal resolution. Now unlike spatial resolution, temporal resolution refers to how quickly we can take images. So something that has low temporal resolution would mean that we can only take a few images over time, while high temporal resolution would mean we can rapidly take many pictures in the same amount of time. So fMRI has relatively low temporal resolution, which means we can only measure your brain's activity every one to three seconds or so. Many commonly used methods for neuroimaging are quite a bit faster, taking images every millisecond. So together, this means that while we can use fMRI can take very sharp images of your brain activity, we can only take those images every few seconds and this means that we're going to miss rapid changes in your brain activity. Going back to the disadvantages fMRI, another major one is that it's actually very expensive to do. A single session can cost hundreds of dollars so doing this kind of experiment can sometimes be very difficult. So that's a bit about fMRI basics.

“Let's move on to talk about the actual fMRI scanner itself. Here's a picture of an fMRI scanning suite. It's actually the one just next door. And as you already know, at the

start of an fMRI scan, you would lie on the scanner bed, and your brain, your head goes inside this weird plastic thing which is actually called radio frequency, or RF, head coil. We're going to talk a lot more about the RF head coil later on, but for now, just know that the RF head coil contains many sensors that go all around your head and do various things like measure certain things in your brain. Now once you're situated on the scanner bed, the entire thing moves inside this tube-like thing, which is actually a very powerful magnet. And even though it's a weird shaped tube, it's a magnet just like any other magnet and so has a direction, or north and South Pole. We're going to indicate the direction of the fMRI magnet as an arrow that points towards the North Pole from now on. It's going to turn out that this direction is going to be really, really important to understanding fMRI physics later on. Now unlike probably most normal magnets you've come across in your life, the MRI magnet is really, really strong. We measure magnetic field strength with the unit called "tesla" and the magnet here that you're in right now is about three Teslas strong. In contrast, the typical fridge magnet is only about point zero, zero one Tesla strong, and so this means the MRI magnet you're in right now is about 3,000 times more powerful than most magnets you've probably come across. You might be wondering why do we need a really powerful magnet? What is it doing? Well to answer the question, we need to get into some of fMRI physics.

"In this next section of the lesson, we're going to try to understand how the three Tesla MRI magnet and the radio frequency, or RF, head coil work together to measure something called net magnetization in your brain. Now later on, we'll explain how net magnetization is related to neural activity in your brain, but for now, let's just try to understand what net magnetization is and how it changes during an fMRI scan. So to start

with, your brain is full of water, which means that it's full of protons. And, you probably already know that protons have a positive charge. What you might not know though is that a proton actually spin about an axis, just like a top spins on a table. Now the interesting thing is that because protons are positively charged, this spinning motion creates something called magnetic moment in the direction of the axis. And all this is just a fancy way of basically saying that these protons that are spinning are like little magnets. And just like any other magnet, they have a north and south pole or a direction, which again I'm going to indicate as an arrow that points towards the North Pole.

“So, if we go back to your brain, and zoom into a little chunk of it, if we could, we would see that little chunk of brain has a ton of protons in it. And those protons are spinning and pulling in totally random directions. Now, we can't actually measure the magnetic pull of any one of these protons. But What we can do though, is measure the average magnetic pull of a whole bunch of protons. And we call this average magnetic pull, or average magnet direction, net magnetization. Visually, you can think of net magnetization as adding up all these little red arrows, arrows that indicate the magnetic pull of each protons. Because red arrows are pointing in random directions, you would see that the arrows all add up to zero .And all this is to say is that under normal circumstances, the net magnetization in any area of the brain is about zero. your brain is generally not magnetized.

“Of course right now is not under normal circumstances because you're inside an fMRI scan. We're putting your brain inside this really, really powerful fMRI magnet, and it turns out the reason we need that fMRI magnet is that it's going to change the direction of those spinning protons. Here's what I mean. Remember that the MR magnet is just like

any other magnet and has north and south pole, which I've indicated with an arrow that points towards the North Pole. Now when I stick your head in the magnet, um, all those protons that used to be pointing in random directions, are going to turn so that they are aligned with the MRI magnet. Most of the protons in your brain are going to turn that they are aligned in the same direction as the MRI magnet, in what's called a parallel, or low-energy state. The remaining fewer protons are going to align so that they're pointing in the opposite direction of the MRI magnet in what's called an anti-parallel, or high energy state. Now remembering anti-parallel and parallel is going to be really important later on, so here's a little mnemonic to help you remember. Parallel protons point in the with the MRI magnet. And because they're going with magnet, they're just going along with the flow, just going along with things doesn't take a lot of energy, so they're in a low energy state. Anti-parallel protons in contrast, are going against the magnet, going against the flow. And of course going against anything takes a lot of energy so anti-parallel protons are in a high energy state. So to remember parallel versus anti-parallel, just remember that parallel protons point with the magnet, so they're low energy, while anti-parallel protons point against the magnet, so they're in high energy state.

“Now if we go back to your brain and we and zoom back in on that little chunk of brain, but now your brain is in the MRI scanner, we see that most of the protons in your brain are aligned in that low energy parallel state. The remaining fewer ones are in that high energy, anti-parallel state. But now because more protons are parallel than anti-parallel state, net magnetization in your brain has also changed. Net magnetization in your brain is no longer zero. It is parallel to the MRI magnet. So now we see how the

MRI magnet has the aligning protons in the parallel or anti parallel state. But because more protons in the parallel than anti-parallel state, net magnetization in your brain changes to be parallel.

“Now what happens when I actually start scanning or collecting data from your brain? Well the start of scanning is a process called excitation where I’m actually going to change the alignment of those protons. Here’s what that looks like. Remember the radio frequency, or RF, head coil contains many sensors that go all around your head. And these sensors do a lot of different things, and one of the things they do is send energy into the brain that causes protons to flip from the low energy state to the high energy state. Remember that when I first stuck your head inside the MRI magnet, most of the protons in your brain turned so they are in the low-energy, parallel state. The remaining fewer protons are in that high energy, anti-parallel state. Because there are more protons in the low energy, parallel state than the high-energy, anti-parallel state, the net magnetization in your brain has changed so is also parallel. Now, this is going to change when I’m actually going to start scanning because I’m going to start a process called excitation, where I use the RF head coil to send RF energy into the brain. And so I’m sending this little burst of energy into the brain, And some of those low-energy parallel protons are going absorb this energy and use it to hop up to the high-energy state. You can kind of think of the process of excitation as like someone drinking some Red Bull. Somehow who is in a low energy state lying around on the couch might drink some Red Bull and get a little burst of energy, and this allows them to hop up and run around in a higher-energy state. In the same way, the excitation process in the MRI scanner send energy into the brain to boost low energy protons into the high-energy state. Now

because we have more protons in the anti-parallel than the low, er, parallel state, you'll notice that net magnetization has reversed directions. It's not anti parallel as well.

“So now we've seen that excitation is the process by which the RF head coil sends pulses of RF energy into the brain, flipping low energy protons to the high-energy state and reversing net magnetization. What happens next? Well the moment I stop exciting the brain, we go into this process called relaxation. which is basically this process. Here what that looks like. So remember during excitation, I sent energy into the brain which allowed some of those low energy protons to hop up into the high energy state. The thing is though that I'm going to excite your brain for but a moment, and the moment I stop exciting your brain, some of those protons in the high-energy state are going slowly lose energy and return back to a low energy state, and this process is called relaxation. And so one of the effects of relaxation is that net magnetization is going to slowly return to each original parallel direction.

“So now we've seen how excitation is the process of sending energy into the brain, flipping low energy protons into the high energy state while relaxation is basically the opposite process by which high energy protons slowly return to the low energy state, restoring the direction of net magnetization. now what happens next? Well what happens next actually happens simultaneously with relaxation, which is that we use the RF head coil to measure net magnetization during relaxation. And the reason we do that is it turns out that relaxation time, and therefore net magnetization, depends in part on the level of neural activity. And we'll talk a lot more about how net magnetization is related to neural activity later, but for now let's just look at how net magnetization changes during the scan in more detail. Here, On the x-axis, I'm going to show time, while the y-axis is

going to be the direction of net magnetization. The MRI magnet as before is going to be pointing up. So, when I first stick your head in the magnet, most of the protons align in the parallel state. while the rest of the protons align anti-parallel state. But because more parallel than anti-parallel protons, the net magnetization in your brain is also parallel. This changes during excitation when I send RF energy into the brain allowing Some of those low energy protons to absorb that energy and hop up into the high-energy state. Now we have more protons in the anti-parallel state than parallel state, net magnetization has also changed and becomes parallel, er, anti-parallel. But only for a moment, because the moment I stop sending energy into the brain, some of those high energy protons are basically going to remember that they're well lazy. They're going to lose that energy and slowly relax back down into their low energy state. And as return to that parallel state, net magnetization is going to slowly drop back down to zero and then back up to is original, parallel state. And it's during this slow relaxation period that we're going to measure net magnetization. and we do this because active and resting brain regions have different relaxation times. Active brain regions have a relatively long relaxation time, while resting brain regions have a relatively short relaxation one. So if we measure net magnetization during relaxation, we're going to see higher net magnetism in active brain regions than resting ones, which ultimately appear as these bright spots in our fMRI images. and that's fMRI physics in a nutshell.

“Let's recap that a little bit. We started out with how your brain is full of protons and how under normal circumstances they're spinning and pointing in random directions. This changes when I stick your head in the MRI magnet though, because the magnet causes those protons to align either in the parallel or anti parallel state . There's more

parallel than anti-parallel though, so net magnetization in your brain changes to parallel. Once I start scanning, though, I'm going to send, I'm going to excite the brain by sending RF energy into the brain, allowing some of those low-energy protons to jump up into the high-energy state, reversing the direction of net magnetization. But only for a moment, because the moment I stop exciting your brain, those protons are going to slowly relax back down into their original low energy parallel state. And as they do this, net magnetization is going to slowly return back down to its original parallel state. and the important thing is that relaxation time in the brain in part on the level of neural activity. So if we measure net magnetization during the relaxation time, we're going to see higher net magnetization in active regions which appear in our fMRI images so you might be wondering: how does net magnetization even relate to neural activity? Why do we care about this?

“To answer that question, we finally get to the biology of fMRI, where we will try to understand how we go from neurons to fMRI. we are going to talk about how we go from a neuron and a neuron's activity to net magnetization and therefore an fMRI image. The path from a neuron to an fMRI image is of course kind of indirect and complicated. But don't worry, just as before, we're going to go through each of these steps one by one. Let's start with a neuron. What does it mean for there to be neural activity? Well during brain activity, what's happening is that these neurons are talking, or signaling, to each other. Neural signaling begins with an electrical signal in a neuron. which is actually called an action potential, and the action potential is mediated by the flow of various charged particles. Now when the action potential reaches the end of the neuron, it triggers the release of neurotransmitters. These neurotransmitters will go on to signal other

neurons and activating them and triggering another action potential. So that's a really broad overview of neural signaling.

“Let's dive deeper into how neurons actually signal to each other. Before we do that, let's go over some neuron structure, just so that we're all on the same page. The big blob here on the left is cell body, and the little branches sticking off of it are called dendrites. Now dendrites are super important because they contain these specialized receptors on them that actually detect signals from other neurons. These signals always appear in the form of neurotransmitters. Now this other long branch off the cell body is called the axon, and the axon is primarily responsible for conducting the electrical signal of the neuron, or the action potential. The end of the axon is called the axon terminal, and when it receives the action, or detects the action potential, it triggers the release of neurotransmitters which go on to signal other neuron. Now the entire neuron is encased in something called the cell membrane. The cell membrane is really really important because it well, keeps stuff from coming in and out of the neuron. It keeps stuff on the inside in and the stuff on the outside, out. And this might sound a really basic, but things the cell membrane does, is it keeps various charged particles in and out of the cell. And as I said before, it's these charged particles are responsible for action potential down a neuron's axon, so it's really, really important that we understand what they are, where they are, and how they move around during neural signaling.

“So let's get started by looking at the neuron at rest. So when the neuron is at rest, not doing anything in particular, it has a higher concentration of sodium on the outside of the neuron. It has a higher concentration of potassium on the inside, as well as a higher concentration of various negatively charged proteins. Now keeping track of these three

things is going to be really really important, so here's a little trick to help you remember. A neuron at rest is like a salty banana. Now, I know this sounds really stupid, but bear with me for a moment. You probably know that bananas are really high in potassium, just like a neurons are. And you probably know that salt is usually made up of sodium, and so a salty banana, is surrounded by sodium, in the same way that a neuron is surrounded by a high concentration of sodium. So if you remember the phrase "salty banana", the only other thing you have to remember is that the inside of the neuron has negatively charged proteins. And the lucky thing is that these negatively charged proteins are never going to move around, so you just have to remember that they're on the inside. So I know that was kind of a stupid analogy, but hopefully it'll help you remember sodium, potassium, and these negatively charged proteins.

"And the reason it's really important to keep track of these are, is that these different concentrations of these charged particles is actually responsible for differences in electrical charge in out of the neuron because the inside of the neuron has a really high concentration of negatively charged proteins, it's actually more negatively charged. The outside is more positively charged because of all of that sodium on the outside. Now I don't have a super great way to help you remember this, but if you remember the "salty banana," you'll remember that there's sodium on the outside and that sodium is positively charged, so the outside of the neuron is also positively charged. Now the inside of anything is of course the opposite of outside, so the inside of the neuron is opposite, and it's negatively charged. Going back to a real neuron, let's zoom in on the cell membrane and take a closer look at what's happening. As we just saw, the outside of the neuron has a higher concentration of sodium, so it's more positively charged. The inside of the

neuron has more potassium but also a lot of that negatively charge protein, so it's more negatively charged on the inside.

“Now the important thing is that difference in charge across the membrane creates a something called an electrochemical gradient. And the thing about electrochemical gradients is that charged particles don't really want them around. they want to get rid of those electrochemical gradients by moving around. In this case, the positively charged sodium ions want to enter the negatively charged inside of the neuron, in order to move that positive charge around so that the cell reaches equilibrium: or equal charge in and out of the cell. The problem is that sodium can't enter to the cell because of the cell membrane. And so we see how the cell membrane is really, really important for maintaining the electrochemical gradient across the neurons, across the neuron.

“Alright, so we now know what the neuron looks like at rest. What happens when a neuron is stimulated? well we saw before that neurons are activated by neurotransmitters that are released by other neurons. in more detail, though, this signaling is occurring across this gap between neurons called the synapse, And so the neuron that sends the signal is called the presynaptic neuron while the neuron that receives the signal is called the postsynaptic neuron. And so in more detail, neural signaling actually starts when the presynaptic neuron has an action potential. which reaches the end of the axon terminal and triggers the release of neurotransmitters. These neurotransmitters cross the synapse until they reach the post synaptic neuron, where they bind to receptors on dendrites. Let's take a closer look at what happens. Remember that when the neuron is at rest, the different concentration of ions in and out of the cell, creating an electrochemical gradient, positive on the outside, negative on the inside. The neuron wants to eliminate

this uh charge by letting sodium into the cell, down its electrochemical gradient to spread its positive charge around, but it can't because it's blocked by the cell membrane.

“Now what I didn't tell you before is that the cell membrane actually has these specialized gates called ion channels that let ions through. Now most of the time these ion channels are closed, but when dendrites detect the present of neurotransmitters, these ion channels will open. So when this happens, at first only the sodium ion channels opens. This allows sodium to travel down its electrochemical gradient and enter the neuron. notice that when this does this, the electrical charge across the cell membrane has changed. the outside of the cell is now more negative, while the inside is more positive. Moments later, the potassium ion channels are going to open, allowing potassium travel down its electrochemical gradient and exit the cell. Notice that when this happens, the charge across the cell membrane has once again reversed: the outside is more positive again while the inside is more negative. So we can kind of think of this process of the action potential as a neuron become an inverted salty banana. Now the inside of the banana is salty, or full of sodium, while the outside is full of potassium. So you can kind of think of neural signaling as the process of going from a salty banana to an inverted salty banana.

“So to go back to real neurons and zoom back out again, we now see how neurotransmitters bind to receptors on dendrites and how this opens various ion channels. this allows sodium to travel into the neuron and for potassium to exit the neuron. And the important thing is that the movement of these charged particles changes the electrical charge across the cell membrane. and it's actually this change in electrochemical gradients that triggers the action potential along the axon. And when the action potential

reaches the axon terminal, it triggers the release of neurotransmitters from the axon terminal, going on to signal other neurons.

“So now we know a bit about neural signaling. What I’m going to show you next is that recovering, or resetting, from neural firing, is going to increase energy demands in active areas of the brain. Here’s why. After an action potential, neurons can’t fire for a while: it’s electrochemical gradients are all messed up. The outside has higher concentrations of potassium and the inside has a higher concentration of sodium. before the neuron can fire again, we need to reset the neuron: we need to get sodium back out and potassium back in. In other words, we need to go back from the inverted salty banana back to the salty banana, and the thing is that this process is going to require a lot of energy. Here’s why. Remember that during electrical, uh, during action potential, ions travel down an electrochemical gradients. What I didn’t tell you before is that no energy needed because ions are traveling down an electrochemical gradient. You can kind of think of it as a ball rolling down a hill. You don’t need to expend energy in order to move that ball down the hill because gravity is just going to pull it down.: In the same way, that negative charge inside the, the neuron is pulling that positive charge inwards, and so you don’t need any energy to move that sodium ion in.

“Now the situation is very different after an action potential. After an action potential, we need to get to get potassium back into the neuron and sodium back out of the neuron. And this process is going to require energy because now we’re moving sodium up an electrochemical gradient. And this process is going to require energy. To go back to the ball on a hill analogy, a ball might roll down a hill on its own because of gravity, but It’s not going to move up on its own, without someone expending energy to

push it back up. In the same way, sodium's not going to move up against its electrochemical gradient into a positively charge without something actively moving it, and that something is going to need a lot of energy in order to do it. now the thing that's actually going to move sodium and potassium up against its electrochemical gradient is called the sodium potassium pump. And the sodium potassium pump is a bit like an ion channel in that they're a way that sodium and potassium to cross the cell membrane. There's a really key difference though, which is that ion channels work by allowing ions to travel passively down their electrochemical gradients. Rolling down that hill. This doesn't require any energy and so these ions move without requiring any energy. Now the sodium potassium pump in contrast, is a way for ions to move up against their electrochemical gradient. But because they are moving up that hill, it requires a lot of energy to work. So while an action potential itself does not require energy because ions are flowing in and out of the neuron using ion channels flowing down electrochemical gradients, resetting after an action potential requires a lot of energy in order to move those ions in and out of the using the sodium potassium pump up against their electrochemical gradients.

“So now we've seen something about neural activity is going to lead to increased energy demands in active brain regions. How does the brain meet these increased energy demands? It turns out by increasing oxygenated blood flow to active brain regions. Let's say that this is your brain. And right now areas of your brain that are involved in learning and memory and complicated things are move active right now because you're listening to this lecture.: so these more active areas are going to have greater energy demands. And to meet these energy demands, they're going to need two things: glucose, and oxygen.

And it can only get these two things through the circulatory system. And so what the brain actually does is increase oxygenated blood flow to, to active brain regions in order to meet the area's increased energy demands. The thing is though that your brain is going to way, way, way over does it. It's going to increase increases oxygenated blood flow to such an extent that the oxygen supplied by increased blood flow is actually going to exceed the amount of oxygen consumed by increased neural activity. And so counter intuitively, even though these active brain regions are consuming more oxygen, they actually have higher levels of blood oxygenation. It's actually the resting, or less active, brain regions that going to have higher levels of deoxygenated blood.

“So we see how the brain increases oxygenated blood flow in order to meet energy demands. This causes blood oxygenation levels to increase in active brain regions. And this is awfully lucky for us because it turns out that it's the blood oxygenation level that's going to lead to that difference in net magnetization in active versus resting brain regions. Remember that all of this is happening inside the really, really powerful MRI magnet. And the interesting thing is that oxygenated blood could not care less. It does not interact with the magnetic field of the magnet at all. It's deoxygenated blood that interacts with it, distorting the magnetic field. And so wherever there's deoxygenated blood, there's going to be distortion of the magnetic field, which weakens net magnetization and shortens relaxation time. So this means that more active areas of the brain have higher levels of blood oxygenation, and so they're going to have less distortion. This is going to result in higher levels of net magnetization, which we can measure with the radio frequency head coil. And so areas that have higher net

magnetization are also going to have higher levels of fMRI signals, which appear as these bright images in our fMRI images. And that's fMRI in a nutshell.

“Let's recap that real quick. We started out with a neuron and we learned how a neuron's activity is caused by changes in electrochemical gradients across the cell membrane. We then learned how this increases energy demands in order to reset the electrochemical gradients of these neurons. In order to meet this increased energy demands, the brain increased oxygenated blood flow to active brain areas. it does so by such an extent though that the blood oxygenation level actually increases in active brain regions. This is very fortunate for us, because oxygenated and deoxygenated blood interact with the magnetic field differently. Oxygenated blood does not interact with it while deoxygenated blood distorts it, lowering net magnetization, and therefore, lowering net magnetization. We then measure net magnetization in the brain using the RF head coils, and we're going to see higher net magnetization in active brain regions producing our fMRI images. Thank you!

Appendix 5

fMRI video review

The fMRI video lesson was developed by Mai Nguyen and can be downloaded from <https://youtu.be/f1-CoySwWqI>.

The full transcript of the story follows below:

“Hi, today we’re going to briefly review a common neuroimaging method called fMRI. We’ll talk about some fMRI basics as well as some fMRI physics, before ending today on the biology of fMRI or how we get from neurons to fMRI images. So let’s get started with some fMRI basics. fMRI stands for functional magnetic resonance imaging and it’s a method for measuring your brain activity. fMRI has a couple major advantages over other neuroimaging methods. fMRI is non-invasive and it has high spatial resolution. On the downside though, fMRI has low temporal resolution and it’s very expensive to operate. So that’s a bit about fMRI.

“Let’s move on to some fMRI physics. When I first stick your head inside the fMRI magnet, the fMRI magnet causes protons in your brain to turn so that they’re aligned with the MRI magnet. More protons align in the anti, er, more protons align in the parallel state than in the anti-parallel state, so net magnetization in your brain changes to be parallel as well to the MRI magnet. Now this changes during excitation, which is the next step of fMRI scanning. During the scan, the RF head coil sends a little pulse of RF energy into the brain every few seconds. Some of those low energy, parallel protons absorb this extra burst of energy, and uses it to hop up in to the high energy, antiparallel

state. Now we have more protons in the anti-parallel state than the parallel state, so one effect of excitation is to change the direction of net magnetization so it's anti-parallel as well. Once I stop exciting the brain though, some of these high-energy protons are going to slowly lose that energy and relax back down over time into their original low energy parallel state. This is going to cause net magnetization to slowly return back to its original parallel state as well. And the reason we care about this relaxation process is that active and resting brain regions have different relaxation times. Active brain regions have a relatively long relaxation time while resting brain regions have a relatively short relaxation time. And so if we measure net magnetization with the RF head coil during the relaxation period, we're going to see higher net magnetization in active brain regions which appear as these bright spots in our fMRI images.

“So, then you might be wondering: how does net magnetization relate to neural activity? Well to answer that question, we're going to get into the biology of fMRI or how we get from neurons to fMRI images. Let's start with a neuron at rest. When the neuron is at rest, and it's not anything in particular, it has a different concentration of charged ions on the inside and the outside. The outside of the neuron has a high concentration of sodium ions, while the inside has a high concentration of potassium and various negatively charged proteins. Now because the neuron has a different concentration of charged particles on the inside and outside, it also has different electrical charge on the inside and outside. Specifically, the outside has a more positive charge because of all the positively charged sodium, while the inside of the neuron is more negatively charged because of all the negatively charged proteins.

“Now this neuron that’s at rest is going to become active when the presynaptic neuron has an action potential down its axon, which triggers the release of neurotransmitters from its axon terminal. The neurotransmitters cross the synapse until they reach the post synaptic neuron, where they will bind to receptors on various dendrites. Now when this happens, the neurotransmitters are going to trigger various ion channels on the cell membrane of the neuron to open. At first, only the sodium ion channels opens, which allows sodium to travel down its electrochemical gradient into the neuron. Notice that when this happens the electrical charge across the cell membrane has changed. Now it’s more negative inside and more positive outside. Moments later, the potassium ion channels open allowing potassium to travel down its electrochemical gradient and exit the neuron. When this happens, the charge once again changes: the outside is more positive again while the inside is more negative again. It’s actually this change in the electrical charge across the cell membrane that’s going to trigger the action potential of the neuron. And when the action potential reaches the axon terminal, it triggers neurotransmitters which go on to signal other neurons. And that’s how a neuron fires.

“Now, the thing is, after a neuron has fired, or has an action potential, it can’t fire again for a while. It needs to reset its electrochemical gradients using the sodium potassium pump. Using the sodium potassium pump is going to require a lot of energy because it actively transporting sodium and potassium against their electrochemical gradients. And so active brain regions are going to have increased energy demands, which the brain is going to meet by increasing oxygenated blood flow to active brain regions in order to bring more glucose and oxygen which the brain consumes for energy.

Now brain is going to overdo this by such an extent though that the oxygen supplied by increased blood flow actually exceed the amount of oxygen consumed by increased neural activity. Which means that active brain regions are actually going to have increased blood oxygenation levels compared to resting brain regions. And this is important to us because oxygenated and deoxygenated blood interact differently with the magnetic field of the MRI scanner. Oxygenated blood does not interact with the magnetic field while deoxygenated blood distorts it, lowering it wherever it is. And so what this means is that active areas of the brain which have higher levels of blood oxygenation are going to have less distortion of the magnetic field, and so are going to have longer relaxation times. This is going to result in higher levels of net magnetization. which is measured by the RF head coil, producing these fMRI images.”

Appendix 6

fMRI quiz questions

1. What is fMRI?
 - a. A method for directly measuring neural activity based on changes in electrical signals
 - b. A method for indirectly measuring neural activity based on changes in blood oxygenation**
 - c. A method for directly measuring neural anatomy based on changes in neurotransmitter concentration
 - d. A method for indirectly measuring neural anatomy based on changes in glucose levels

2. fMRI has ____ spatial resolution, which means it ____.
 - a. High; can precisely measure the location of neural activity**
 - b. Low; cannot precisely measure the location of neural
 - c. High; can detect fast changes in neural activity
 - d. Low; cannot detect fast changes in neural activity

3. fMRI has ____ temporal resolution, which means it ____.
 - a. High; can precisely measure the location of neural activity
 - b. Low; cannot precisely measure the location of neural
 - c. High; can detect fast changes in neural activity

- d. Low; cannot detect fast changes in neural activity**
4. Compared to other neuroimaging methods, what is an advantage of fMRI?
- a. It is cheap to run.
 - b. It requires very little equipment.
 - c. It measures neural activity at the level of a neuron.
 - d. It is non-invasive**
5. The strength of the MR magnet is approximately ____ times the strength of the typical fridge magnet.
- a. 30
 - b. 300
 - c. 3,000**
 - d. 30,000
6. During fMRI scanning, the MRI magnet ____.
- a. Aligns protons parallel or anti-parallel to the MRI magnet**
 - b. Randomizes proton spin direction
 - c. Increases density of proton spin in brain
 - d. Knocks proton spin out of synchrony
7. Net magnetization is ____.
- a. Electromagnetic energy needed to stimulate active neurons

- b. Electromagnetic energy consumed by proton spin
 - c. Total magnetic strength of the MRI magnet and RF head coil
 - d. Average magnetic pull of a group of protons**

- 8. During excitation, the RF headcoil sends pulses of RF energy into the brain, which _____.
 - a. Demagnetizes the protons in resting brain regions
 - b. Boosts protons from the low-energy state up to the high-energy state**
 - c. Stimulates protons in active brain regions so they release magnetic moments
 - d. Causes increased oxygenated blood flow

- 9. During relaxation _____.
 - a. Anti-parallel protons return to parallel state**
 - b. Oxygenated blood flow falls back to normal levels
 - c. The MRI magnet briefly turns off
 - d. Net magnetization becomes random

- 10. Neural activity affects ____, leading to differences in ____ between active and resting brain regions.
 - a. The speed of proton spin; net magnetization
 - b. Proton relaxation time; net magnetization**
 - c. The proportion of parallel protons; blood oxygenation

- d. Proton excitation; blood oxygenation

11. Which part of the neuron is indicated by the arrow?

- a. Dendrite**
- b. Neurotransmitter vesicle
- c. Myelin sheath
- d. Potassium transporter

12. What is the role of dendrite receptors in neural signaling?

- a. Release neurotransmitters
- b. Stimulate vesicle formation
- c. Establish an electrochemical gradient
- d. Detect signals from other neurons**

13. What is the role of the potassium-sodium pump in neural signaling?

- a. Detect the concentration of different ions
- b. Reset the electrochemical gradient**
- c. Close ion channels
- d. Move protons out of the neuron

14. What is the role of the axon in neural signaling?

- a. Physically connect signaling neurons
- b. Propagate the action potential**

- c. Store surplus glucose for later consumption
 - d. Synthesize neurotransmitters
15. BEFORE an action potential, there is greater concentration of ____ INSIDE the neuron and greater concentration of ____ OUTSIDE the neuron.
- a. **K⁺; Na⁺**
 - b. Na⁺, K⁺
 - c. K⁺; neurotransmitter
 - d. Neurotransmitter; K⁺
16. At REST, the INSIDE of the neuron is more ____ charged. The OUTSIDE of the neuron is more ____ charged.
- a. **Negatively; positively**
 - b. Positively; negatively
 - c. Positively; positively
 - d. Negatively; negatively
17. The neuron that sends a signal is called the ____ neuron, while the neuron that detects the signal is called the ____ neuron.
- a. Synaptic; resting
 - b. Active; inactive
 - c. **Presynaptic; postsynaptic**
 - d. Preposition; postpositioned

18. The movement of ions in and out of the neuron during an action potential does not require energy because _____.

- a. They are moving down the electrochemical gradient.**
- b. They are using ion pumps on the cell membrane
- c. They are balanced by the release of neurotransmitters
- d. They are stimulating increased blood flow to the neuron

19. When neurotransmitters bind to receptors on dendrites, _____.

- a. Ion channels open**
- b. The sodium-potassium pump activates
- c. The cell membrane closes
- d. Oxygen transport begins

20. AFTER an action potential, there is greater concentration of ____ INSIDE the neuron and _____ OUTSIDE the neuron.

- a. K^+ ; Na^+
- b. Na^+ ; K^+**
- c. K^+ ; neurotransmitter
- d. Neurotransmitter; K^+

21. Substantial energy is consumed by the neuron when it _____.

- a. Stimulates increased blood flow before an action potential

- b. Resets ion gradients following an action potential**
- c. Releases neurotransmitters during an action potential
- d. Repairs the cell membrane after an action potential

22. Why does blood flow increase to active regions of the brain?

- a. To replenish stores of neurotransmitters
- b. To provide a source of extracellular sodium
- c. To facilitate the exchange of charged ion
- d. To meet increased energy demands**

23. Blood oxygenation levels are ____ in active brain regions compared to resting brain regions because ____ in active brain regions.

- a. Lower; oxygen supply is depleted by increased oxygen consumption
- b. The same; oxygen supply is carefully regulated to match oxygen consumption
- c. Higher; oxygen supply surpasses oxygen consumption**
- d. Initially higher; oxygen supply briefly surpasses oxygen consumption before consumption increases

24. How do oxygenated and deoxygenated blood affect the magnetic field?

- a. Oxygenated blood does not interact with the magnetic field; deoxygenated blood distorts it**

- b. Oxygenated blood distorts the magnetic field; deoxygenated blood does not interact with it
- c. Both oxygenated and deoxygenated blood distort the magnetic field
- d. Neither oxygenated or deoxygenated blood interacts with the magnetic field

25. What does fMRI measure?

- a. Changes in electrical currents from the movement of sodium and potassium
- b. Changes in electrical currents from the consumption of oxygen
- c. Changes in local magnetization from the synthesis of neurotransmitters
- d. Changes in local magnetization from changes in blood oxygenation level**