57 An Introduction to Time-Resolved Decoding Analysis for M/EEG

THOMAS A. CARLSON, TIJL GROOTSWAGERS, AND AMANDA K. ROBINSON

ABSTRACT The human brain is constantly processing information in order to make decisions and interact with the world for tasks ranging from recognizing a familiar face to playing a game of tennis. These complex cognitive processes require communication between large populations of neurons. The noninvasive neuroimaging methods of electroencephalography (EEG) and magnetoencephalography (MEG) provide population measures of neural activity with millisecond precision that allow us to study the temporal dynamics of cognitive processes. However, multisensor M/EEG data is inherently high dimensional, making it difficult to parse important signals from noise. Multivariate pattern analysis (MVPA) or *decoding* methods offer vast potential for understanding high-dimensional M/EEG neural data. MVPA can be used to distinguish between different conditions and map the time courses of various neural processes, from basic sensory processing to high-level cognitive processes. In this chapter we discuss the practical aspects of performing decoding analyses on M/EEG data, as well as the limitations of the method, and then review some applications for understanding representational dynamics in the human brain.

One of the most remarkable aspects of the human brain is its speed and processing efficiency. A familiar friend is recognized in an instant, speech communication feels like a natural dynamic exchange, and we can make split-second decisions in life-threatening situations. A crucial question in cognitive neuroscience is how the brain manages such complex tasks with ease. As noninvasive brain-imaging techniques with millisecond resolution, magnetoencephalography and electroencephalography (M/EEG) offer unparalleled potential in capturing the dynamics of human cognition. Multivariate pattern analysis (MVPA or *decoding*) methods, in conjunction with M/EEG data, give insight into the temporal dynamics of information processing in the brain. Such methods yield insight not only into the time course of specific cognitive processes but also into the relative order of different cognitive processes. This can lead to increased understanding of the temporal dynamics of brain representations—for example, how low-level visual representations are transformed into high-level object representations (Contini, Wardle, & Carlson, 2017).

When MVPA was first introduced to functional magnetic resonance imaging (fMRI), it gave cognitive neuroscientists the unprecedented capacity to decode information in the brain (for recent reviews, see Haxby, 2012; Haynes, 2015; Tong & Pratte, 2012). MVPA in combination with fMRI's excellent spatial resolution enabled researchers to target and scrutinize the information represented in different brain areas. MVPA methods for M/EEG data complement fMRI findings by providing valuable insight into the time course of neural processing. Accordingly, M/EEG decoding methods are becoming increasingly popular in the cognitive neuroscience community.

Most cognitive neuroscientists are aware of decoding methods, as they have been actively used in fMRI research for two decades. Interestingly, methods for decoding mental states from the brain predate this work by a quarter century. Vidal first developed these methods for EEG in the context of brain-computer communication, asking the provocative question "Can observable electrical brain signals be put to work as carriers of information in man-computer communication or for the purpose of controlling such external apparatus as prosthetic devices or spaceships?" (Vidal, 1973, p. 157). Vidal's and subsequent work focused on the practical applications of having access to the brain's internal mental states. Many useful applications for decoding methods exist in M/EEG research, including brain-computer interfaces (BCI; e.g., Wolpaw et al., 2002), lie detection (e.g., Davatzikos et al., 2005), and the diagnosis of brain disorders (e.g., Ewers et al., 2011). Practical applications such as these place greater weight on performance over explanation. If diagnostic accuracy for schizophrenia using a particular method is increased by 2%, this represents a significant achievement that can have wide-reaching benefits. In contrast, the cognitive neuroscientist is interested in how the brain works. Here, the priority is to elucidate the neural mechanisms underlying cognitive processes, and metrics like prediction accuracy are of lesser importance (Hebart & Baker, 2017). The cognitive neuroscience application of decoding methods for M/EEG is

the topic of this chapter. Henceforth, we will use the term *decoding* to reference the use of this method for understanding the brain.

This chapter focuses explicitly on time-resolved decoding analysis for M/EEG. MVPA can be applied to M/EEG data (frequency analysis, wavelet decomposition, connectivity, and more) in a variety of domains. We will limit the scope of our discussion to time-series decoding, but many of the same principles apply to other domains. This research uses paradigms that closely resemble EEG's event-related potential (ERP) or MEG's event-related field (ERF) research. In these experiments, participants are presented with stimuli (e.g., images, sounds, and others) from different conditions (e.g., faces vs. houses; attended vs. unattended), and the analysis is time locked to the presentation of the stimulus. MVPA and ERP/ERF analyses both seek to determine any differences between experimental conditions. The main difference is that ERP/ERF research uses a univariate approach, such as testing for differences between conditions from a single recording site or from data averaged across multiple sites. MVPA, in contrast, utilizes pattern information across recording sites and thus is often a more sensitive measure. As the topic of this chapter is MVPA, we will not discuss further differences between the approaches except to make illustrative points (for more information, see Grootswagers, Wardle, & Carlson, 2017; Hebart & Baker, 2017). For those interested in ERP/ERF analysis or more general information about the analysis of EEG recordings, we refer the reader to the excellent text on ERP research by Luck (2005).

The fundamental goal of the decoding analysis is to learn what and how information is represented by the brain (see chapters 56 and 58). M/EEG recordings are a measure of brain activity with high temporal resolution, and MVPA is a sensitive measure to determine whether or not any information in the recordings can distinguish between experimental conditions. The combination of the two can tell the researcher if there is information in the recordings of brain activity that can distinguish experimental conditions for each point in time. Going beyond this to address the more critical questions of "what" and "how" information is represented in the brain requires specific knowledge of the field of inquiry, which requires the development of robust experimental paradigms.

The goal of this chapter is to inform readers with little MVPA experience about the mechanics of applying these methods to M/EEG data and guide researchers on using these methods to answer the critical questions of "what" and "how" information is represented in the brain. We assume the reader has a basic knowledge M/EEG and thus will focus on the practical details of decoding analyses rather than aspects of conducting M/EEG experiments. We also note that M/EEG-MVPA is a relatively new field of research. On the one hand, this means that many exciting new theoretical questions can be explored using these powerful analysis methods. On the other hand, it means the techniques themselves are still evolving, and there is currently limited technical guidance (see Grootswagers, Wardle, & Carlson, 2017 for an advanced tutorial). At times, we will give recommendations based on our own experience that lack a scientific reference. We expect many of these recommendations to be followed up empirically in the coming years.

In this chapter, we first provide a foundation by situating pattern analysis methods in the context of M/EEG data. We then discuss some of the intricacies of using decoding methods for analyzing M/EEG data. We then conclude with a section describing advanced methods to familiarize the reader with more sophisticated approaches to time-resolved decoding.

The Basics of Time-Resolved Decoding Analysis for M/EEG

To give intuition to the analysis and a framework for discussion, we describe a hypothetical experiment. Consider a standard EEG system with 64 sensors (figure 57.1*A*), each of which continuously measures local electrical currents on the scalp (figure 57.1*B*). In each trial, participants are shown an image of either an *X* or an *O*, and there are 40 trials for each condition. To simplify the description of the analysis, we start by considering only two electrodes centered on the occipital cortex (e.g., O1 and O2) and data from a single time point (e.g., 100 ms poststimulus onset). This narrow slice of data can be represented as a two-dimensional scatterplot (figure 57.1*C*). In the plot, the two axes are the measurements from the two electrodes in microvolts (μV) for the single time point. The 80 data points are the measurements recorded for the individual trials, denoted by *X*s and *O*s. In this simplified scenario, the EEG data take the form of the familiar diagram given in virtually all texts describing pattern classification analysis (e.g., Duda, Hart, & Stork, 2001).

Pattern classification analysis is used to determine if there is information in the EEG measurements that can predict whether the participant was viewing an *X* or an *O* on a given trial. The analysis is a two-step process. The first step is to train the classifier to find a decision boundary that best separates the *X*s and *O*s using the EEG data. For the purposes of this initial description, we use a linear classifier, which uses a linear decision

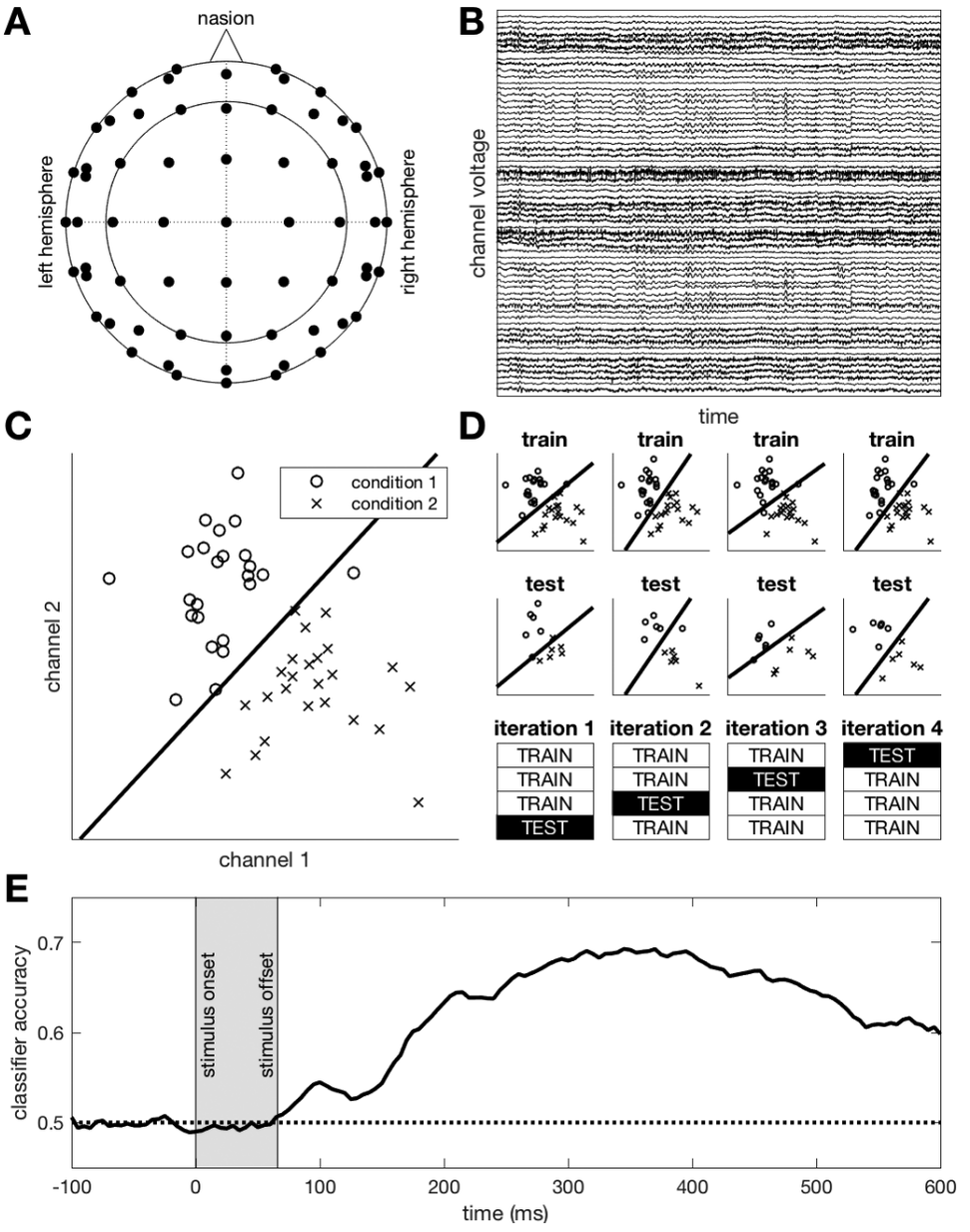


FIGURE 57.1 Time-resolved decoding analysis for M/EEG. *A*, Top-down view of a channel layout for a 64-channel EEG. *B*, Sample recording data from an EEG experiment. *C*, Scatterplot showing hypothetical EEG data from a single time point for a decoding experiment. The two axes are the measured current from two EEG electrodes. The points are individual trials for the *X* and *O* conditions. The line denotes the optimal decision boundary for classifying *X*s and *O*s derived from the generative model. *D*, Four iterations of cross-validation using 75% of the data to train the classifier and 25% to test the classifier. The top row of plots show the training data along with the classification boundary computed using linear discriminant analysis (LDA) for each iteration. The plots below show the test data for each iteration with the decision boundary computed from the training data. *E*, An example of a time-resolved decoding analysis showing classifier accuracy averaged over participants. Data from Grootswagers, Ritchie, Wardle, Heathcote, and Carlson (2017).

boundary for classification. In the plot in figure 57.1*C*, the observations to the left of the boundary are classified as *O*s, and those to the right of the boundary are classified as *X*s.

The second step is testing the classifier, in which a trained classifier is used to predict whether the participant was viewing an *X* or an *O* in a new set of trials. It is essential to use independent data for the two steps (training and test) to show that the classifier can generalize to independent data outside the training sample. For this, we use cross-validation, which involves splitting the data into training and test sets. In this example, we divide the data into four blocks and use 75% of the data to train the classifier and the other 25% to test the classifier. The classifier is trained four times, each time using three blocks of data to train the classifier and the remaining block to test the classifier (figure 57.1*D*). The average prediction accuracy (performance) of the classifier across the four iterations is taken as an estimate of decoding performance for classifying *X*s and *O*s from the EEG data at this time point. Importantly, if the classifier accuracy is above chance (i.e., 50%), this is taken as evidence there were different patterns of activation across the two channels when the participant viewed an *X* or an *O*.

To bring this simplified scenario back to a full M/EEG data set is only a matter of scaling up. First, instead of limiting the analysis to two channels, all of the channels or a selected subset of the channels is used for the analysis. Second, the analysis is repeated for every time point instead of only on a single time point. The result is a time-varying measure of stimulus/condition decodability (figure 57.1*E*). Finally, the experiment is repeated in multiple subjects to enable statistical inference, treating the subject as a random effect (see the section on statistical assessment).

The above provides a broad overview of the critical features of decoding analysis for M/EEG. In the following sections, we go into more detail about the steps in the analysis and discuss specific issues for the analysis of M/EEG data.

Practical Aspects of M/EEG Decoding Analysis

On preprocessing M/EEG data for multivariate pattern analysis M/EEG data is inherently noisy, picking up electrical activity from a variety of nonbrain sources, such as muscle activity, eye movements, and environmental noise. M/EEG research uses many preprocessing procedures to reduce this noise, such as filtering, resampling, and artifact rejection. The goal of preprocessing is to increase the signal-to-noise ratio (SNR), reducing the probability of obtaining a false negative (i.e., failing to observe an effect that is present). Preprocessing pipelines can vary between different laboratories (and even across individual researchers in a laboratory), as there is no consensus on preprocessing pipelines. The variability in preprocessing pipelines has been identified as an important issue for reproducibility in neuroimaging research (Poldrack et al., 2017). In developing a preprocessing pipeline, one thus needs to balance the desire to maximize SNR with minimizing the number of preprocessing steps to ensure the results are robust.

The analysis of ERP/ERF in EEG/MEG, respectively, has been around for decades, while decoding methods are relatively new. A natural starting point for thinking about a preprocessing pipeline for M/EEG decoding analysis might be to adopt established methods from ERP/ERF research (Luck, 2005). However, MVPA is different from ERP/ERF analyses, and thus the preprocessing steps need not be the same. In particular, decoding is more robust to artifacts than these other methods. Standard classifiers, such as linear discriminant analysis (LDA) and support vector machines (SVM), implicitly model the noise in the data. Therefore, classifiers can replace steps that typically require laborious manual inspection, such as bad sensors and artifact rejection. To understand this, we need to look at how machine-learning classifiers like LDA and SVM maximize the prediction.

First, the classifier assigns weights to each sensor. High weights enhance information in the measurement that informs the classification, and low weights suppress uninformative information in the measurement. This aspect is relevant because it potentially makes some preprocessing procedures redundant. In the context of EEG, an example is a broken or very low impendance electrode. In a standard ERP analysis, this electrode would be removed or interpolated during preprocessing. For MVPA, this step can be omitted because the classifier would learn in training that the noisy electrode is not informative for the prediction and would assign it a low weight. Note also that channel interpolation is a linear combination of surrounding electrodes and thus adds no additional information to the classifier. Eyeblink and muscular artifacts are other examples. For ERP analysis, these artifacts can have a significant effect on the quality of the data and thus need to be removed. For MVPA, these artifacts are of lesser concern. If the artifact is not informative for the prediction, the classifier will suppress the artifacts by giving low weight to the component of the signal related to the artifacts. Removing trials with eyeblink artifacts is thus another preprocessing step that potentially could be removed from the MVPA preprocessing pipeline (Grootswagers, Wardle, & Carlson, 2017). Note that artifacts confounded with

the conditions of interest are a serious concern for all (MVPA and ERP) analyses!

Second, classifiers can cancel noise in the data to improve the prediction. This is relevant when considering the effect of environmental noise. For ERP analysis, a noisy channel increases variance in the estimate of the average evoked response, and thus it is advisable to interpolate noisy channels or exclude them from the analysis. In contrast, classifiers (e.g., LDA) can use the information about the noise to improve prediction. For example, suppose we have two hypothetical channels. One channel contains brain activity differentiating the two experimental conditions and also environmental noise. The second channel is far from the source of brain activity and only contains the environmental noise. The classifier can use an estimate of the environmental noise from the second channel and subtract it from the first channel so that what remains is the signal that differentiates the two experimental conditions. Including the noisy channel can therefore give the classifier *more* information about the signal of interest than if the channel had been removed (Averbeck, Latham, & Pouget, 2006; Haufe et al., 2014).

In summary, MVPA is a relatively new approach to analyzing M/EEG data that does not have stringent noise-reduction requirements, as do typical ERP analyses. The critical lesson to be taken concerning preprocessing is that MVPA can inherently deal with some noise and artifacts in the data. This is an advantage of the decoding approach because it reduces data-processing steps relying on subjective criteria (e.g., the interpolation of noisy channels identified by visual inspection). While systematic research is still required, M/EEG MVPA research will need to develop its own guidelines for preprocessing. These new standards will need to take into account how classifiers operate to strike a balance between optimizing SNR and minimizing the number of preprocessing steps and options for reproducible research (Poldrack et al., 2017; for early explorations into the effect of preprocessing pipelines on M/EEG decoding, see Grootswagers et al., 2017).

Classifier selection Decoding analysis generally favors linear classifiers. As the name implies, linear classifiers use a linear boundary (e.g., see figure 57.1) or a hyperplane for greater than two dimensions for classification. Compelling arguments have been made for the use of linear classifiers in fMRI research (Kriegeskorte, 2011; Misaki et al., 2010; Muller, Anderson, & Birch, 2003; Mur, Bandettini, & Kriegeskorte, 2009; Pereira, Mitchell, & Botvinick, 2009; Schwarzkopf & Rees, 2011). In contrast, there has been little systematic discussion of this for M/EEG. Below, we give an overview of the different classifiers available for M/EEG decoding research and step into the fMRI community's argument for linear classifiers, discussing them in the context of M/EEG.

Although nonlinear classifiers have the added capacity to fit more complex class boundaries using nonlinear terms, in practice they generally perform equal to or worse than their linear counterparts in neuroimaging decoding studies (Misaki et al., 2010). This is because the increased flexibility of nonlinear classifiers comes at the cost of overfitting, which limits generalization performance. Furthermore, the results from nonlinear classifiers are harder to interpret. As previously noted, it is also essential to keep in mind that cognitive neuroscience applications of MVPA focus on understanding the brain, where improvements in performance are of marginal value (see the introduction to this chapter). A small increase in performance is a hefty price for the loss of interpretability, which we discuss below. Thus, unless there is substantial justification, linear classifiers are the preferred method for cognitive neuroscience applications of MVPA for the analysis of M/EEG data.

Linear classifiers have two interpretive advantages from a cognitive neuroscience perspective. First, linear classifiers are a biologically plausible form of a "readout" (DiCarlo & Cox, 2007). Specifically, the information used by a linear classifier could also be "read out" by a single downstream neuron (Kriegeskorte, 2011; Misaki et al., 2010). In the context of M/EEG, this is less compelling. Embedded in this reasoning is the fact that the downstream neuron has access to the same information as the classifier (Carlson et al., 2017). A single EEG channel records the aggregated activity of thousands of neurons from different brain regions, and the entire EEG cap has access to most of the neurons on the cortical surface. It is not biologically plausible that a single neuron has access to all the information represented on the cortical surface.

The second interpretive advantage is that linear classifiers produce weight maps that can be visualized to gain insight into the source of decodable information (Kamitani & Tong, 2005). An fMRI-MVPA experiment studying faces, for example, might show that voxels in the fusiform face area (FFA; Kanwisher, McDermott, & Chun, 1997) are given high weights. The weight maps thus can be used to infer that the FFA is a strong candidate source of decodable information. Similarly, for M/EEG, the classifier weights can be projected back to the topographical map of the sensors to gain insight into the source of decodable information (but see Haufe et al. [2014] and the discussion below).

The final point of discussion is the choice of classifier. In a recent study, we compared the performance of

five classifiers when decoding minimally preprocessed MEG data: LDA, Gaussian naive Bayes (GNB), linear SVMs, Spearman's rank correlation, and Pearson's correlation (Grootswagers, Wardle, & Carlson, 2017). The study found that SVM, LDA, and GNB performed the best, suggesting that they are all excellent choices for M/EEG decoding studies.

Localizing the source of decodable information in M/EEG Cognitive neuroscientists are often interested in both spatial (i.e., where) and temporal (i.e., when) signatures of neural processing. Recovering the underlying source(s) of neural activity is a long-standing challenge for M/EEG research. This challenge extends to M/EEG decoding research, although advanced decoding methods do offer at least one possible solution (see the section on representational dynamics). In this discussion it is essential to make a distinction between resolution and spatial precision. Resolution is the capacity to resolve two points in space. Using MVPA, MEG studies have shown that it may be possible to resolve activation patterns spanning V1 columns that are just a millimeter in width (Cichy, Ramirez, & Pantazis, 2015; Wardle et al., 2016). These results highlight the remarkable sensitivity of MVPA for detecting subtle differences in patterns of activation. However, localizing the precise source of this neural information (i.e., spatial precision) remains an ongoing challenge.

Broadly, three methods can be used to localize sources in the brain using MVPA. Below, we will discuss two of these methods; the third will be presented later in the chapter. The first approach is weight projection. MVPA returns both a performance metric (e.g., percent correct) and the weights used by the classifier to make the prediction. The more informative the sensor, the higher the weight. One straightforward means of identifying the source of decodable information is to plot the weights on the scalp map. When interpreting these maps, it is important to consider the definition of *informative* (c.f., de-Wit et al., 2016; Haufe et al., 2014). Previously, we discussed the three ways classifiers optimize their performance. In addition to weights being assigned to distinguish condition-specific information, weights are also used by the classifier to suppress noise. When using the weight projection method to interpret the underlying sources, it is thus essential to consider only weights that reflect the differences between conditions (Grootswagers, Wardle, & Carlson, 2017; Haufe et al., 2014). The second approach is to extract multivariate brain activity from a region of interest (ROI) for MVPA. The most straightforward variant of this is to select sensors located above the ROI—for example, using ten sensors over the occipital cortex to study early visual processing. This can also be applied in a sensor-searchlight approach (similar to fMRI; see Haynes et al., 2007; Kriegeskorte, Goebel, & Bandettini, 2006), in which the analysis is repeated on local clusters of sensors, providing a scalp map of decoding accuracies (e.g., see Collins, Robinson, & Behrmann, 2018; Kaiser, Oosterhof, & Peelen, 2016).

These two methods are subject to interpretive criticism arising from the fact that brain activity will propagate outside the M/EEG sensors located above a source (information loss), and brain activity from nearby areas will propagate into the selected sensors (information leakage). Nevertheless, it can be a useful means for gross localization (e.g., left vs. right hemisphere) in some circumstances, and the limitations in its interpretability are transparent. Alternatively, the sensor-level data can first be projected into source space using source reconstruction methods such as minimum norm estimate (MNE; Hamalainen & Ilmoniemi, 1994) and beamforming (Hamalainen & Ilmoniemi, 1994; Van Veen et al., 1997). Multivariate time-series data can be reconstructed using these methods using virtual voxels in an ROI, and decoding analyses can be performed for different ROIs. As with univariate analysis, the quality of the source reconstruction can be improved using anatomical MRI data from the participant and fMRI data to precisely define the ROI. Notably, these more sophisticated methods for localization are also subject to issues of information loss and leakage (Brookes, Woolrich, & Barnes, 2012; Gohel et al., 2018; Hipp et al., 2012; Nolte et al., 2004; Sato et al., 2018), but the algorithms will attempt to minimize their effect.

In summary, MVPA is a sensitive measure for differentiating conditions based on the evoked spatial distribution of activity from M/EEG. The MVPA approach, however, remains limited for localizing sources, which stems from the fundamental inverse problem for M/EEG.

Statistical assessment of information at the group level Time-resolved MVPA tests the presence of information at the group level. That is, we want to know whether across our sample of subjects, classifier performance is higher than would be expected by chance. There are multiple proposed methods to assess this, and there is no consensus about the optimal approach. Group-level classifier accuracies can, for example, be tested against chance using a t-test, or nonparametric tests such as a sign-rank test (Wilcoxon, 1945) or permutation test (Oostenveld et al., 2011). Because these tests are repeated at each time point, they must also be corrected for multiple comparisons—for example, using Bonferroni, false discovery rate (FDR; Nichols & Holmes, 2002), or cluster-based corrections (e.g., Oostenveld et al., 2011;

Smith & Nichols, 2009; Stelzer, Chen, & Turner, 2013). As with the choice of test statistic, there is no consensus on the optimal method for correcting multiple comparisons. Thus, the choice of statistical analysis must be guided by the experimental questions (see Allefeld, Gorgen, & Haynes, 2016; Hebart & Baker, 2017; Thirion et al., 2015).

Advanced Methods for M/EEG Decoding: Representational Dynamics

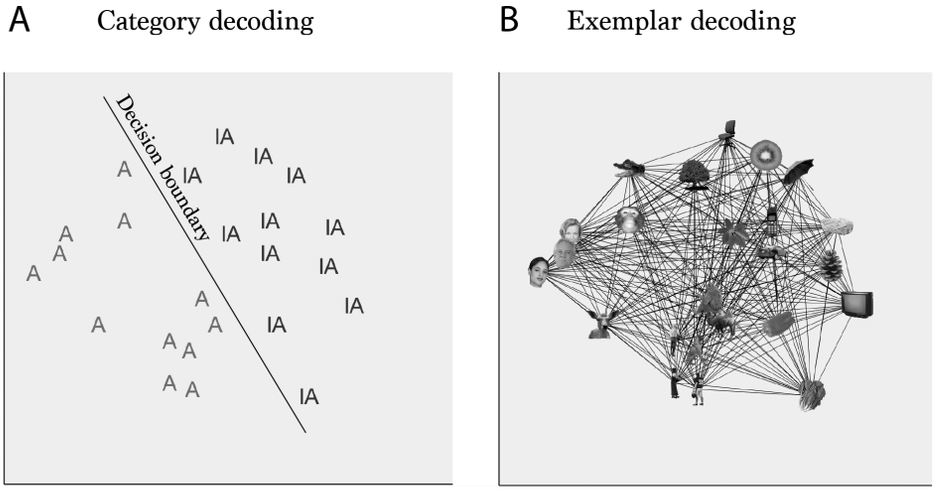
Exemplar-based decoding methods expand on categorical approaches by studying the structure of how information is represented. Figure 57.2*A, B*, shows the difference between the two approaches. The two parts of the figure show a reconstruction of the brain's representation of 24 objects from MEG data 140 ms after the onset of the stimulus (data from Carlson et al., 2013). The stimulus set included 12 animate objects (e.g., camel and alligator) and 12 inanimate objects (e.g., chair and kiwi fruit). In the category decoding approach (figure 57.2*A*), the stimuli are treated as equivalent class members (labeled *A* = animate or *IA* = inanimate), and the analysis is conducted using standard methods (see the section on the basics of time-resolved decoding). If the analysis shows the classifier can decode animate and inanimate objects from the MEG data, this is evidence that the brain representation of the stimuli is encoding object animacy.

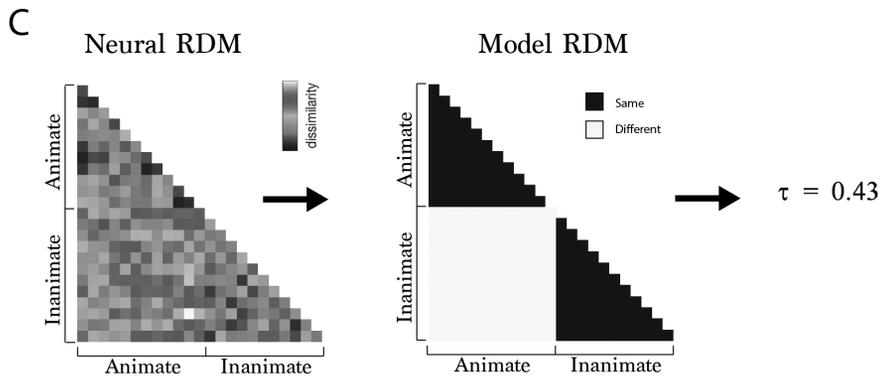
Exemplar-based decoding methods study how the individual exemplars of each category are represented. Here, the decoding analysis is run for all possible pairwise combinations of exemplars—including within-category. The decodability (i.e., classification performance) of each pair is taken as a measure of "distance" between the two items in the neural representation (Walther et al., 2016). Figure 57.2*B* demonstrates this approach to plotting the individual stimuli. A line connecting the object exemplars in the figure indicates the decodability of the connected pair. The length of the line is proportional to decodability—that is, stimuli that are close to one another (a human and a monkey face) are less decodable than stimuli that are far apart (e.g., a monkey and a television) at the given point in time during visual processing.

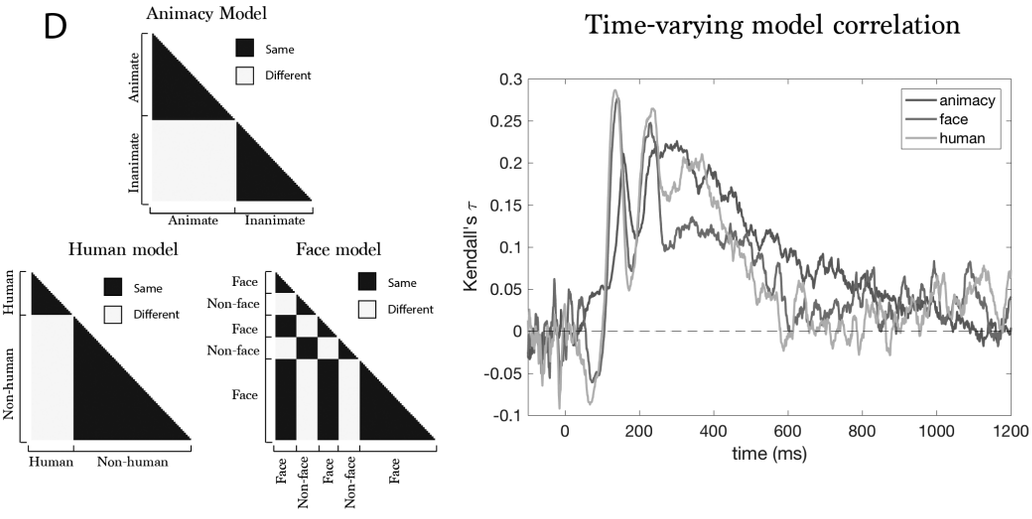
Representational similarity analysis (RSA; Kriegeskorte & Kievit, 2013; Kriegeskorte, Mur, & Bandettini, 2008) is the standard framework for hypothesis testing for the exemplar-based decoding approach. In the RSA, the pairwise decodability of the stimuli from the brain recordings is encoded in a representational dissimilarity matrix (RDM; figure 57.2*C*). The rows and columns of the matrix correspond to individual exemplars, and each cell is the neural decodability between the two exemplars. Hypotheses are tested by constructing model RDMs that make predictions about the structure of the brain representation. An animacy model, for example, predicts that animate and inanimate objects form separable clusters in the representation. Formally, the animacy model predicts that objects within the animate and inanimate object categories will be close to one another (distance = 0), and objects that span the category boundary will be far apart (distance = 1; see figure 57.2*C*). To test the model, the entries of the observed neural RDM are correlated with the model RDM. For the animacy model, a high correlation between the neural RDM and the model allows us to conclude that animacy is represented in the brain.

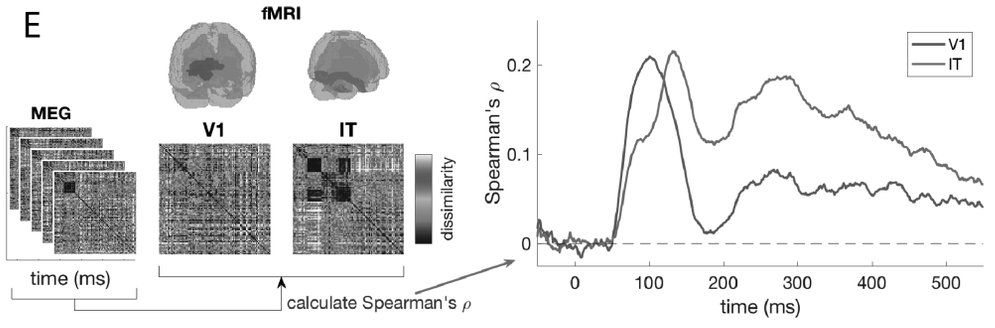
The exemplar-based decoding approach taken in RSA makes the structure of the representation and the models more explicit, which has many advantages (Kriegeskorte & Kievit, 2013; Kriegeskorte, Mur, & Bandettini, 2008; Nili et al., 2014). The exemplar-based analysis can be used to study the time-varying structure of brain representations using M/EEG. As before, the analysis and model testing are performed on all of the time points. By examining how brain representations emerge over time, we gain a deeper understanding of how information is dynamically transformed and represented in the brain. Figure 57.2*D* shows the time-varying analysis from a MEG data set with a more extensive stimulus set using several different category models (92 exemplars; Cichy, Pantazis, & Oliva, 2014). The figure shows that category structure for different categories (animacy, human, face) emerge at different time points in the brain's emerging representation of objects. By studying the emergence of different categories in time, studies have shown that basic-level category information emerges first (e.g., human face), and abstract-level category information (e.g., animacy) comes at the later stages of object processing (Carlson et al., 2013; Cichy, Pantazis, & Oliva, 2014; Contini, Wardle, & Carlson, 2017).

Exemplar-based decoding methods and the RSA framework also provide a unique solution to the challenge of acquiring data that have the spatial resolution to study regional brain activity and the temporal resolution to explore the fine-grained temporal dynamics of neural activity (Cichy, Pantazis, & Oliva, 2014, 2016). This novel approach takes advantage of the fine-grained structure of information in brain representations, as indexed by MEG and fMRI, and integrates the data to estimate regional brain activity. The fMRI and MEG experiments are run using identical stimuli; thus, the two imaging methods produce RDMs of equal size that are directly comparable. The MEG RDM describes the brain representation at each time point, which includes activity from multiple brain regions. The fMRI RDM









from an ROI describes that brain area's representation. Importantly, different brain areas have unique representations, although some areas will be similar, such as V1 and V2. By correlating a brain area's fMRI RDM with the MEG RDM for each time point, we get a time-varying estimate of that area's contribution to the MEG signal across time. This method was used to study the contribution of early visual cortex (V1) and inferior temporal cortex (IT) to the MEG signal in time for object recognition (figure 57.2*E*; Cichy, Pantazis, & Oliva, 2014).

Conclusion

M/EEG decoding methods provide a powerful set of tools for cognitive neuroscientists to gain insight into perceptual and cognitive functions by revealing the brain's processing dynamics with millisecond resolution. In this chapter we described the fundamental aspects of M/EEG decoding analysis, practical considerations in running these analyses, and advanced methods to study representational dynamics. These methods have been incorporated into a variety of MVPA toolboxes (Bode et al., 2018; Fahrenfort et al., 2018; Gramfort et al., 2014; Hanke et al., 2009; Meyers, 2013; Oostenveld et al., 2011; Oosterhof, Connolly, & Haxby, 2016). We expect M/EEG decoding to have a broad impact on the future of cognitive neuroscience research. For researchers interested in taking the next step and learning more about these methods, we refer the reader to advanced tutorials (Grootswagers, Wardle, & Carlson, 2017; Lemm et al., 2011).

Acknowledgments

We would like to thank the reviewers and Nick McNair, Denise Moerel, and Selene Petit for their feedback and suggestions.

REFERENCES

Allefeld, C., Gorgen, K., & Haynes, J. D. (2016). Valid population inference for information-based imaging: From the second-level t-test to prevalence inference. *Neuroimage, 141*, 378–392.

Averbeck, B. B., Latham, P. E., & Pouget, A. (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience, 7*, 358–366.

Bode, S., Feuerriegel, D., Bennett, D., & Alday, P. M. (2018). The decision decoding ToolBOX (DDTBOX)—a multivariate pattern analysis toolbox for event-related potentials. *Neuroinformatics, 17*(1), 27–42.

Brookes, M. J., Woolrich, M. W., & Barnes, G. R. (2012). Measuring functional connectivity in MEG: A multivariate approach insensitive to linear source leakage. *Neuroimage, 63*, 910–920.

Carlson, T., Goddard, E., Kaplan, D. M., Klein, C., & Ritchie, J. B. (2017). Ghosts in machine learning for cognitive neuroscience: Moving from data to theory. *Neuroimage, 180*, 88–100.

Carlson, T., Tovar, D. A., Alink, A., & Kriegeskorte, N. (2013). Representational dynamics of object vision: The first 1000 ms. *Journal of Vision, 13*(10), 1–19.

Cichy, R. M., Pantazis, D., & Oliva, A. (2014). Resolving human object recognition in space and time. *Nature Neuroscience, 17*, 455–462.

Cichy, R. M., Pantazis, D., & Oliva, A. (2016). Similarity-based fusion of MEG and fMRI reveals spatio-temporal dynamics in human cortex during visual object recognition. *Cerebral Cortex, 26*, 3563–3579.

Cichy, R. M., Ramirez, F. M., & Pantazis, D. (2015). Can visual information encoded in cortical columns be decoded from magnetoencephalography data in humans? *Neuroimage, 121*, 193–204.

Collins, E., Robinson, A. K., & Behrmann, M. (2018). Distinct neural processes for the perception of familiar versus unfamiliar faces along the visual hierarchy revealed by EEG. *Neuroimage, 181*, 120–131.

Contini, E. W., Wardle, S. G., & Carlson, T. A. (2017). Decoding the time-course of object recognition in the human brain: From visual features to categorical decisions. *Neuropsychologia, 105*, 165–176.

Davatzikos, C., Ruparel, K., Fan, Y., Shen, D. G., Acharyya, M., Loughead, J. W., Gur, R. C., & Langleben, D. D. (2005). Classifying spatial patterns of brain activity with machine

FIGURE 57.2 Advanced methods for M/EEG decoding. *A, B*, Plots demonstrate the difference between category and exemplar-based decoding methods. Data for the plots come from a single time point (140 ms poststimulus onset) of an MEG experiment studying the brain's representation of objects (Carlson et al., 2013). *A*, Category-decoding approach with stimuli labeled by their category (A = animate object; IA = inanimate object) and the decision boundary from a linear classifier trained to discriminate animate and inanimate objects from the MEG data. *B*, Exemplar-decoding approach with individual stimuli displayed as images. The lines in the plot represent the pairwise decodability of individual exemplars for all possible pairwise comparisons. The distance (*line length*) indicates the relative decodability of exemplar pairs. *C*, Representational similarity (RSA) model testing applied to the MEG data. The entries of the neural RDM are correlated with the animacy model to study whether the brain represents object animacy at 140 ms poststimulus onset. *D*, RSA used to test multiple candidate models from another MEG experiment investigating the brain's representation of objects. Data from Cichy, Pantazis, and Oliva (2014). Plotted is the time correlation between the time-varying RDMs from the MEG and animacy, human, and face models. *E*, Graphical depiction of the MEG-fMRI fusion approach estimating regional time-varying neural activity from fMRI regions of interest (ROI). The time-varying MEG RDMs are correlated with RDMs from ROIs (obtained using fMRI). For each ROI, this gives a time-varying correlation indexing neural activity in the ROI. Shown is the estimated neural activity for V1 and inferior temporal cortex. Data from Cichy, Pantazis, and Oliva (2014). (See color plate 63.)

learning methods: Application to lie detection. *Neuroimage, 28*, 663–668.
de-Wit, L., Alexander, D., Ekroll, V., & Wagemans, J. (2016). Is neuroimaging measuring information in the brain? *Psychonomic Bulletin & Review, 23*, 1415–1428.
DiCarlo, J. J., & Cox, D. D. (2007). Untangling invariant object recognition. *Trends in Cognitive Sciences, 11*, 333–341.
Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (2nd ed.). New York: Wiley.
Ewers, M., Sperling, R.A., Klunk, W. E., Weiner, M. W., & Hampel, H. (2011). Neuroimaging markers for the prediction and early diagnosis of Alzheimer's disease dementia. *Trends in Neurosciences, 34*, 430–442.
Fahrenfort, J. J., van Driel, J., van Gaal, S., & Olivers, C. N. L. (2018). From ERPs to MVPA using the Amsterdam decoding and modeling toolbox (ADAM). *Frontiers in Neuroscience, 12*, 368.
Gohel, B., Lim, S., Kim, M. Y., Kwon, H., & Kim, K. (2018). Dynamic pattern decoding of source-reconstructed MEG or EEG data: Perspective of multivariate pattern analysis and signal leakage. *Computers in Biology and Medicine, 93*, 106–116.
Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., & Hamalainen, M. S. (2014). MNE software for processing MEG and EEG data. *Neuroimage, 86*, 446–460.
Grootswagers, T., Ritchie, J. B., Wardle, S. G., Heathcote, A., & Carlson, T. A. (2017). Asymmetric compression of representational space for object animacy categorization under degraded viewing conditions. *Journal of Cognitive Neuroscience, 29*, 1995–2010.
Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. *Journal of Cognitive Neuroscience, 29*, 677–697.
Hamalainen, M. S., & Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: Minimum norm estimates. *Medical & Biological Engineering & Computing, 32*, 35–42.
Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. (2009). PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics, 7*, 37–53.
Haufe, S., Meinecke, F., Gorgen, K., Dahne, S., Haynes, J. D., Blankertz, B., & Biessmann, F. (2014). On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage, 87*, 96–110.
Haxby, J. V. (2012). Multivariate pattern analysis of fMRI: The early beginnings. *Neuroimage, 62*, 852–855.
Haynes, J. D. (2015). A primer on pattern-based approaches to fMRI: Principles, pitfalls, and perspectives. *Neuron, 87*, 257–270.
Haynes, J. D., Sakai, K., Rees, G., Gilbert, S., Frith, C., & Passingham, R. E. (2007). Reading hidden intentions in the human brain. *Current Biology, 17*, 323–328.
Hebart, M. N., & Baker, C. I. (2017). Deconstructing multivariate decoding for the study of brain function. *Neuroimage, 180*, 4–18.
Hipp, J. F., Hawellek, D. J., Corbetta, M., Siegel, M., & Engel, A. K. (2012). Large-scale cortical correlation structure of spontaneous oscillatory activity. *Nature Neuroscience, 15*, 884–890.
Kaiser, D., Oosterhof, N. N., & Peelen, M. V. (2016). The neural dynamics of attentional selection in natural scenes. *Journal of Neuroscience, 36*, 10522–10528.
Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience, 8*, 679–685.
Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302–4311.
Kriegeskorte, N. (2011). Pattern-information analysis: From stimulus decoding to computational-model testing. *Neuroimage, 56*, 411–421.
Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 3863–3868.
Kriegeskorte, N., & Kievit, R. A. (2013). Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences, 17*, 401–412.
Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis—connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience, 2*, 4.
Lemm, S., Blankertz, B., Dickhaus, T., & Muller, K. R. (2011). Introduction to machine learning for brain imaging. *Neuroimage, 56*, 387–399.
Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
Meyers, E. M. (2013). The neural decoding toolbox. *Frontiers in Neuroinformatics, 7*, 8.
Misaki, M., Kim, Y., Bandettini, P. A., & Kriegeskorte, N. (2010). Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *Neuroimage, 53*, 103–118.
Muller, K. R., Anderson, C. W., & Birch, G. E. (2003). Linear and nonlinear methods for brain-computer interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 11*, 165–169.
Mur, M., Bandettini, P. A., & Kriegeskorte, N. (2009). Revealing representational content with pattern-information fMRI—an introductory guide. *Social Cognitive and Affective Neuroscience, 4*, 101–109.
Nichols, T. E., & Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: A primer with examples. *Human Brain Mapping, 15*, 1–25.
Nili, H., Wingfield, C., Walther, A., Su, L., Marslen-Wilson, W., & Kriegeskorte, N. (2014). A toolbox for representational similarity analysis. *PLOS Computational Biology, 10*, e1003553.
Nolte, G., Bai, O., Wheaton, L., Mari, Z., Vorbach, S., & Hallett, M. (2004). Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical Neurophysiology, 115*, 2292–2307.
Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience, 2011*, 156869.
Oosterhof, N. N., Connolly, A. C., & Haxby, J. V. (2016). CoSMoMVPA: Multi-modal multivariate pattern analysis of neuroimaging data in Matlab/GNU octave. *Frontiers in Neuroinformatics, 10*, 27.

Pereira, F., Mitchell, T., & Botvinick, M. (2009). Machine learning classifiers and fMRI: A tutorial overview. *Neuroimage, 45*, S199–209.

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafo, M. R., Nichols, T. E., Poline, J. B., Vul, E., & Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience, 18*, 115–126.

Sato, M., Yamashita, O., Sato, M. A., & Miyawaki, Y. (2018). Information spreading by a combination of MEG source estimation and multivariate pattern classification. *PLoS One, 13*, e0198806.

Schwarzkopf, D. S., & Rees, G. (2011). Interpreting local visual features as a global shape requires awareness. *Proceedings of the Royal Society B: Biological Sciences, 278*, 2207–2215.

Smith, S. M., & Nichols, T. E. (2009). Threshold-free cluster enhancement: Addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage, 44*, 83–98.

Stelzer, J., Chen, Y., & Turner, R. (2013). Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. *Neuroimage, 65*, 69–82.

Thirion, B., Pedregosa, F., Eickenberg, M., & Varoquaux, G. (2015). Correlations of correlations are not reliable statistics: Implications for multivariate pattern analysis. ICML Workshop on Statistics, Machine Learning and Neuroscience (Stamlins), Lille, France.

Tong, F., & Pratte, M. S. (2012). Decoding patterns of human brain activity. *Annual Review of Psychology, 63*, 483–509.

Van Veen, B. D., van Drongelen, W., Yuchtman, M., & Suzuki, A. (1997). Localization of brain electrical activity via linearly constrained minimum variance spatial filtering. *IEEE Transactions on Biomedical Engineering, 44*, 867–880.

Vidal, J. J. (1973). Toward direct brain-computer communication. *Annual Review of Biophysics and Bioengineering, 2*, 157–180.

Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of dissimilarity measures for multi-voxel pattern analysis. *Neuroimage, 137*, 188–200.

Wardle, S. G., Kriegeskorte, N., Grootswagers, T., Khaligh-Razavi, S. M., & Carlson, T. A. (2016). Perceptual similarity of visual patterns predicts dynamic neural activation patterns measured with MEG. *Neuroimage, 132*, 59–70.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin, 1*, 80–83.

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. *Clinical Neurophysiology, 113*, 767–791.