Communication Neuroscience: Theory, Methodology and Experimental Approaches

Ralf Schmälzle & Dar Meshi

To cite this article: Ralf Schmälzle & Dar Meshi (2020): Communication Neuroscience: Theory, Methodology and Experimental Approaches, Communication Methods and Measures, DOI: 10.1080/19312458.2019.1708283

To link to this article: https://doi.org/10.1080/19312458.2019.1708283

Published online: 28 Jan 2020.
Communication Neuroscience: Theory, Methodology and Experimental Approaches

Ralf Schmälzle 1,* and Dar Meshi 2,*

1Department of Communication, Michigan State University, East Lansing, USA; 2Department of Advertising and Public Relations, Michigan State University, East Lansing, USA

ABSTRACT

The human brain is our primary biological organ of communication. The brain acts as both the sender and receiver of messages and underpins our fundamental ability to communicate and interact with others. Communication scholars can, therefore, study the brain to gain a more complete understanding of communication phenomena. Our goal with the present manuscript is to promote neuroscience research to communication scholars in the following ways: (1) We provide rationale for studying communication from a neural perspective. (2) We delineate the various advantages and challenges that neuroscience methods present. (3) We describe three distinct methodological entry points for communication scholars to approach the field. Specifically, we illustrate how neuroscience measures can be incorporated into communication research as dependent variables, mediators, or predictors. We then close with a forward-looking perspective on future developments in measurement, analysis, and theory, which we expect will have a profound influence on communication science.

The human brain is our primary biological organ of communication (Weber, Eden, Huskey, Mangus, & Falk, 2015; Weber, Sherry, & Mathiak, 2009). The brain acts as both the sender and receiver of messages and underpins our fundamental ability to communicate and interact with others. Communication scholars can, therefore, study the brain to gain a better understanding of communication phenomena (Boster & Sherry, 2010; Sherry, 2004). Historically, there has been momentum for this approach when investigating speech, hearing, nonverbal communication, and media (Kempter & Bente, 2004; Lang, 2014; Reeves et al., 1985). More recently, with the advent of new methods to image the brain (see neuroimaging below), there has been a strong movement promoting a neuroscientific approach across the field of communication research. Several papers which capitalized on these neuroimaging methods have appeared in communication journals (for review see Weber, Fisher, Hopp, & Lonergan, 2018), and about twice as many have appeared in psychology journals with topics that focus on core communication questions (for review see Falk & Scholz, 2017; Meshi, Tamir, & Heekeren, 2015). The annual number of these publications on communication topics is increasing, and in 2016, under the leadership of René Weber, an interest group for Communication Science and Biology (CSaB) was founded within the International Communication Association. This growth trend is also reflected in the strategic hiring decisions of leading communication departments around the world, where over 10 professors with a neuroimaging research agenda have been hired in the past 5 years. Thus, while measuring brain activity has not been considered a classical research tool of the communication discipline (Craig, 1999), we are witnessing the birth of a new approach to communication inquiry: Communication Neuroscience.
Importantly though, the neuroscience theory and methods used to investigate human communication are complex, and this can lead to miscommunication about what neuroscience research can and cannot achieve. Our goal here is to provide a theoretical perspective on the role of neuroscience research in communication science and to communicate the benefits and barriers to conducting this research in a systematic way (DeAndrea & Holbert, 2017). We advocate for communication scholars to embrace this new paradigm to describe, explain, and predict communication phenomena, but with the caveat that the contribution of neuroimaging may come in different forms than what newcomers might expect.

Rationale for Studying Communication from a Neural Perspective

The emergence of Communication Neuroscience is part of a larger development of interdisciplinary science in the 21st century. This development is characterized by (1) combinations of historically disconnected disciplines, across hierarchical levels of explanations (e.g., Churchland & Sejnowski, 1990; Cacioppo, Berntson, et al., 2000), (2) refined measurement and appreciation of the resulting increase in complexity (e.g., Barabási, 2012; Weber et al., 2009), and (3) an emphasis on mechanistic explanation as opposed to covering-law-based approaches (e.g., Craver, 2007). The field of biochemistry provides an exemplar that illustrates this development for communication scholars. Before biochemistry came into being, biologists interested in the principles of “living matter” often invoked constructs like vitalism, which posits that living phenomena cannot be fully explained by the principles of basic sciences like physics and chemistry (Greco, 2005). Vitalism and its relatives ultimately did not provide solid explanations. The early chemists, however, could not explain these phenomena either because the basic elements they knew from the periodic system proved to be too simplistic. It was only when the amino acids and their role in DNA were discovered that the link between the two fields began to make sense: theories became aligned, they made extremely precise predictions, practical control of compounds became possible, and industrial applications flourished. Viewed from today’s perspective, we can see that the early biologists had approached the problem from a level which could be described as “too high”, whereas the early chemists had entered the field on a level which could be described as “too low” (Bechtel, 2008) – one can imagine these two fields existing on horizontal planes, with a gap in between. Importantly, however, the new field of biochemistry offered the right level of granularity to explain the mechanisms governing the behavior of organic molecules. Therefore, biochemistry filled the gap between the two planes. At the same time, both biology (e.g., systems biology) and chemistry (e.g., quantum chemistry) continued to exist and thrive.

The field of cognitive neuroscience is also a merger between two fields, neuroscience and psychology. Cognitive neuroscience is currently on a trajectory similar to biochemistry in the 20th century. For example, the field has established independence from its source disciplines and it has formed a new research community along with specialized journals, conferences, and recognition by funding bodies (e.g., a dedicated NSF Cognitive Neuroscience program). Cognitive neuroscience advocates for a theoretical middle-ground between neuroscience and psychology by “elevating” the neuroscientists’ focus from single neurons and “lowering” the psychologists’ focus on hypothetical constructs that have always remained somewhat mentalistic in nature (Slaney & Racine, 2013). We note, however, that this endeavor proves more challenging than the illustrative example of biochemistry because it cuts right to the core of the mind-body problem, which has caused many philosophical controversies. The mind-body problem consists of debate on whether (or the extent to which) biological substrates enable or comprise the human mind (Cummins, 2000; McGinn, 1989). Progress made to date in cognitive neuroscience has already substantially increased our understanding of cognitive phenomena (Engel, 2008; Mather, Cacioppo, & Kanwisher, 2013). More broadly, the theoretical confluence between neuroscience and cognitive science/psychology proves as one of the most fertile areas of scientific progress these days (Hassabis, Kumaran, Summerfield, & Botvinick, 2017).

These developments in cognitive neuroscience, and its two closest allies – social and affective neuroscience – have implications for communication or at the very least for communication
subfields that make claims about mental processes. Although the quirks of the mind-body problem make it sometimes tempting to view the realm of mental and social phenomena as separate from the physical, chemical, and biological domains, there is overwhelming evidence that this “dualism” has led science astray (Dennett, 1991, 2017), and that the complex social phenomena that we aim to explain in communication science are rooted in and affected by biological mechanisms that have been strongly shaped by evolution. This includes attachment and procreation behaviors (Harlow, 1958; Insel, 2010); nonverbal interaction patterns; information-sharing systems (from birdsong to language; Ghazanfar, 2002); social perceptions, inferences, and preferences (face reading, theory of mind, impressions of attractiveness; Lieberman, 2010); moral systems (Wright, 1995); and hierarchical social structures (Tooby & Cosmides, 1992; Zerubavel, Bearman, Weber, & Ochsner, 2015), to name only a few. With this biological perspective in mind, truly understanding human communication requires us to study it from a multi-level perspective that jointly focuses on the biological processes, that is the brain activity, as well as the social behaviors they produce (Adolphs, 2009; Brothers, 1990; Cacioppo, Berntson, Sheridan, & McClintock, 2000; Cappella, 1996; Krakauer, Ghazanfar, Gomez-Marín, Maclver, & Poeppel, 2017; Ochsner & Lieberman, 2001; Stanley & Adolphs, 2013).

Neuroimaging: Capturing Communication’s Hidden Processes

The term “neuroimaging” describes the measurement of brain activity. With neuroimaging, we can look into the “black box” of the human brain (Bolls, Weber, Lang, & Potter, 2019; Geiger & Newhagen, 1993; Schramm, 1971; Weber et al., 2009) to investigate theories about communication phenomena. Different neuroimaging methods can be used to measure brain activity, such as electroencephalography (EEG), magnetoencephalography (MEG), functional near-infrared spectroscopy (fNIRS), and functional magnetic resonance imaging (fMRI). The specific details of how these methods assess neural activity are beyond the scope of this article and are discussed extensively elsewhere (Baars & Gage, 2012; Gazzaniga, 2009; Huettel, 2008; Luck, 2005). What all these methods have in common, however, is that they take moment-to-moment measurements simultaneously across different brain sites. Hence, they provide spatiotemporal data that are able to reveal which brain regions are active, or the interactions between regions, at specific timepoints. In addition, it is important to mention that different methods can measure different phenomena related to the firing of neurons in the brain – some methods directly measure electrical brain activity, while other methods indirectly measure brain activity. For example, the actual electrical activity of firing neurons is recorded with EEG and MEG (Biasiucci, Franceschillo, & Murray, 2019; Gross, 2019). Alternately, changes in blood flow around firing neurons are recorded with fNIRS and fMRI (Huettel, 2008; Logothethis, 2008). Regardless of the measurement, the result is similar, each of these techniques creates an “image” of the brain activity that it is recording – hence, the term neuroimaging is used. In the discussions below, we focus mainly on fMRI because it has been the most widely used method for recent breakthroughs, but the general principles we discuss extend to other methods.

Neural measures have specific advantages over more established behavioral methodologies in communication research (but also limitations that we will describe below as well). First, this type of data collection does not require the explicit questioning of a participant. Therefore, the measurement itself does not interfere with the thought process that is measured, for example, by (1) interrupting the experience, (2) by evoking specific thoughts through the questioning itself, or (3) by requiring attention allocation for answering questions. Second, neuroimaging measures are process-based biological signals that can be recorded continuously. This temporal dimension is indispensable for studying communication processes but is often absent with other methods (Poole, 2013). Third, just as neuroimaging circumvents the need for overt questions, the “answers” (i.e., the recordings from individual brain regions) are not verbally filtered either. This feature is critical for studying processes that are detached from the language system, which is the case for most phenomena related to emotion, motivation, and other implicit processes (Lieberman, 2000). As such, neural measures can
better bypass social desirability bias and are not subject to retrospective distortion, which are both major concerns in self-report (Stone, Bachrach, Jobe, Kurtzman, & Cain, 1999). It is important to note though, that measurement with fMRI can introduce confounds as well. For example, the noise of the fMRI machine, or the relatively snug space within the machine, may induce cognition that researchers do not intend to measure. Overall, however, neuroimaging has many desirable measurement characteristics and thus holds potential to advance the discipline as a whole (Greenwald, 2012).

Finally, it is also important to note how current neuroimaging methods are distinct from classic psychophysiological methods. Psychophysiological methods measure physiological responses (e.g., heart rate) that result from nervous system activity (Potter & Bolls, 2011), and these methods also possess the three, above-described advantages of neuroimaging methods (Lang, Bradley, Chung, & Lee, 2003; Lang, Potter, & Bolls, 2008). Psychophysiological methods have a long history in communication research, promoting a biological and information-processing approach to our discipline (Kemper & Bente, 2004; Lang, 2014; Potter & Bolls, 2011; Ravaja, 2004). Neuroimaging can be viewed as an extension of this endeavor. Importantly, however, neuroimaging reveals the location and intensity of brain activity. Therefore, neuroimaging can directly expose the inner workings of the brain as it is responding to or producing communication. In this way, neuroimaging methods are able to capture communication phenomena where they arise.

**Extant Neuroimaging Research into Communication Phenomena**

Neuroimaging research has already substantially contributed to our understanding of communication phenomena. Here, we provide a very brief and selective overview of areas in which contributions have been made. Of note, most of this research connects the planes of the communication and neuroscience fields, and the resulting studies were published in neuroscience and psychology journals, beyond the surveillance of many communication scholars. First, research has revealed how the brain perceives when others display or send out social communications. Since the seminal discovery of a brain area that is sensitive to faces – the so-called fusiform face area (Kanwisher, 2017; Kanwisher, McDermott, & Chun, 1997) – work on processing faces, facial expressions and emotion, body posture, biological motion and actions, touch and many other static and dynamic nonverbal social signals has rapidly expanded. Overall, this has led to substantial knowledge about the brain regions involved in social perception (Jack & Schyns, 2017; Lieberman, 2010, 2013). Next, in parallel work, neuroscientists have explored the minutiae of speech perception and production, including the socially important functions of voice and prosody detection (Ethofer, Van De Ville, Scherer, & Vuilleumier, 2009; Hensel, Bzdok, Müller, Zilles, & Eickhoff, 2015; Hickok & Poeppel, 2016; Petkov et al., 2008). Third, neuroscientists have focused on understanding how we think about and respond to others while we interact. Termed “social cognition” in the literature, research has revealed brain regions, such as the temporoparietal junction, involved in making inferences about others’ mental states (Alcalá-López et al., 2017; Saxe & Kanwisher, 2003; Schmälzle, Brook O’Donnell, et al., 2017; Schurz, Radua, Aichhorn, Richlan, & Perner, 2014), as well as when we have empathy for others, or make social and moral judgments (Stanley & Adolphs, 2013), or observe others’ actions (Caspers, Zilles, Laird, & Eickhoff, 2010; Schilbach et al., 2013). Several social neuroscience studies have also been conducted to understand social-affective phenomena such as social exclusion, self-related processing, social comparison, or social reward (Doré & Ochsner, 2015; Eisenberger, Lieberman, & Williams, 2003; Fliessbach et al., 2007; Mesi, Morawetz, & Heekeren, 2013; Northoff, 2014; Saxe & Kanwisher, 2003; Schmälzle, Brook O’Donnell, et al., 2017). Cutting-edge work has also been done in perhaps one of the most central areas of communication research – interpersonal communication. A recent technological advancement allows researchers to record neuroimaging data from two brains during on-going interaction, making it possible to study the dyadic interaction between two physically separated brains. Such research has revealed commonalities in brain activity between senders and receivers during successful message transmission (Babiloni & Astolfi, 2014; Schilbach et al., 2013; Stephens, Silbert, & Hasson, 2014).
2010). These studies demonstrate that profound neuroscience research examining communication phenomena has already taken place – although most have been labeled as “social neuroscience”.

While the above examples were primarily conducted outside the communication discipline, an increasing body of research shows how neuroimaging can be applied within core communication contexts. For example, researchers have examined the reception and processing of health and risk communication messages (Cascio, Dal Cin, & Falk, 2013; Cooper, Bassett, & Falk, 2017; Imhof, Schmälzle, Renner, & Schupp, 2017; Langleben et al., 2009; Ramsay, Yzer, Luciana, Vohs, & MacDonald, 2013; Schmälzle, Häcker, Renner, Honey, & Schupp, 2013; Schmälzle, Renner, & Schupp, 2017; Wang et al., 2013; Weber, Huskey, Mangus, Westcott-Baker, & Turner, 2015), as well as responses to mass media, political communication, and computer-mediated communication (Huskey, Craighead, Miller, & Weber, 2018; Klasen, Zvyagintsev, Weber, Mathiak, & Mathiak, 2008; McKnight & Coronel, 2017; O’Donnell et al., 2015; Schmälzle, Häcker, Honey, & Hasson, 2015). Other studies have investigated social media use (Baek, Scholz, O’Donnell, & Falk, 2017; Meshi et al., 2013, 2016; Scholz et al., 2017) or the interplay between brain function and social network variables (Falk & Bassett, 2017; Schmälzle, Brook O’Donnell, et al., 2017). Finally, a host of studies exists on the neuroscience of video-game play (Huskey, Wilcox, & Weber, 2018; Kühn et al., 2011; Mathiak & Weber, 2006), neurocinematics (Hasson et al., 2008; Hasson, Nir, Levy, Fuhrmann, & Malach, 2004), and virtual reality (Pfeiffer et al., 2014; Schilbach et al., 2010; Wilms et al., 2010). These pioneering studies demonstrate that neuroscience holds promise to “speak to” and “receive input from” almost every subfield of communication (Huskey et al., in press).

The Barriers and Challenges of Communication Neuroscience

Although we advocate that more communication scholars can integrate neuroimaging methods into their research, there are several barriers to entry (e.g., Hopp & Weber, in press; Weber, Mangus, & Huskey, 2015). Here, we provide detail on two general challenges: (1) operational challenges, such as the practical burdens of cost, time, and training, and (2) theoretical challenges, such as incommensurability between paradigms, which can result in institutional inertia and resistance. By describing these barriers and how to navigate them, we achieve a key goal of this article, which is to provide guidance and affirmation to aspiring communication neuroscientists.

Several operational challenges need to be addressed when communication scientists adopt neuroscience approaches. One key operational barrier consists of acquiring the appropriate training to conduct this research. This includes knowledge in neuroanatomy, methodological skills, computational skills, neuroscientific theories, and other general competencies. Training in neuroanatomy can be achieved with reading and/or coursework. Several excellent books cover general neuroanatomy (i.e., the brain’s geography) (Nieuwenhuys, Voogd, & van Huijzen, 2013) and functional neuroanatomy (Nolte, 2002). In addition, there are dedicated neuroscience courses, though typically offered outside communication departments (i.e., neuroscience or psychology departments). Next, although neuroimaging has the reputation of being complicated, our own experience from teaching several undergraduate and graduate introductory courses shows that, with some dedicated effort, communication students can learn the nature of what is being measured and how it is measured. After all, the core concepts of biological measurement do not differ between neuroimaging and established psychophysiological approaches (Cacioppo, Tassinary, & Berntson, 2000; Potter & Bolls, 2011). Granted, neuroimaging involves technical jargon, but once students are familiarized with the meaning of these terms, they are able to proceed, much in the same way that current non-neuroscience communication students acquire knowledge of statistics and how to set up online surveys, or learn to scrape data from social media postings. In sum, if one is committed to master this new approach, then one can certainly do it.

Another operational barrier to entry is the logistics of engaging in communication neuroscience projects, for example, in terms of costs, duration, and labor. Regarding costs, if choosing to use fMRI to collect data, access to scanners is limited and researchers will need to pay around $600 per hour in the U.S.
Often, each study participant requires at least 1 hour of scanner time, and with sample sizes usually between 30 and 60 individuals, experimental costs are higher than the usual costs in communication research (e.g., surveying students). Also, sample sizes in fMRI research continue to increase and debate is ongoing regarding correct sample sizes with certain experimental designs (Turner, Paul, Miller, & Barbey, 2018). Practically, this means that researchers need grant support if they want to collect primary fMRI data. However, possibilities for secondary data analyses are abundant as neuroimaging is currently undergoing a data-sharing revolution. Several large-scale fMRI datasets have already been made available. For example, efforts like the Human Connectome Project and the Human Brain Project, as well as large-scale clinical initiatives on autism, ADHD, and datasets during movie viewing are paving the way for population fMRI studies (Falk, Berkman, & Lieberman, 2012). These datasets will become the neuroimaging equivalent of large-scale survey research that are currently standard and widely used across the social sciences. This situation creates many opportunities, especially for researchers who may lack funding or resources for their own data collection. When researchers can use secondary data to address novel theoretical questions and employ state-of-the-art reproducible methods (e.g., Van Horn & Gazzaniga, 2013; Weston, Ritchie, Rohrer, & Przybylski, 2019), then the researchers, the field, and the taxpayer all benefit.

Another operational challenge to collecting neuroimaging data is the time required to complete a study – from start to finish, a single neuroimaging study often takes 1–2 years. This is because recordings have to be done on an individual basis and analysis can be lengthy – neuroimaging data require intensive and specialized preprocessing and usually several approaches to analysis are taken. This makes communication neuroscience studies much more slow-paced than other communication studies, and they, therefore, require added investment in doctoral training. In support of this, postdoctoral positions are the norm in neuroscience and psychology, because it takes a number of years to conduct research and receive this specialized training, but postdoc positions are rare in communication. Furthermore, PhD’s in communication tend to be shorter, allowing for less opportunity to publish if focusing on neuroimaging studies. This potential reduced output could then negatively impact graduate students’ career opportunities as they finish their degree and intend to move on to faculty positions. The field needs to be very aware of these important practical issues or it will run risk of losing connection with this emerging field.

We now discuss theoretical challenges in embracing a communication neuroscience approach, which are arguably the most prominent difficulty. Researchers encounter this difficulty when trying to align theories in communication and neuroscience. Communication Neuroscience is an inherently interdisciplinary and integrative endeavor, and therefore these challenges emerge at the intersection between established paradigms of communication and neuroscientific inquiry (Kuhn, 2012). Although excellent theoretical introductions to neuroscience exist (Gazzaniga, 2009; Kandel, Schwartz, & Jessell, 2000), including introductory neuroscience textbooks for course adoption (Ward, 2015, 2016), it is fair to say that researchers who want to embrace a communication neuroscience approach will face existential “incommensurabilities” (Kuhn, 2012). In particular, a key source of misunderstanding is that traditional neuroimaging methodologies cannot perform “mind-reading” to answer questions like “what is the content of a person’s thoughts or feelings?” This crucial misunderstanding, which we believe is still surprisingly widespread in communication, occurs because of the human tendency to make, what cognitive neuroscientists call, reverse inferences (Cacioppo, Tassinary, et al., 2000; Poldrack, 2006). To address this issue, we will next describe in detail what reverse inference means and how to avoid it.

**Reverse Inference: The Problem of Connecting the Results of Multiple Studies**

Before we can illustrate the fallacy of reverse inference, we will first introduce the notion of standard, forward inference in neuroimaging. Imagine an fMRI study that attempts to answer “which brain region is involved in the feeling of love?” To do this, researchers could put a group of people in an fMRI scanner and show them pictures of people they love – assuming that an individual feels love when they see the person they love – and pictures of familiar people that they do not love.
Importantly, all other variables between these two conditions of pictures would be held constant, such as familiarity and trust with the people depicted in the images, as well as image characteristics, like contrast and brightness. As a result, the independent variable that is manipulated is love, and the dependent variable is brain activation. After collecting data, the researchers can statistically compare brain activation when participants looked at loved ones versus looking at familiar others, revealing brain regions that are involved in processing love. Of note, a variation of this study was actually already conducted almost two decades ago, revealing such brain regions as the caudate, hippocampus and insula (Bartels & Zeki, 2000). To sum up this approach, in traditional neuroimaging research, investigators want to know what a brain region is doing, so they manipulate a stimulus, thought or feeling as the independent variable of the experiment and reveal the effect of this manipulation on the dependent variable, that is regional brain activity. These findings from the study are generalized through statistical forward inference (for example, please see Figure 1).

Now let’s consider a case of reverse inference, which can occur when another study, with a different independent variable, uses forward inference to reveal an identical/overlapping brain region. Imagine a team of researchers conducted an fMRI study that attempts to answer “which human brain region is involved when seeing one’s iPhone?” To do this, the researchers recruit a group of iPhone owners and show them images of iPhones, similar to the study above about love. Specifically, the independent variable manipulated in this study is seeing iPhones, and the dependent variable is brain activation. After collecting data, researchers analyze brain activation when participants looked at iPhones, revealing brain regions involved in processing these images. A forward inference can be made, in that the brain regions revealed by this analysis are active in the population of iPhone owners around the world when they see iPhones. Similar to the above example, a variation of this study has actually been done (Lindstrom, 2011) and it revealed that the insula, and other brain regions, are more active when viewing videos of iPhones ringing. So far this is an appropriate forward inference, namely that the insula and other regions in iPhone owners are involved in processing images of iPhones. Critically, the attentive reader may have noticed that the results from this study – insula activation – overlap with the above study on love. Now the temptation to make a “reverse inference” becomes quite strong, namely to link the independent variables of the two studies – love and iPhones – to claim that the brain data demonstrate that people love their iPhones. The person who conducted the iPhone study did, in fact, make this reverse inference between the

![Figure 1](image-url). Examples of forward and reverse inference. (a) Various stand-alone studies have investigated cognitive processes (independent variable) while measuring brain activation (dependent variable) to reveal the involvement of the insula in each process. Solid lines depict the appropriate forward inference conclusions from these studies, which is that the insula is involved in all of these types of cognition/perception. (b) When a study (e.g., on iPhone perception) reveals an overlap with other studies in regard to the dependent variable, in this case the insula, it is tempting to make an incorrect reverse inference. Dotted lines depict these inappropriate reverse inferences with respect to the iPhone study. These reverse inferences are incorrect because a link can be drawn to all other cognitive processes previously investigated that revealed the same brain region. Therefore, we cannot conclude if people love, fear, or are disgusted by their iPhones.
two studies and the article he wrote for the New York Times conveyed this conclusion, titled “You love your iPhone. Literally.” (Lindstrom, 2011). For illustration of this reverse inference, please see Figure 1.

However, does this overlap in insula activation really mean that people love their iPhones? No, it does not, and a group of prominent neuroimaging researchers responded to and criticized the New York Times piece (Poldrack, 2011). This is because the independent variable in the iPhone study was not love, it was viewing iPhones, and one-to-one mappings between brain regions and mental states are extremely rare – the norm is that one brain region is involved in many processes (a one-to-many mapping) so choosing to conceptually link back to the investigated mental state (independent variable) of a previous study with an overlapping dependent variable (e.g., insula activation) is incorrect (Cacioppo, Tassinary, et al., 2000). For example, many fMRI studies have revealed the insula to be active in response to a wide variety of stimuli, thoughts or feelings that were manipulated as the independent variable in each study, such as disgust, fear, pain, anxiety, decision-making, speech generation, hearing and more (for review Uddin, Nomi, Hébert-Seropian, Ghaziri, & Boucher, 2017). Therefore, one could easily have made the same reverse inference link between the iPhone study and disgust, and claimed that because the insula is active when viewing iPhones and also when experiencing disgust, people are disgusted by their iPhones. Equally plausible, one may rely on the association between fear and insula activity, and claim that people are afraid of their iPhones. It becomes clear that when we make reverse inference claims, linking the independent variables of two, separate neuroimaging studies because of a similar finding with the dependent variable (e.g., overlapping brain region), we get into murky waters.

The human tendency to make reverse inferences drives much of the above-mentioned misunderstanding about the role neuroimaging can play in communication science. The issue occurs, in part, because even the most advanced neuroimaging methods, such as fMRI, have limited spatial resolution – and thus cannot determine the exact neurons that are firing, but only larger clusters of thousands of neurons. So one study about love finds, for example, the three brain regions mentioned above – the caudate, hippocampus and insula – and another study, with a different independent variable, could also find the insula, along with several other regions. Then, there is an overlap in findings at the insula and this tempts researchers to draw an incorrect link between the independent variable in these two studies.

This concept of reverse inference is crucially important for understanding what neuroimaging can and cannot do, and how it can advance communication theory. Currently, one cannot design a traditional neuroimaging study and expect to know what people are thinking or feeling by interpreting the dependent variable – the brain region that is revealed as active – as this would require reverse inference. In most cases, conclusions cannot be made by linking one’s neuroimaging findings to other independent studies that came before. There are notable exceptions, however. For example, using previous neuroimaging publications, a Bayesian probability can be calculated which reveals the likelihood that if regional brain activation is observed, a certain mental state was present – therefore obtaining the elusive one-to-one mapping of brain region activation and mental state mentioned above. This has been done, for example, with the nucleus accumbens and rewards (Ariely & Berns, 2010). The nucleus accumbens is a small brain region contained within a larger region called the ventral striatum (Haber & Knutson, 2010), and numerous studies have shown that the receipt of a reward activates the nucleus accumbens/ventral striatum (Clithero & Rangel, 2013). Ariely and Berns used Bayesian statistics to calculate that there is a 90% chance, if one observes nucleus accumbens activity in a study, that the study participants were rewarded. Without this type of Bayesian calculation based on myriad previous papers, however, a reverse inference is not possible. So, for example, if one wanted to investigate media effects by showing Instagram images to individuals in the scanner, and this revealed activity in the insula, this researcher would not be able to conclude what the participants were thinking or feeling when viewing these images. This researcher would, however, be able to claim that the insula is involved in the processing of Instagram-posted media images, whereby the arrow of causality goes from the presentation of the
images to the activity in the brain. In sum, misunderstanding of this critical reverse inference issue can result in a mismatch of expectancies about what communication neuroscience can provide, and what other fields of communication want. Despite these important caveats, neuroimaging has the potential to markedly advance communication science, as we describe in the next section.

**Approaches to Improve Explanation and Prediction of Communication Phenomena**

Communication science aims to provide an explanation and/or prediction of communication phenomena. Pioneering work in communication neuroscience, such as the studies mentioned above, has already contributed to this endeavor. The next generation of communication scholars can build on this groundwork by capitalizing on three different approaches for using neuroimaging to test or expand communication theory. These approaches differ by the conceptual role that the neural measures take in the study: (1) that of a dependent variable, (2) that of a mediator/moderator, or (3) that of a predictor. These approaches align well with those already taken in communication science with other forms of data. In the following section, we introduce each approach and provide examples of how they can be implemented, as well as what communication scholars can learn from their implementation.

**The Brain as a Dependent Variable**

First, if one’s primary goal is to understand how the brain responds to and processes communication phenomena, then researchers can take the “brain as dependent variable” approach (Figure 2a). This approach uses neuroimaging to test hypotheses about underlying neural correlates of communication (Huettel, 2008), and it is exemplified above with the study about “love”. By using “love” as the independent variable and measuring the effect of its manipulation on brain activity, this approach can reveal brain regions or networks involved in specific thoughts, emotions and perceptions (e.g., most communication phenomena). In other words, the goal of these brain-as-DV studies is to identify the “who, what, when, where and how” concerning the brain and particular thoughts, emotions or perceptions. During the early days of neuroimaging, researchers primarily used this approach to “map out” how different brain regions respond to manipulations, for example, in response to various perceptual (e.g., brightness), cognitive (e.g., decision-making), or emotional (e.g., love) phenomena. Research of this sort still comprises the majority of

![Figure 2. Examples of the different ways that communication researchers can capitalize on neuroimaging data. Neuroimaging data can help explain communication phenomena by serving as (a) a dependent variable, (b) a mediator/moderator, and as (c) a measure to predict communication outcomes.](image-url)
studies in cognitive, affective, or social neuroscience. Of note, different neuroimaging metrics can be used to apply the brain-as-DV approach, such as regional brain activity, connectivity between regions, or inter-subject correlations (Huskey et al., 2018; Mesihi et al., 2013; Schmälzle et al., 2013). Importantly, these same measures can also be used to apply the other below-described conceptual approaches (i.e., brain-as-mediator, brain-as-predictor). Studies taking a brain-as-DV approach typically have titles like “Neural correlates of…” and then they list the perception, cognition or emotion that was studied. The findings of these studies contribute to an ever-growing knowledge base about how the brain responds to psychological manipulations. In fact, by aggregating data across many studies, it has even become possible to establish meta-analytic platforms that continue to grow as more work is published (Fox & Lancaster, 2002; Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011).

Combining Computational Models with Neural Data: Insights into Hidden Processes

Despite the enormous growth of neuroimaging and the cumulative body of knowledge it provides, some critics have found this approach unsatisfying (Coltheart, 2013). These critics argue that such “brain mapping” studies contribute little to theory, or at least do not provide the sort of contributions they would like to see (e.g., Uttal, 2015). More advanced forms of neuroimaging, called computational or model-based neuroimaging, strive to overcome such criticisms. Specifically, these approaches relate the measured brain activities to parameters (internal variables) of formal computational models.

To give an example, we return to the brain region we mentioned above when discussing reward processing, the ventral striatum. Many studies have demonstrated that rewards activate the ventral striatum (Clithero & Rangel, 2013). This is done, for example, by comparing brain activity while rewarding stimuli are presented (i.e., winning money) to brain activity while neutral control stimuli are presented. As discussed above, such designs enable a forward inference, namely concluding that winning money activates the ventral striatum. Beyond such descriptive statements, however, psychological theories strive for deeper explanations especially regarding how reward outcomes are computed to result in learned associations over time. Standard theories of reward learning, going way back to the classical models of Rescorla and Wagner (1972), posit that: (1) the brain makes an actual computation, called a reward-prediction error, which is the quantifiable difference between the actual reward amount and the amount one had expected, and (2) that learning depends on the size of the reward prediction error (termed error- or difference-based learning). Importantly, models of reward learning have been mathematically formalized so that they include numerical values of latent cognitive variables (like reward prediction errors, which is often denoted as delta). This is an important improvement compared to most psychological theories that stay at a verbal-descriptive level. In a computational, model-based neuroimaging study of reward learning, one can, therefore, use the trial-by-trial predictions for such a latent reward-prediction error variable and test whether and/or where it correlates with the measured brain activity. Of note, strong evidence implicates the neurotransmitter dopamine and the ventral striatum in these functions (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006; Schultz, Dayan, & Montague, 1997). In this way, the computational model provides a bridge between behavioral and neural data. Using neuroimaging in this manner arguably gets us much closer to answering the question of “how” brain activity gives rise to cognitive phenomena (Forstmann & Wagenmakers, 2015; O’Reilly & Mars, 2011; Palmeri, Love, & Turner, 2018).

This computational neuroimaging approach represents an advanced subtype of the brain-as-DV approach and holds great promise to advance communication theory. For example, the concept of reward prediction errors not only stimulated much research on reinforcement learning within cognitive and affective neuroscience, but it has inspired work in social neuroscience (Cheong, Jolly, Sul, & Chang, 2017; Hackel & Amodio, 2018; Stanley & Adolphs, 2013). For example, computational, model-based neuroimaging has been used to study how we learn vicariously through observing others (Burke, Tobler, Baddeley, & Schultz, 2010; Olsson, Nearing, & Phelps, 2007), or how the brain keeps track of social norm violations (Zinchenko & Arsalidou, 2018) – both of which are central topics within the communication discipline. Similarly, the general notion of difference-
based computation features prominently in theories of social comparison (see Luo, Eickhoff, Hétu, & Feng, 2018 for neuroimaging studies) and could easily be extended toward Expectancy-Violation-Theory in interpersonal contexts (Burgoon & Hale, 1988), although we are not aware of any studies to date. Within the domain of persuasion, relatively well-specified theories like Social-Judgment Theory (Sherif, Sherif, & Nebergall, 1965) and Information-Integration Theory (Kaplowitz, Fink, Armstrong, & Bauer, 1986) are likewise promising candidates for computational studies. Overall, although the level of formalization of socially focused theories still lags behind that of other areas, particularly the highly advanced models that exist in visual neuroscience (e.g., Engel, 2008; Itti & Koch, 2001; Kriegeskorte & Douglas, 2018; Wu, David, & Gallant, 2006), we expect that a computational approach to communication neuroscience can make key contributions on phenomena that have henceforth remained elusive. In summary, the brain-as-DV approach can build the necessary bridge between neuroscience and communication. Communication scholars can use this approach to specifically answer questions about the brain in regard to communication phenomena. Thus, this approach is useful across many subfields of communication research.

**The Brain as a Mediator**

Second, researchers can take the “brain as mediator” approach (Figure 2b). In this type of research, one can ask if a specific brain activity mediates a particular relationship between two communication-relevant concepts (Falk, Cascio, & Coronel, 2015). Statistical mediation analyses are very common across the social sciences and generally regarded as a tool to identify processes and explain mechanisms between the input and output variables of a system (Baron & Kenny, 1986; O’Keefe, 2003; Spencer, Zanna, & Fong, 2005). For example, a psychological variable obtained by the survey may be identified as a statistical mediator of two communication input and output variables. Importantly, this relationship can be improved upon using neuroimaging measures because, logically, it must be the brain that mediates communication input and output variables, as all thoughts and actions arise from brain activity. Therefore, brain-as-mediator studies can be performed to reveal specific brain mechanisms that account for the link between input and output. For instance, this approach has been productive in studying the brain mechanisms that mediate the input of painful (nociceptive) stimuli and subjective reports of pain; this approach identified particular regions of the anterior cingulate cortex and the periaqueductal gray as key neural mediators (Geuter et al., 2018). To illustrate how a brain-as-mediator approach can be taken in a communication context, we consider the following example. There is some evidence suggesting that reading literary fiction improves theory of mind performance (Kidd & Castano, 2013; Mar, Oatley, & Peterson, 2009). In parallel, ample evidence points to the so-called mentalizing network – a set of brain regions that includes the temporoparietal junction – as a brain system that is involved in inferring the mental state of others (Saxe & Kanwisher, 2003; Schurz et al., 2014). Thus, one may assume that training theory of mind via fiction-reading should also be represented by changes in the brain – similar to the effects that learning a motor skill (e.g., playing piano) has on the brain’s motor system (Bassett, Yang, Wymbs, & Grafton, 2015; Dayan & Cohen, 2011). To test this hypothesis, one could examine whether the relationship between reading and theory-of-mind skills is mediated by activity in the temporoparietal junction. If a significant mediation is observed, then this would provide more specific insight into how literary reading actually causes theory-of-mind skills to improve.

However, the term “causes” should be used with caution because when neuroimaging measures are used in this way, they provide only correlational evidence. Thus, to make a fully causal argument, one would need another study, such as one that disrupts activity in the temporoparietal junction (e.g., with transcranial magnetic stimulation) to show that this breaks the association between reading and theory-of-mind skill improvement. Nevertheless, even though this neuroimaging approach remains correlational, this mediator-identification strategy can greatly aid communication science’s search for mechanisms by providing more specific insights into the relationship between communication input variables, the regional brain responses they evoke, and the resulting output variables. In this sense, communication...
neuroscience helps us to build better, more complete theories, and as such, the brain-as-mediator approach will prove invaluable to anchor hypothetical constructs in observable brain mechanisms. Of note, just like in the brain-as-DV approach, the inclusion of formal computational models would further boost the explanatory power of this brain-as-mediator approach. Thus, if a computational model could make specific predictions about how the information gleaned from fiction reading improves mentalizing functions, and if temporoparietal junction activity tracked with such a model’s parameters, then this would offer an even more convincing explanation. At present, however, computational modeling of mentalizing and other social processes is still at an early stage (Cheong et al., 2017; Gonzalez & Chang, 2019; Hackel & Amadio, 2018) and our example remains hypothetical.

**The Brain as Predictor**

Finally, researchers can take the “brain as predictor” approach (Figure 2c). In this type of research, scholars can use brain activity to predict future behaviors (Falk et al., 2015; Weber et al., 2018). More specifically for communication researchers, brain activity can be measured at the time of message exposure and either the raw brain activity or derivative measures can then be used to predict a future outcome, for example, a behavioral measure collected at a later time point that is related to the message. Importantly, this approach can be applied to both within-sample behaviors (predicting the future behavior of the study participants who provided the brain data) and out-of-sample behaviors (predicting the behavior of a population from the brain data of a small group of study participants). For example, with regard to within-sample predictions, Falk and colleagues used fMRI to measure the brain response of smokers while viewing anti-smoking messages and linked these responses to reductions in participant smoking 1 month later (Falk, Berkman, Whalen, & Lieberman, 2011). Important to note, in this study, brain data were able to predict message effectiveness above and beyond self-report. This further demonstrates the benefit of adding neuroimaging methods to one’s research program. With regard to out-of-sample predictions, Falk and colleagues followed their within-sample study by using the group’s brain response to anti-smoking messages to predict message effectiveness at the population level (Falk et al., 2012). Specifically, the brain data predicted which anti-smoking advertisements motivated more people to contact an anti-smoking hotline. In another example of out-of-sample prediction research, Weber et al. (2015) measured the brain response to anti-drug messages in a group of study participants and forecasted message effectiveness in two other groups of individuals. Of note, with the brain-as-predictor approach, the function of the brain region used to predict behavior does not necessarily need to be germane to the research topic. In practice, however, the choice of brain region must either be guided by a-priori considerations, or if a data-driven approach is taken, the predictive models must be cross-validated to guard against spurious results (Varoquaux et al., 2017). This brain-as-predictor approach is popular beyond the field of communication due to its obvious significance for other domains. For instance, brain-as-predictor studies have been used in neuroeconomics to predict commercially relevant outcomes such as music album sales and crowdfunding of projects (Berns & Moore, 2012; Genevsky, Yoon, & Knutson, 2017; Knutson & Genevsky, 2018), and there is great interest in developing brain markers for clinical diagnoses or subjective experiences such as pain and other similar phenomena (Gabrieli, Ghosh, & Whitfield-Gabrieli, 2015; Rosenberg, Casey, & Holmes, 2018; Woo, Chang, Lindquist, & Wager, 2017).

**Future Developments in Communication Neuroscience**

As described above, communication neuroscience strives to understand how the brain responds to and produces communication phenomena. Brain activity can also be used to predict communication outcomes. Each of the three experimental approaches outlined above (see Figure 2) can produce new insights into communication phenomena, and we expect that future developments will expand our understanding of the neural basis of communication effects, mechanisms, and predictive markers.
Therefore, we next propose three potential developments that stand to further deepen the impact of communication neuroscience.

First, measurement of brain activity will greatly improve in the future. Driven by funding from both government and private industry, brain imaging methods are constantly improving – examples of this include the BRAIN initiative by the US, the EU’s Human Brain Project, the China Brain Project, and Facebook’s significant investment in neurotechnology development. Thus, current methods will continue to evolve, for example, by making existing technologies more sensitive, advancing their portability, or decreasing their costs. In addition, new methods will also appear. For example, fMRI will not be the final significant neuroimaging technology developed, entirely new imaging methods will be created that allow us to image the brain with improved spatial and temporal resolution.

Next, advances in the analysis of neuroimaging data will improve our ability to extract knowledge from data. This will greatly enhance the sensitivity of existing methods. For example, an important recent development is the application of machine learning to identify patterns of brain activity (whether they serve as dependent variable, mediators, or predictors; Chang, Gianaros, Manuck, Krishnan, & Wager, 2015; Haxby et al., 2001; Norman, Polyn, Detre, & Haxby, 2006). Similarly, the rise of graph-theoretical analyses promises to yield a new understanding of the brain as a complex network (Bassett & Sporns, 2017). In addition, emerging approaches allow for analyses of shared experiences in pairs and groups of individuals, which is novel when compared to traditional, one-brain-at-a-time analytical approaches. For example, analyses can examine shared neural responses for an audience watching a film (Cohen et al., 2017; Schmälzle & Grall, in press). An in-depth discussion of these emerging techniques for Communication Neuroscience has recently been provided by Turner, Huskey, and Weber (2019).

Finally, at the intersection of novel measurement and analysis methods, we see brain-computer-interfaces (BCIs) as another emerging trend (Guger, Allison, & Edlinger, 2013). In brief, BCIs are systems that record data from the live brain and immediately analyze the results to make a prediction about behavior, for example, by letting patients who cannot speak select letters or words to express themselves, or to control limb prosthetics. As of yet, most BCI applications lie within motor and elementary cognitive domains, and the systems are invented by engineers, but rarely used for scientific purposes. However, it seems possible that applications in social and affective contexts could be developed and the field could become more integrated with communication neuroscience.

Finally, and most importantly, we foresee that the theoretical advances spurred by research within communication neuroscience will create dynamic feedback loops with the general communication field to fuel future theory and research. Evidence of such a development can be seen in cognitive neuroscience, where neurally specified computational models are gradually filling in the black box of the brain with computationally defined mechanisms (Hassabis et al., 2017; Kriegeskorte & Douglas, 2018). In a similar vein, as computational approaches to social neuroscience are becoming more prominent (Stanley & Adolphs, 2013), they will provide us with a deeper understanding of how social processes emerge from our neural biological substrate.

**Summary**

In conclusion, the brain – the biological organ of communication – has a firm foundation in the future of communication research, as communication neuroscience has developed into an established approach to study communication. Without much effort, communication scholars can harness the power of neuroimaging measurements to address key theory and research questions in their specific domains. Communication scholars can capitalize on the many advantages provided by these methods, improving our ability to explain and predict communication phenomena. However, the opportunity does not come without costs: projects in communication neuroscience take longer, special training is required, and data are more difficult to collect, analyze, and interpret. We have highlighted the pervasive issue of reverse inference as one constraint on how neuroimaging can and cannot contribute to explanation-focused communication theory. Despite these challenges, the future is bright for communication neuroscience. To support this, we provided three approaches that communication scholars can take to address their
research questions with neuroimaging data. These approaches will allow researchers to test theories about hypothesized communication mechanisms and stimulate the creation of new theories. Overall, this upcoming research will greatly expand our knowledge of communication phenomena. In parallel, methods and theories in cognitive neuroscience will continue to advance at a rapid pace. These forthcoming developments will also influence the field. In sum, we look forward to the contributions of future generations of communication neuroscientists.

Acknowledgments

We thank Maral Zakharia for creating the graphical illustrations, Kjerstin Thorson for providing helpful feedback on an earlier version of the manuscript, and three anonymous reviewers for their helpful feedback and discussion.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Ralf Schmälzle http://orcid.org/0000-0002-0179-1364
Dar Meshi http://orcid.org/0000-0003-0772-037X

References


