Why More *Is* Better: Simultaneous Modeling of EEG, fMRI, and Behavioral Data

Brandon M. Turner[a,*], Christian A. Rodriguez[b], Tony M. Norcia[b], Samuel M. McClure[c], Mark Steyvers[d]

[a]*Department of Psychology, The Ohio State University*
[b]*Department of Psychology, Stanford University*
[c]*Department of Psychology, Arizona State University*
[d]*Department of Cognitive Science, University of California, Irvine*

Abstract

The need to test a growing number of theories in cognitive science has led to increased interest in inferential methods that integrate multiple data modalities. In this manuscript, we show how a method for integrating three data modalities within a single framework provides (1) more detailed descriptions of cognitive processes and (2) more accurate predictions of unobserved data than less integrative methods. Specifically, we show how combining either EEG and fMRI with a behavioral model can perform substantially better than a behavioral-data-only model in both generative and predictive modeling analyses. We then show how a trivariate model – a model including EEG, fMRI, and behavioral data – outperforms bivariate models in both generative and predictive modeling analyses. Together, these results suggest that within an appropriate modeling framework, more data can be used to better constrain cognitive theory, and to generate more accurate predictions for behavioral and neural data.

Keywords: joint modeling framework, Bayesian modeling, EEG, fMRI, Linear Ballistic Accumulator model

*Corresponding Author
Email address: turner.826@gmail.com (Brandon M. Turner)

This research was supported by National Science Foundation grant 1358507. We would like to thank Nathan Fackler for helpful discussions that improved an earlier draft of this manuscript.

1. Introduction

The field of cognitive science is faced with many options for studying how experimentally-derived variables are systematically related to the dynamics underlying a cognitive process of interest. To date, much of our understanding of cognition has been advanced by two dominant, but non-interacting groups. The largest group, cognitive neuroscientists, rely on statistical models to understand patterns of neural activity. These models are typically purely data-mining techniques, and often disregard the computational mechanisms that might detail a cognitive process. The other group, mathematical psychologists, is strongly motived by *theoretical* accounts of cognitive processes, and instantiates these theories by developing formal mathematical models of cognition. The models often assume a system of computations and equations intended to characterize the process assumed to take place in the brain. To formally test their theory, mathematical psychologists rely on their model's ability to fit behavioral data. A good fit is thought to reflect an accurate theory, whereas a bad fit would refute it.

Although both groups are concerned with explaining behavior, they tend to approach the challenge from different vantage points. Thinking in terms of Marr (1982)'s levels of analysis, mathematical psychologists tend to focus on the computational and algorithmic levels, whereas cognitive neuroscientists focus more on the implementation level. Although progress can be made by maintaining a tight focus on one level, certain opportunities are missed. As a result of their single-level focus, both approaches suffer from critical limitations (Love, 2015), and these limitations have inspired researchers to combine neural and behavioral measures in an integrative fashion. The primary payoff of this endeavor is the localization of mechanisms assumed to underly the computations supporting task-specific behavior within the brain. The importance of solving the integration problem has spawned several entirely new statistical modeling approaches developed through collaborations between mathematical psychologists and cognitive neuroscientists, collectively forming a new field often referred to as "model-based cognitive neuroscience" (e.g., Forstmann and Wagenmakers, 2015; Daw, 2011; Forstmann et al., 2011b; van Maanen et al., 2011; Turner et al., 2013a; Mack et al., 2013; Boehm et al., 2014; Love, 2015; Palmeri et al., 2015; Turner et al., 2015b).

So far, the field of model-based cognitive neuroscience has established links between both structural measures such as diffusion-weighted imaging (Forstmann et al., 2010; Turner et al., 2013a), and functional measures such

as electroencephalography (EEG; e.g., Nunez et al., 2015), functional magnetic resonance imaging (fMRI; e.g., van Maanen et al., 2011; Mulder et al., 2012; Summerfield and Koechlin, 2010; Bai et al., 2007; Forstmann et al., 2008, 2010; White et al., 2014; Mack et al., 2013; Turner et al., 2015b) and magnetoencephalography (e.g., Amano et al., 2006). These efforts, while elucidating, have focused on relating model parameters to one modality of neural measure. However, the limitations of these modalities in their application to understanding cognition are well known. Superficially, measures like EEG provide excellent temporal resolution with a continuous read out of brain activity propagated to the scalp. These measures are well suited for studying fine-grained details of a decision process where many brain regions interact to perform a given task, but they are limited by the specificity of spatial information they can provide. On the other hand, measures like fMRI provide good spatial resolution, but offer limited information on the temporal properties of the interesting neural signature.

Given the orthogonal nature of the benefits and limitations of EEG and fMRI, it is a natural question to ask whether these measures can be combined in an effort to capitalize on the temporal and spatial resolutions provided by each modality. There currently exist several methods for "fusing" multimodal functional neuroimaging data together, such as partial least squares correlation (Lin et al., 2003; Martinez-Montes et al., 2004), independent component analysis (Beckmann and Smith, 2005; Liu and Calhoun, 2007; Calhoun et al., 2006, 2009; Eichele et al., 2009; Franco et al., 2008; Teipel et al., 2010; Xu et al., 2009; Calhoun et al., 2011; Calhoun and Adali, 2009), structural equation modeling (Astolfi et al., 2004; Hamandi et al., 2008), multiple regression (De Martino et al., 2010; Eichele et al., 2005), and canonical correlation analysis (Correa et al., 2010a,b, 2008). However, the focus of these methods is either source localization (for reviews, see Sui et al., 2012; Dähne et al., 2015), or relating behavioral measures (e.g., response times) to brain data (for a review, see Krishnan et al., 2015). While these methods have elucidated several theoretically interesting questions (for reviews of applications, see Krishnan et al., 2015; Sui et al., 2012), they are unable to directly inform cognitive models in a quantitative fashion.

The goal of this article is to develop a statistical method for linking neural data from EEG and fMRI to the (latent) parameters of a cognitive model. To accomplish this, we extend the recently-proposed joint modeling framework (Turner et al., 2013a) to multimodal functional measures. Our method is unique because it allows a theoretical account of the cognitive process to

drive the analysis of the neural and behavioral measures. Furthermore, our method is holistic, putting forth a single model to account for all aspects of our data. As we will detail below, such a strategy provides numerous advantageous on both generative and predictive modeling fronts.

We begin by first describing the joint modeling framework and explaining how it can be extended to multiple neural measures. We then describe an experiment on an intertemporal choice task that was collected for the purposes of testing the framework on imbalanced experimental designs. To provide clarity about our modeling approach, we first describe the experimental effects present in the EEG, fMRI, and behavioral data separately. We then illustrate some of the benefits of the joint modeling framework by showing how a "bivariate" model taking into account only one functional measure can be used to boost performance of a behavioral-data-only model. In all of our analyses, we show this boost from both a generative modeling perspective through analyses of the estimated parameters (i.e., posterior distributions), and a predictive modeling perspective by assessing the models' performance in a cross-validation test. We then show how the extended "trivariate" model that incorporates both functional neural measures further enhances the performance of the model relative to both bivariate models and the behavioral-data-only model. We close with a discussion of limitations and future directions for our approach to integrating data from multiple modalities.

1.1. Joint Modeling

Given the technical hurdles involved in simultaneous EEG/fMRI, in this article we propose an alternative for performing EEG/fMRI integration. Instead of obtaining the neural measures simultaneously, our method works around potential artifacts by simply aggregating the effects of interest across experimental contexts. That is, our method explicitly models changes in neural activity present in EEG and fMRI by binding these effects to changes in the behavioral data. In this way, the behavioral data serve as a bridge connecting the two modalities to one another, even if the EEG and fMRI data were obtained at different points in time for the same subject, or say EEG data was collected for some subjects and fMRI data was collected for others.

Our method builds on a recently developed joint modeling framework (Turner et al., 2013a). This framework was originally proposed as a way to integrate neural and behavioral data in one cognitive model, but in Turner et

al. only structural data were considered. Here, we extend the original framework to functional measures and to trivariate scenarios (i.e., including EEG, fMRI, and behavior). A typical joint model consists of three components, which we will discuss in turn.

First, the behavioral data B are described in terms of a behavioral model, such as the Linear Ballistic Accumulator (LBA; Brown and Heathcote, 2008) model or classic signal detection theory model (Green and Swets, 1966). Importantly, the behavioral model consists of a set of model parameters θ, which are of lower dimensionality than the full behavioral data set B. For example, in the SDT model, θ would consist of parameters such as discriminability (d') and bias (β). Ideally, the chosen behavioral model should consist of mechanisms that provide insight into how the data arise, and how experimental manipulations predictably affect these mechanisms (a property known as selective influence; Heathcote et al., 2015).

Second, the neural data N are described in terms of a neural model. The key property of the neural model is that it should consist of a set of parameters δ that describe the important parts of the neural data N in a way that is of significantly lower dimensionality. For example, the neural model could describe increases in neural activity across time through a general linear model (GLM). Alternatively, the neural model could also be of a more mechanistic nature, describing the computations purportedly implemented by a brain region.

Third, the lower-dimensional features of both the neural and behavioral data are linked by establishing an explicit relationship between the behavioral model parameters θ and the neural model parameters δ. Although there are many types of linking functions one could use, one particularly successful linking approach has been a hierarchical multivariate Gaussian distribution (Turner et al., 2013a; Turner, 2015; Turner et al., 2015b).

The joint modeling framework extends other exploratory type approaches in model-based cognitive neuroscience (see Turner et al., 2015a) by formalizing these three components into a single model. Furthermore, fitting the joint model via Bayesian statistics allows the model to inherit many of the advantages of Bayesian modeling, which have been discussed by many authors (e.g., Shiffrin et al., 2008; Lee and Wagenmakers, 2013). In the context of model-based cognitive neuroscience, most of these advantages center around flexibility. First, joint models are agnostic – there is no commitment to any particular behavioral model, neural model, or linking function, and similarly, there is no commitment to any particular behavioral (e.g., response time,

accuracy, confidence) or neural (e.g., EEG, fMRI) measurements. Second, the parameter estimates for the behavioral model are affected by the information contained in the neural measurements. Because this is a hierarchical Bayesian model, information from the neural measures first affects the neural model parameters, and if this effect is consistent across subjects (or even trials, see Turner et al., 2015b) it affects the parameters of the linking function. Once the linking function has been affected, the behavioral model parameters are systematically altered to account for these changes. This process also goes in the other direction, where changes in the behavioral model parameters affect the neural model parameters. Third, joint models are fit to data within one single regime instead of using separate parameter estimation stages in an exploratory manner (e.g., Forstmann et al., 2010, 2008, 2011a; Ho et al., 2012; van Maanen et al., 2011). Furthermore, joint models are unique in that they do not require elaborate retraining or refitting strategies to generate predictions for either neural or behavioral measures. Fourth, due to the Bayesian component of the estimation process, joint models can naturally handle sparse, imbalanced, or missing data structures. This aspect is extremely useful when the number of observations that can be obtained is limited, or the cost of obtaining neural measures is high. Given all of these advantages, joint models seem ideally suited to tackle the problem of integrating behavioral, EEG, and fMRI measures into one cohesive model.

Figure 1 depicts a graphical diagram of the models we will investigate in this article. Graphical diagrams are useful in making variable dependencies explicit by illustrating their relationship through connections in the graph. Observable variables are represented as gray boxes (e.g., B), whereas unknown (latent) model parameters are shown as empty circles (e.g., θ). Although all of the parameters influence one another when fitting the model to data, only explicit dependencies among variables in the graph are represented by the arrows. So, for example, the graphical diagram shows that the behavioral data B is a function of the behavioral model parameters θ, which are functions of other model parameters Ω known as *hyper* parameters. The orange plate represents the behavioral data/model, the green plate represents the EEG data/model, and the blue plate represents the fMRI data/model. All of the models we will use in the present article can be represented via some combination of the plates in Figure 1. In previous work (Turner et al., 2013a), we made the connection between a single modality of neural data and the behavioral data explicit. So, for example, the orange plate was connected to the green plate. In the present article, the goal is to formally extend this

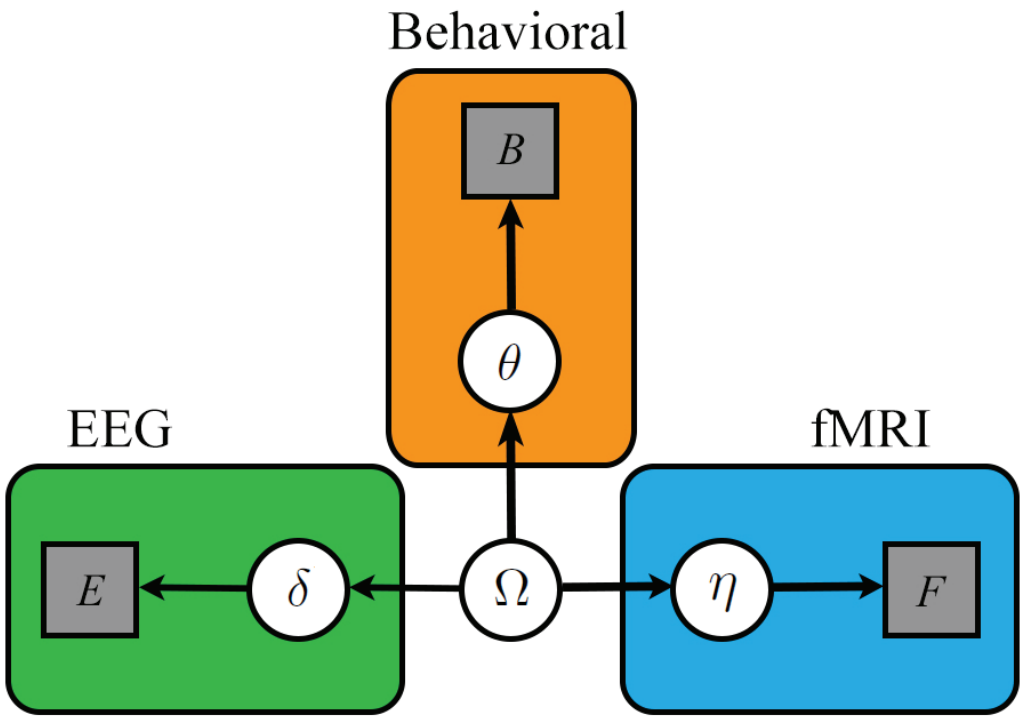

Figure 1: Graphical diagram of models presented in the text. Observable data are represented as gray boxes, whereas unknown (latent) variables are represented as empty circles. The orange plate represents the behavioral data/model, the green plate represents the EEG data/model, and the blue plate represents the fMRI data/model.

framework to two modalities of neural data, namely fMRI and EEG. The motivation for our endeavor is clear: by combining the temporal resolution of EEG with the spatial resolution of fMRI, the link between multiple measures of cognition (i.e., EEG, fMRI, and behavior) can be better understood.

We begin by describing our three experiments, specifically detailing the imbalance of our task design. We reiterate the basic patterns present in our data, some of which have been detailed extensively for other purposes (Rodriguez et al., 2014, 2015b,a). To better illustrate some of the advantages joint modeling provides, we begin our analyses by applying bivariate joint models to each data modality separately. In the final section, we apply a trivariate joint model to the full (imbalanced) data set and highlight many of the advantages that are realized when EEG, fMRI, and behavioral data

are fused. In each joint modeling application, we evaluate the model from a generative and predictive perspective, showing that not only do joint models help in the interpretation of our data, they also improve the predictive performance of the behavioral model.

2. Materials and Methods

In total, we conducted three separate intertemporal choice experiments, collecting data on three important variables: behavior, EEG, and fMRI. However, because we originally had no intention of integrating these three measures *simultaneously*, some experiments only targeted two of the three measures. The structure of our data is illustrated in Figure 2. In all, we have 54 subjects (rows) and three measures (columns). The first experiment consisted of 23 subjects who provided both behavioral and EEG data. In this case, the fMRI data can be thought of as "missing", and this arrangement is illustrated in Figure 2 with red cells. The second experiment also consisted of 23 subjects, but these subjects provided both behavioral and fMRI data, and no EEG data (again illustrated by red blocks in Figure 2). The final experiment consisted of 8 subjects who provided both EEG and fMRI data in two separate conditions (randomly counterbalanced). For each of these conditions, we also collected behavioral data, and this structure is illustrated in Figure 2 as striped cells. In the modeling results below, it was useful to withhold some elements of the data selectively as a test of the model's generalizability. These withheld data are illustrated in Figure 2 as orange cells.

2.1. Data collection

2.1.1. Subjects

Sixty-one healthy adults participated in this study (32 females, ages 19-46 years, median 24 years). All subjects gave written informed consent. Stanford University's Institutional Review Board approved the study. Two subjects were excluded because their behavior did not allow us to estimate reliable temporal discounting parameters. Another five subjects were excluded because of data collection problems. Data from a total of fifty-four subjects were analyzed (30 females, ages 19-46 years, median 25 years).

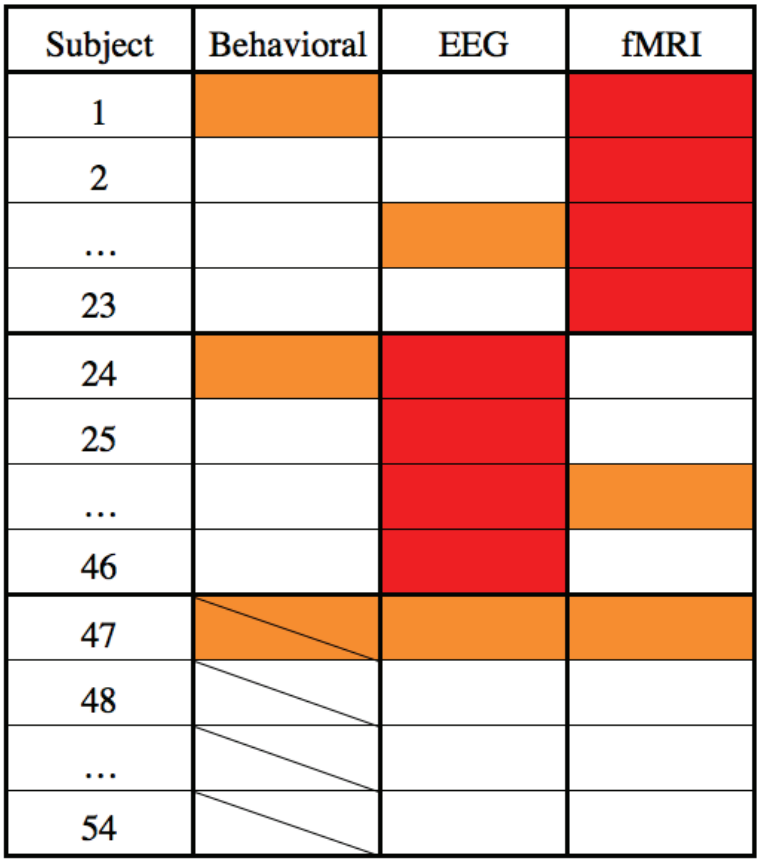

Figure 2: Design matrix for the trivariate model simulation. Our data consist of behavior, EEG, and fMRI measures (columns) for different combinations of subjects (rows). The red cells represent observations that were not recorded, the orange cells represent observations that were withheld from the model in various simulation studies (see text for details), and striped cells represent dual-purpose observations (i.e., we collected behavioral measures during both fMRI and EEG sessions).

2.1.2. Temporal discounting model and task design

The first two experiments were conducted over two sessions, whereas the third experiment consisted of three sessions. The purpose of the first session was to estimate each individual's discount rate using a hyperbolic discounting model. In the first experiment ($n = 23$), the second session consisted of an electroencephalography (EEG) experiment. In the second experiment ($n = 23$), the second session consisted of a functional magnetic resonance imaging (fMRI) experiment. In the third experiment ($n = 8$), subjects performed in both the EEG and fMRI experiments in sessions separated by one week. The purpose of the one week lag was to reduce any practice effect that might occur, while assuming only small deviations in the temporal discounting behavior (Kirby, 2009). In the first session, we began by estimating each individual's discount rate, exactly as in the first two experiments. However, during the third session, we assumed this discount rate was equivalent to what had been estimated one week prior. The task for the first session used a staircase procedure to measure each individual's discount rate k, assuming a hyperbolic discounting function

$$V_D = \frac{r}{1 + kt} \tag{1}$$

where V_D is the subjective value of the delayed reward, r is the monetary amount offered, and t is the delay. The staircase procedure required participants to select between a delayed reward (of r dollars available at delay t) and a fixed immediate reward of \$10 (V_I). For any choice, indifference between the immediate and delayed options implies a discount rate of $k = (r - V_I)(V_I t)^{-1}$. We refer to this implied equivalence point as k_{eq}; our procedure amounted to varying k_{eq} systematically until indifference was reached. Specifically, we began with k_{eq}=0.02. If the subject chose the delayed reward, k_{eq} decreased by a step size of 0.01 for the next trial. Otherwise, k_{eq} increased by the same amount. Every time the subject chose both a delayed and an immediate offer within five consecutive trials, the step size was reduced by 5%. Participants completed 60 trials of this procedure. We placed no limits on the response time, and presented both offers on the screen, as "\$10 now" on the left side, and "\$$r$ in t days" on the right.

After completing the first session, we fit a softmax decision function to participants' choices. We assumed that the likelihood of choosing the delayed reward was given by

$$P_D = \frac{1}{1 + e^{-m(V_D - V_I)}}, \tag{2}$$

where m accounts for sensitivity to changes in discounted value.

We used individually determined values of k and m to generate choices for the second session. At every trial, t was randomly selected from a range of 30-45 days. We then calculated and offered an amount r that would give P_D of 0.1, 0.3, 0.5, 0.7, or 0.9 (Figure 3a-b). The EEG group completed 30 trials at every P_D level, except at P_D=0.5, for which they completed 60 trials. The fMRI group completed 40 trials at every P_D level, except at P_D=0.5, for which they completed 80 trials. Non-uniform trial distributions as a function of P_D were introduced to allow us to study the effects of choice difficulty on EEG and fMRI measures, with equal numbers of trials at each difficulty level. The results of these analyses have been reported elsewhere (Rodriguez et al., 2014, 2015b,a). Trial types were randomized and counterbalanced over two blocks for the EEG group and over four blocks for the fMRI group. We also counterbalanced the mapping between choices and button presses for every subject. During the first half of the second session, approximately half of subjects (17 in EEG, 15 in fMRI) indicated choices of the delayed reward by pressing a button with their left index finger and immediate choices by pressing a different button with their right index finger. The other subjects indicated their choices by the inverse left-right mapping. All subjects switched the initial response mapping during the second half of the session.

We used a sequential presentation of delay and amount during the second session (Figure 3c). The rationale for this sequential presentation has been explained elsewhere (Rodriguez et al., 2014, 2015b). We report RT as measured from the onset of the decision period, 1000ms into the trial. The duration of the decision period was fixed at 4000ms. When subjects made choices in less than 4000ms the amount information disappeared and the screen remained blank until 4000ms elapsed. Trial length was thus fixed at 5000ms. We discarded any trial in which a response was made in less than 200ms or fell outside of the decision period. To optimize experimental time and separability of neural signals across trials for both groups, we introduced a long inter-trial-interval (ITI) for the fMRI sessions (between 4-10s), and a shorter ITI for the EEG sessions (100-350ms). In exchange for participation subjects received \$10 cash for the EEG session or \$20 for the fMRI session, plus an additional amount, determined by their choice in a randomly selected trial.

EEG data was collected using a 128 channel Geodesic Sensor Net (Electrical Geodesics, Inc., Eugene OR, USA), with a 500 Hz sampling rate, using the vertex as reference. During pre-processing, we re-referenced to the

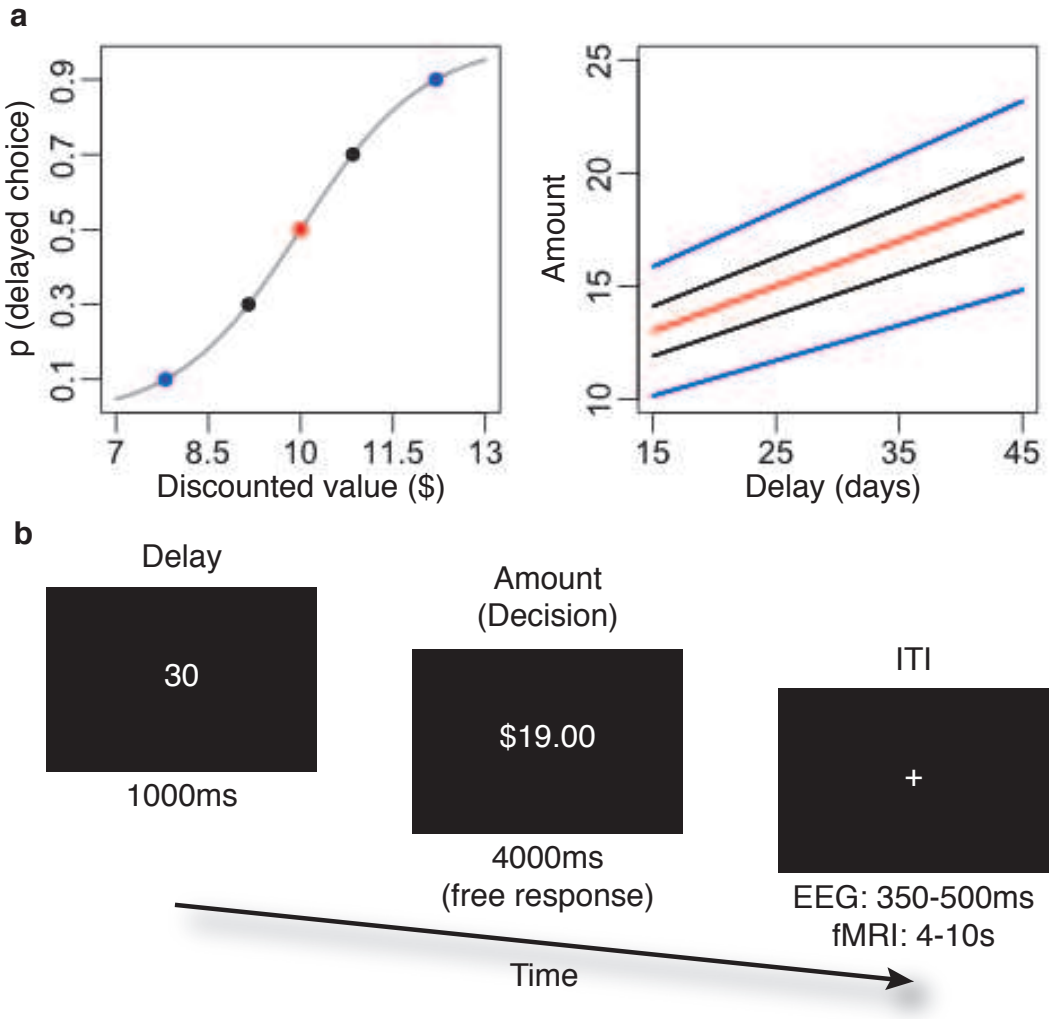

Figure 3: Experimental design: (a) Delayed reward offers corresponded with one of five different levels of discounted value. Each level of discounted value corresponded to one of five probabilities of choosing the delayed reward: 0.1, 0.3, 0.5, 0.7 or 0.9. (b) Every delay could be combined with any of five different amounts to yield a different discounted value and probability of choosing the delayed reward. (c) Delay and amount information was presented sequentially. Delays were presented first for 1000ms. Amounts were presented second, replacing the presentation of the delay and remaining on the screen for a maximum of 4000ms. After every trial, a fixation cross was presented on the center of the screen for a randomly chosen inter-trial-interval in the order of hundreds of milliseconds during the EEG experiment and several seconds during the fMRI experiment.

average reference, epoched trials from -1500 to +6500ms around the onset of delay presentation, baseline corrected trials using the average from 0 to 5000ms, and band-pass filtered the data at 0.5-200Hz. Trials were visually inspected and rejected if excessive artifacts were present. Normally occurring artifacts were rejected using an independent component analysis algorithm from the EEGLab toolbox (Delorme and Makeig, 2004). Epochs were then transformed to current source density (CSD) using the CSD toolbox (Kayser and Tenke, 2006).

fMRI data was collected using a GE Discovery MR750 Scanner. fMRI analyses were conducted on gradient echo T2*-weighted echoplanar functional images with BOLD-sensitive contrast (42 transverse slices; TR, 2000 ms; TE, 30 ms; 2.9mm isotropic voxels). Slices had no gap between them and were acquired in interleaved order. The slice plane was manually aligned to the anterior-posterior commissure line. The total number of volumes collected per subject varied depending on random inter-trial intervals. The first 8 seconds (4 volumes) of data contained no stimuli and were discarded to allow for T1 equilibration. In addition to functional data, we collected whole-brain, high-resolution T1-weighted anatomical structural scans (0.9mm isotropic voxels). Image analyses were performed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/). During pre-processing, we first performed slice-timing correction, and realigned functional volumes to the first volume. Next, we co-registered the anatomical volume to the realigned functional scans, and performed a segmentation of grey and white matter on the anatomical scan. Segmented images were then used to estimate non-linear Montreal Neurological Institute (MNI) normalization parameters for each subject's brain. Normalization parameters estimated from segmented images were used to normalize functional images into MNI space. Finally, normalized functional images were smoothed using a Gaussian kernel of 8 mm full-width at half-maximum.

2.2. Description of the Data

To provide some rationale about the specific assumptions we made in developing our models, we will briefly discuss the primary details of the behavioral and neural data, as well as a "folding" process we used to make a seamless connection from brain to behavior.

2.2.1. Behavioral Data

The behavioral data consisted of choice response time measures on each trial, and the aggregate of these measures is illustrated in Figure 4. We illus-

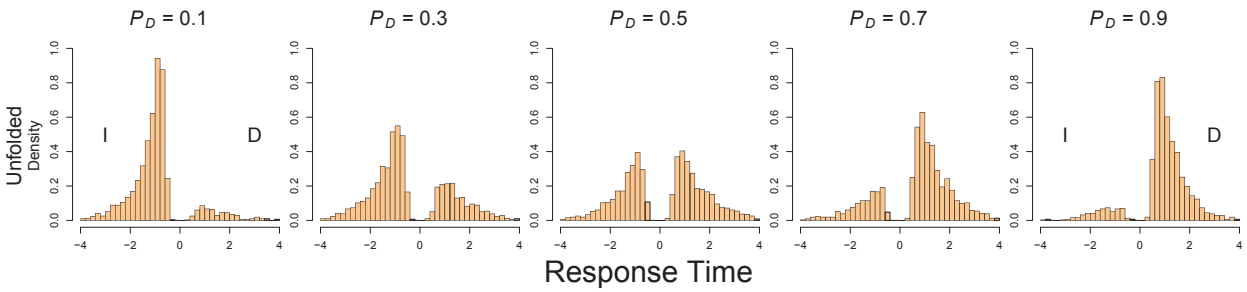

Figure 4: Summary of behavioral data from the experiment. Moving from left to right, each histogram represents the choice response time distribution corresponding to the five levels of probability of delayed reward P_D arranged in increasing order. Response times corresponding to the immediate reward option (I) are shown on the negative x-axis, whereas response times corresponding to the delayed reward option (D) are show on the positive x-axis.

trate the choice response times by plotting the response time distributions for the immediate option (i.e., \$10 now; I) on the negative x-axis, and the response time distributions for the delayed option (i.e., \$$X$ in Y days; D) on the positive x-axis. Plotting the distributions in this way allows us to assess the relative choice probabilities by comparing the heights of the two distributions. Figure 4 shows these response time distributions for each of the P_D values: $P_D = 0.1$, 0.3, 0.5, 0.7, and 0.9, left to right, respectively. The figure shows that at low values of P_D, the probability of choosing the immediate option is high, and as P_D increases, so does the probability of choosing the delayed option. This indicates that our hyperbolic discounting model captured the choice process well enough to show a behavioral effect within the experimental manipulation.

Given the linear relationship to P_D and symmetric properties of the neurophysiological measures reported below, it was convenient to fold our data symmetrically around $P_D = 0.5$. The folding process created three new experimental variables: $V1$ consisting of data from $P_D = \{0.1, 0.9\}$, $V2$ consisting of data from $P_D = \{0.3, 0.7\}$, and $V3$ which was equivalent to data from the previous $P_D = 0.5$ condition. To elaborate, we recoded immediate-option decisions in the $P_D = 0.1$ condition and delayed-option decisions in the $P_D = 0.9$ condition to a variable called "subjectively high valued" (SHV), and delayed-option decisions in the $P_D = 0.1$ condition and immediate-option decisions in the $P_D = 0.9$ condition to a variable called "subjectively low val-

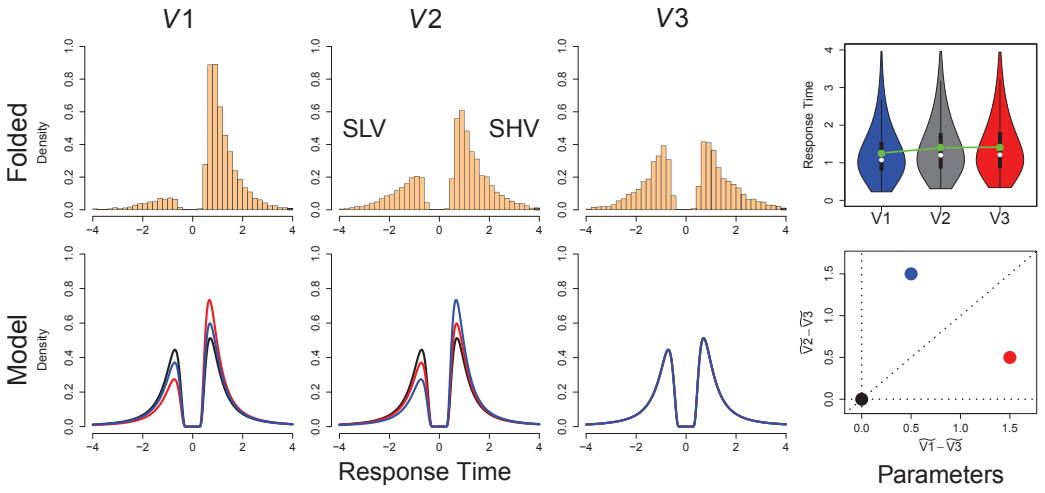

Figure 5: Summary of behavioral data from the experiment. The top row shows the choice response time distributions for the three new conditions $V1$ (first column), $V2$ (second column), and $V3$ (third column) after the folding process (see main text). The response times corresponding to the subjectively low-valued option (SLV) are shown on the negative x-axis, whereas the response times corresponding to the subjectively high-valued option (SHV) are shown on the positive x-axis. The fourth column shows a violin plot of the response times collapsed across the response choice for each condition, with the green line/dots representing the mean, and the white dots representing the median. The bottom row shows predictions from the LBA model for each condition (corresponding to the top row) according to three different sets of shift parameters in the drift rate, illustrated in the fourth column.

ued" (SLV). The same mapping was applied to the $P_D = 0.3$ and $P_D = 0.7$ conditions, but no mapping was necessary for the $P_D = 0.5$ condition. The top row of Figure 5 shows the response time distributions for the variables $V1$ (first column), $V2$ (second column), and $V3$ (third column), where SLV choices are shown on the left and SHV choices are shown on the right. The figure shows that as the levels of V increase, preference of the SHV option decreases until indifference in condition $V3$. In addition to a relationship between response probability and condition, the top-right panel of Figure 5 shows a violin plot of the response time distributions for each condition, collapsed across response probability. The white dots on each violin plot represent the median response time, whereas the green dot represents the mean response time. The figure shows that the response times increase from $V1$ to $V2$, and from $V2$ to $V3$, but to a lesser degree.

In previous work (Rodriguez et al., 2014), we modeled the (unfolded) intertemporal choice data within a sequential sampling framework, using the Linear Ballistic Accumulator (LBA; Brown and Heathcote, 2008) model to explain this choice behavior. In Rodriguez et al. (2014), we fit a variety of different models positing different theoretical accounts of how the choice process should change across values of P_D in order to best capture the behavior. We compared these models on the basis of conventional model fit statistics, taking into account model complexity and fit, and found that the best model allowed both a drift rate and nondecision time parameter to vary as a function of P_D values. In subsequent analyses, we found justification for these choices by comparing the drift rate parameters across the levels of P_D, and concluded that the drift rate was systematically related (i.e., sigmoidally) to P_D across subjects. Furthermore, this relationship was symmetric about $P_D = 0.5$. For the nondecision time parameter, we did not find any substantive relationship with P_D, but did find relationships with the median response time in each P_D condition. Our conclusion was that the nondecision time parameter was not systematically related to P_D, but that the nondecision time parameter afforded us a level of flexibility that enhanced the model's fit to the data.

Given that the best-fitting LBA model from our previous work relegated the systematic effects of P_D to the drift rate, when folding the data, we maintained a symmetric relationship between drift rate and value condition. We also allowed the nondecision time to vary across value conditions. To parallel the neural model we describe below, we reparameterized the model such that a "base" drift rate accounted for the choice response times in Condition $V3$, a parameter $\eta^{(1)}$ additively shifted the model's drift rate between Condition

$V3$ to $V2$, and another parameter $\eta^{(2)}$ additively shifted the model's drift rate between Condition $V3$ to $V1$. Although Appendix A contains all of the mathematical details of our implementation, the basic equations of our drift rate parameterization for the subjectively high valued option are

$$\begin{aligned} \mu_{V3} &= \text{logit}^{-1}(\nu), \\ \mu_{V2} &= \text{logit}^{-1}\left(\nu + \eta^{(1)}\right), \text{ and} \\ \mu_{V1} &= \text{logit}^{-1}\left(\nu + \eta^{(2)}\right), \end{aligned}$$

where μ_k is the drift rate for the subjectively high valued option in condition k, and ν represents the "base" drift rate. This drift rate reparameterization is mathematically equivalent to our previous account (Rodriguez et al., 2014), although this new formulation of the nondecision time component in our model has two fewer parameters.

To illustrate how the LBA model accounts for the folded data, we simulated the model – described in Appendix A – for three different settings of the shift parameters for each of the three value conditions, shown in the bottom row of Figure 5. The settings of the shift parameters are represented in the bottom-right panel of Figure 5, where the x-axis corresponds to $\eta^{(2)}$ (i.e., the shift from $V3$ to $V1$), and the y-axis corresponds to the $\eta^{(1)}$ (i.e., the shift from $V3$ to $V2$). The first parameter setting – illustrated by the black dot in the bottom right panel of Figure 5 – was $(0, 0)$. This setting produces no change in drift rate across any of the value conditions. The first, second, and third columns show the predicted choice response time distributions from the LBA model (black lines) for the value conditions $V1$, $V2$, and $V3$, respectively. Here, because no change in drift rate occurred, the model makes equivalent predictions across value conditions. Other parameter settings are more interesting. For example, the red dot in the bottom-right panel of Figure 5 shows the model predictions under the setting $(\eta^{(2)}, \eta^{(1)}) = (1.5, 0.5)$. Moving from $V3$ (third column) to $V2$ (second column), we see a marked increase in the probability of endorsing the SHV option. This change is produced by the 0.5 unit increase (i.e., on the logit scale) in the drift rate for the SHV option. Moving from $V3$ (third column) to $V1$ (first column), we see an even larger increase in the probability of endorsing the SHV option. Again, this increase follows directly from the value of $\eta^{(2)} = 1.5$. The third parameter setting produces response time distributions that are similar to the second parameter setting, but the magnitude of the effect is in the opposite order with respect to value conditions. These predictions are not in

line with what the pattern present in the data, shown in the top row of Figure 5, whereas the simulation produced from the second parameter setting does closely match the empirical data. In interpreting the parameter values in subsequent sections, we should keep in mind that in general, parameters falling in this lower triangular area are more reflective of the pattern in the observed data than are parameters in the upper triangular area.

2.2.2. Neural Data

Our neural data consisted of both EEG and fMRI recordings. In a previous analysis (Rodriguez et al., 2015a), we first identified the elements of the EEG data that were significantly related to P_D. Figure 6a shows a topographic plot of the GLM analysis relating EEG activity by electrode to the levels of P_D. The strongest relationship was observed at an area of the scalp that roughly corresponds to the dorsal medial frontal cortex (dmFC), and these electrodes are illustrated in Figure 6a with white circles (i.e., electrodes E5, E6, E11, and E12). Following this result, we investigated the temporal properties of this signature by plotting the average EEG signal for each of the levels of V, shown in Figure 6b. This analysis revealed a key time window where the average EEG signal was differentiated across the levels of V ranging from 300-850 milliseconds, illustrated by dashed vertical lines. Figure 6c, 6d, and 6e show the average EEG signal for the four electrodes in Figure 6a in the 300-850 milliseconds time window for $V1$, $V2$, and $V3$, respectively, color coded according to the key on the right hand side. The figures show that as V increases, the average EEG signal decreases. Having identified the key spatial and temporal aspects of the neural signal, we can then use these results to facilitate our joint behavioral-neural analyses below by reducing the dimensionality of the neural data. Specifically, the single-trial EEG data observed at the four electrode locations in Figure 6a and in the time window 300-850 milliseconds were used as input into the models below.

A similar procedure was used to identify the key neural signature in the fMRI data (Rodriguez et al., 2015b). The first analysis investigated which brain regions were linearly related to the levels of V through a GLM analysis. Figure 6f shows that the dmFC was significantly related to V across subjects, a finding that is consistent with our EEG analysis in the top row (Rodriguez et al., 2015a). Having identified this area, we estimated single-trial βs for the neural activity in the dmFC by fitting a general linear model to the blood oxygenation level dependent (BOLD) responses. Each GLM specified the onsets of the delay presentation and and the subjects' response in every trial,

plus the onset of the amount presentation for a single trial. Events in all three onset regressors were modeled as impulse delta functions and convolved with the canonical hemodynamic response function (HRF). In addition, the model included six regressors corresponding to the motion parameters estimated during data preprocessing and constants to account for the mean activity within each of the four sessions over which the data were collected. Figure 6g shows the average of the single-trial βs in $V1$ (blue), $V2$ (black), and $V3$ (red), and this average activity is also illustrated in the brain on the same axial slice given in Figure 6f for $V1$ (h), $V2$ (i), and $V3$ (j). Collectively, Figures g-j show that as the levels of V increase, activation of the dmFC increases, an effect that is opposite in direction to our EEG data. In our analyses below, we will use the single-trial β estimates in the dmFC as input to our joint behavioral-neural model.

3. A Bivariate Joint Model of Value Accumulation

Our ultimate goal is to demonstrate that a trivariate joint model can be useful in modeling EEG and fMRI data from a descriptive and predictive perspective. However, to show that this is a useful extension, we must also demonstrate that the trivariate model provides something beyond bivariate joint models (Turner et al., 2013a). Furthermore, it might be useful to first show some of the advantages a bivariate joint model provides through two pedagogical examples. Hence, we begin our analyses by examining the extent to which each of the neural measures are related to the latent parameters of the behavioral model. To accomplish this, we constructed a joint model that simultaneously captures the effects present in the choice response time data and the neural recordings. Because we are only examining two sets of variables (i.e., behavior and EEG measures), we will refer to this model as a bivariate joint model. Figure 1 illustrates the Trivariate model that subsumes the bivariate model presented in this section; that is, in this section we are only concerned with relating the behavioral data (orange plate) to either the EEG data (green plate; first analysis) or the fMRI data (blue plate; second analysis). We will now conceptually summarize each of the three components of the bivariate model, and refer the interested reader to Appendices A and B for the formal details.

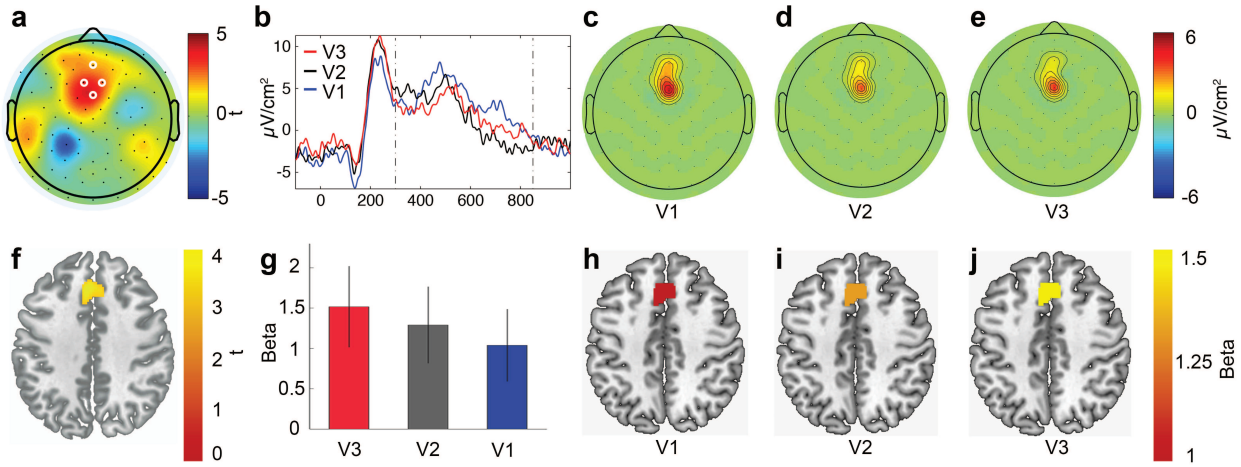

Figure 6: Summary of neural data from the experiment. (a) Topographic plot of the test value statistics relating EEG activity to the levels of P_D in a GLM analysis. The electrodes with the strongest relationship are represented with white circles. (b) Mean stimulus-locked event-related potential (ERP) for the three value conditions: V1 (blue line), V2 (black line), and V3 (red line). The time window used in the analysis was from 300 to 850 milliseconds, which is indicated by the dashed vertical lines. (c, d, e) Topographic plots of the average EEG signal in the four electrodes shown in (a) during the time window in (b), for each of the three value conditions: V1 (c), V2 (d), and V3 (e). These trial-by-trial EEG values were used in the modeling applications below. (f) Axial view of the brain region (i.e., the dmFC) found to significantly correlate with value condition. (g) Average of the single-trial beta estimates for each value condition, based on a GLM analysis of dmFC activity during the decision period. (h, i, j) Similar information in Panel g, but illustrated in the dmFC, and separated across the three value conditions: V1 (h), V2 (i), and V3 (j). These trial-by-trial β values were used in the modeling applications below.

3.1. Details of the Model

The joint modeling approach is particularly well-suited for situations in which the measures of interest can be described independently of one another, and the research goal is to infer and constrain the model parameters describing each of these measures (see Turner et al., 2015a, for a survey of different statistical strategies for linking neural and behavioral measures). These independent "submodels" are then linked together, comprising an entirely new model with different – albeit similar – properties of its constituent parts. Furthermore, the linking process inherently changes the submodels, due to the added constraint provided by the additional measures. That is, data from one submodel affects the parameter estimates of the other submodel, and vice versa. Perhaps more interesting is that the full model is fit to the full data set in one setting, so elaborate retraining techniques are not necessary in generating flexible predictions (e.g., for missing data). The bivariate joint model can be described in three parts: (1) the neural submodel which succinctly describes the important statistical properties of the neural data, (2) the behavioral submodel consisting of latent model parameters assumed to produce an observer's set of choice response time data, and (3) a linking structure that binds the two submodels together. We will discuss each of these pieces of the model in turn.

3.2. The Neural Submodel

The first piece of the bivariate joint model is the neural submodel describing the neural data. The choice of a neural submodel is an important one, but in our experience, the most important factor in this choice is reducing the neural data to a (smaller) set of neural model parameters. The neural submodel can be generative, expressing how the neural data arise from a theoretically-oriented process (e.g., Manning et al., 2014; Friston et al., 2003; Gershman et al., 2011), or statistical, describing the statistical properties of the neural data without theoretical considerations (e.g., Norman et al., 2006). One effective strategy within the joint modeling framework for functional data is to average the trial-to-trial neural signal across an important time window in the decision process (e.g., Turner et al., 2015b). This procedure allows one to easily justify choices of the linking function via the central limit theorem. Specifically, averaging the neural data produces a random variable that is Gaussian in distribution, making multivariate Gaussian linking functions both justified and convenient (see Appendix B for details).

Our data consist of two neural measures – EEG and fMRI – and so to examine the modality-specific relationships present in each neural measure, we will use the bivariate joint model in two separate analyses: one that includes behavior and only EEG data, and one that includes behavior and only fMRI data. As discussed above, Figure 6 illustrates the EEG data from the experiment. Panel a shows the group average across value conditions in the critical time period. To reduce the dimensionality of our data, we focused on the four electrodes centered over dorsal medial frontal cortex (dmFC), which are represented in Panel a as white circles. Panel b shows the average stimulus-locked EEG signal for each value condition across time, where the critical time period is marked by the dashed vertical lines. Panels c, d, and e show the average topography for each electrode in the same window, but separated according to value condition (i.e., $V1$, $V2$, and $V3$, respectively). We used the single-trial measures of EEG activity as input to the model on a particular trial; hence, for a particular subject on a particular trial, our neural measures were a vector consisting of four values: the time-averaged EEG signal amplitude between 300 and 850 milliseconds for each of the four electrodes represented in Panel a of Figure 6.

We used a similar strategy for the fMRI data. Panel f in Figure 6 shows the t-statistic obtained for linear trend across value conditions (a statistic that we had previously obtained in Rodriguez et al., 2015b). This GLM analysis identified the dmFC as a key brain area in the deliberation of intertemporal choices, which motivated its use in the analyses in this article. Having identified the dmFC in this group analysis, we could then establish single-trial β estimates of dmFC activation. Next, we can group these single-trial values according to the value condition. Panel g shows these average beta estimates across value conditions in a bar graph, whereas Panels h, i, and j show these same average values in the identified dmFC area in Panel f, to be consistent with the top row of Figure 6. We used the single-trial β estimates of dmFC activity as input to the bivariate joint model. Compared to the EEG data, our fMRI data are more sparse (i.e., consisting of a single value rather than four values).

As described in Appendix B, the bivariate joint model uses a parameter to capture the mean neural (i.e., either EEG or fMRI) signal in the $V3$ condition, where the choice options have approximately the same subjective value, as well as additive parameters that capture the change in the mean neural signal from value condition $V3$ to $V2$, and from $V3$ to $V1$. At the individual subject level, these parameters are called $\delta_j^{(1)}$, and $\delta_j^{(2)}$, respec-

tively, for Subject j. However, at the group level, they are referred to as $\delta_{\mu}^{(1)}$, and $\delta_{\mu}^{(2)}$. These parameters are important because they are linked to parameters in the behavioral model that capture an analogous additive effect in the behavioral data. We will use these parameters interchangeably in the analyses that follow to describe differences in the EEG signal or β values across conditions.

3.2.1. The Behavioral Submodel

To characterize the behavioral data, we used the Linear Ballistic Accumulator (LBA; Brown and Heathcote, 2008) model. While the technical details of the model are described in Appendix A, the concept of the LBA model is similar to all other sequential sampling models of choice response time (e.g., Ratcliff, 1978; Usher and McClelland, 2001; Shadlen and Newsome, 2001; Busemeyer and Townsend, 1993). The model assumes that, upon presentation of a stimulus, a competition ensues among the choice alternatives to gather evidence sequentially, leading to a threshold amount of evidence. In the LBA model, this accumulation process occurs ballistically and independently for each choice alternative. Once one of the alternatives gathers enough evidence to reach a threshold, a choice is made by the observer to correspond to the winning accumulator. Rodriguez et al. (2014) showed that the LBA model provided an explanation of decision-making behavior in intertemporal choice that facilitated an understanding of how reward and delay information simultaneously mapped to choice probabilities and response times.

In our model, the parameters of greatest interest are the parameters that correspond to the rate of evidence accumulation (i.e., the "drift rate" parameters) across value conditions. As described in Appendix A (see also Figure 5), our model uses a parameter to capture the base drift rate in the $V3$ condition, where the options are equally preferable, as well as parameters that capture the additive shift in the drift rate from $V3$ to $V2$, and from $V3$ to $V1$. At the individual subject level, these parameters are called $\eta_j^{(1)}$, and $\eta_j^{(2)}$, respectively, for Subject j. However, at the group level, they are referred to as $\eta_{\mu}^{(1)}$, and $\eta_{\mu}^{(2)}$.

3.2.2. The Linking Structure

The final component of the model simply links the parameters of the neural and behavioral submodels together. To accomplish this, we exploit the Gaussian properties of our neural submodel, and directly connect the param-

eters $\delta_{\mu}^{(1)}$, $\delta_{\mu}^{(2)}$,$\eta_{\mu}^{(1)}$, and $\eta_{\mu}^{(2)}$ by way of a multivariate Gaussian distribution (i.e., Equation B.1). The consequence of the linking function is a formal association between the shift parameters in the neural and behavioral setting. Specifically, if the model learns that larger changes in $\delta^{(2)}$ are accompanied by larger changes in the drift rates $\eta^{(2)}$, then for a new subject, the model will predict changes in the behavioral data that are proportional to the changes observed in the neural data between conditions $V3$ and $V1$. The multivariate Gaussian distribution is convenient because it provides clear interpretations for the parameters of interest. Namely, the variance-covariance matrix can be transformed into a correlation matrix that carries with it a quantification of the magnitude and direction of the relationship between pairs of model parameters. This is particularly useful in our analyses because it directly informs us about the type of associations that exist between the parameters of the neural and the behavioral submodels.

3.3. Application To EEG Data

The first analysis only considers the effects of adding EEG data to the analysis of behavioral data. To examine this, we focus on the first 23 subjects in our data (i.e., the first 23 rows shown in Figure 2). These subjects will be used again in the Trivariate Joint Model analysis below. We did not use the last eight subjects (i.e., Subjects 47-54) in our experiment, even though they did provide EEG data in one condition.

3.3.1. Results

We present our results in two parts. First, we discuss our results from a generative modeling perspective, by evaluating the relative merits of our modeling approach through an examination of the posterior distributions. To do this, we compared the parameter estimates of the joint model, which takes into account both the neural and behavioral data, to the parameter estimates of a behavioral-data-only model, which is effectively the behavioral submodel of the bivariate joint model discussed above. Such a comparison allows us to evaluate the magnitude of the constraint that the neural data provide on the behavioral model (cf. Turner, 2015; Cassey et al., 2015). Second, we discuss our results from a predictive modeling perspective, where we use out-of-sample predictions to assess the generalizability of the bivariate joint model relative to the behavioral-data-only model. Such an analysis allows us to assess the importance of the neural data in making predictions about

behavioral data, and avoid problems of interpretation that are associated with overfitting (Wilson et al., 2015).

Generative Model Analysis. A central theme in the joint modeling framework is that the neural data should guide the behavioral model in the inference process. Ideally, the joint model would identify a statistical pattern in the neural data that resembles the mechanism(s) assumed by the behavioral model. Due to the particular assumptions of our linking function, if a statistically similar (i.e., correlated) mechanism to the behavioral model is found in the neural data, the amount of information inferred about the latent variable from the full data set will increase, which can affect both the mean and variance of the posterior distribution (Turner, 2015). Both of these posterior properties are important because they can affect the accuracy of the predictions about new data. One way to examine the amount of information gained from the neural data is to compare the parameter estimates obtained from fitting a behavioral-data-only model to the behavioral data to the estimates obtained from fitting the bivariate joint model to the behavioral and neural data. Differences in the estimates of the behavioral model parameters are attributable to information from the neural data that has propagated upward through the hyperparameters, and downward to the behavioral submodel parameters.

As discussed above, the key parameters in our analysis are the parameters that capture the additive effects of evidence accumulation across value conditions. Because our model assumes a structural relationship from subject-to-subject, the important parameters to examine are $\eta_j^{(1)}$ and $\eta_j^{(2)}$. To examine the effects of such a linking structure, one strategy is to selectively withhold and reveal some aspects of the neural data from the model. To implement this, we arbitrarily removed the behavioral data for Subject 1 before either model was fit, but still modeled Subject 1's parameters in the fitting routine. We did not withhold the neural data for Subject 1 because these data will have no influence on the parameters of the behavioral-data-only model (but they will influence the parameters of the bivariate joint model). When fitting the behavioral-data-only model to these data, the resulting parameter estimates are akin to a prior predictive distribution. However, when fitting the bivariate joint model to data, the resulting parameter estimates are still the prior predictive distribution, but importantly, they are now a conditional prior predictive distribution because the parameters of the neural submodel are estimable. The resulting predictive distributions can then be used to make predictions about data across all value conditions in the experiment.

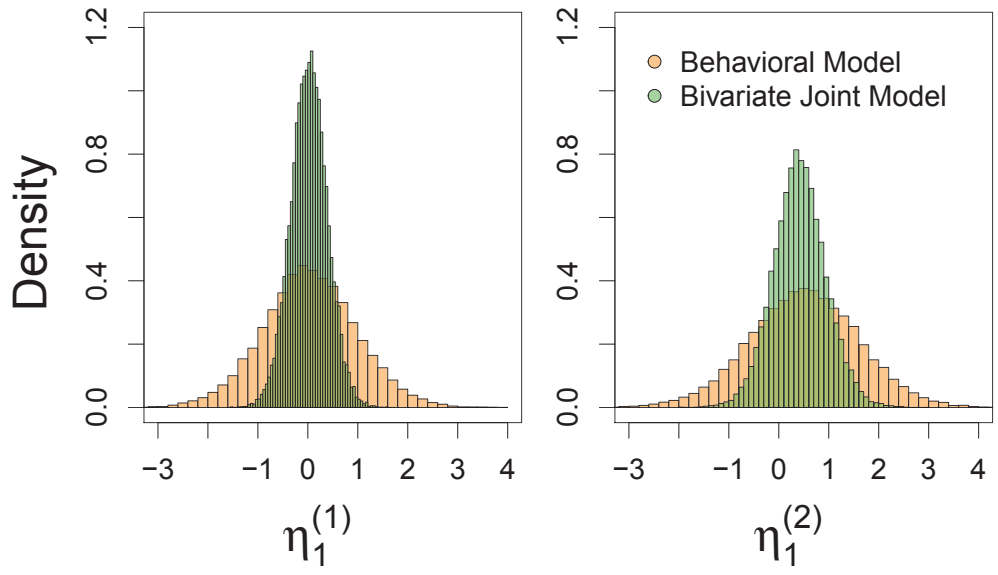

Figure 7: Comparison of the value effect parameters $\eta_1^{(1)}$ (left panel), and $\eta_1^{(2)}$ (right panel) obtained from the behavioral-data-only model (orange) and the bivariate joint model (green) for the withheld subject in the EEG experiment.

We first fit the behavioral-data-only (i.e., a hierarchical LBA) model to the behavioral data, and then fit the bivariate joint model to the full data. The algorithmic details of our fitting procedure are outlined in Appendix B. Figure 7 shows the estimated posterior distributions for the parameters $\eta_1^{(1)}$ (left), and $\eta_1^{(2)}$ (right) for Subject 1. In both panels, the orange histograms correspond to the estimated parameters for the behavioral-data-only model, whereas the green histograms correspond to the bivariate joint model. Recall that $\eta_1^{(1)}$ and $\eta_1^{(2)}$ represent the increase in the drift rate from the $V1$ (i.e., the most difficult) condition to $V2$ and $V3$, respectively. Increases in the drift rate from one condition to another reflect increases in the subject's propensity to choose the response that is subjectively larger in value. Hence, we should expect to see increasingly positive values for $\eta_1^{(1)}$ and $\eta_1^{(2)}$ such that $\eta_1^{(2)} > \eta_1^{(1)}$ (Rodriguez et al., 2014). However, we should keep in mind that this intuition is derived from group-level analyses, and does not necessarily describe the data for each individual subject. Furthermore, the posterior distributions in Figure 7 have not been estimated from behavioral data. Instead, they represent the model's ability to generalize the information contained in the remaining 22 subjects in the data subset.[2]

Figure 7 shows that for both models, the prediction for $\eta_1^{(1)}$ is centered on zero, whereas for $\eta_1^{(2)}$, the posteriors are shifted slightly upward. This indicates that both models were capable of using reasonably sparse data (i.e., just 22 subjects) to generalize predictions for a subject whose data were not observed in a way that is reflective of our previous behavioral analyses on a full data set (consisting of 46 subjects; Rodriguez et al., 2014). However, the important question was whether the bivariate joint model was more sensitive to the effects of value condition due to its ability to exploit the information in the neural signal. To answer this, we only need to compare the orange and green posterior predictions in both panels. Visually, it is clear that the green histograms have less variance, and the locations of the green and orange histograms are approximately equal. For $\eta_1^{(1)}$, the means are 0.393 and 0.437 and the standard deviations are 1.005 and 0.500 for the behavioral-data-only and the bivariate joint model, respectively. For $\eta_1^{(2)}$, the means are 1.068 and 0.946 and the standard deviations are 1.147 and 0.628 for the behavioral-data-only and the bivariate joint model, respectively. We can

[2]For the bivariate joint model, if you count Subject 1, who only had neural data, then the posteriors reflect a generalization from 23 subjects.

also evaluate the models' predictions for the probability that $\eta_1^{(1)}$ and $\eta_1^{(2)}$ are greater than zero. The behavioral-data-only model predicts that $p(\eta_1^{(1)} > 0) = 0.657$ and $p(\eta_1^{(2)} > 0) = 0.828$, whereas $p(\eta_1^{(1)} > 0) = 0.819$ and $p(\eta_1^{(2)} > 0) = 0.937$ for the bivariate joint model. This suggests that the relative generalizability to patterns across conditions is greater in the bivariate joint model, although it is worth emphasizing that this generalization may not result in better predictions for behavioral data. To address that question, we will use predictive modeling techniques in the next section.

Another generative modeling analysis would be to inspect the correlation parameters between the neural and behavioral submodels. In the bivariate models, two correlation parameters exist that describe the relationship between the additive shift parameters in the neural and behavioral submodels. Specifically, the parameter ρ_1 describes, across subjects, the degree to which $\eta_j^{(1)}$ is related to $\delta_j^{(1)}$ for all j, and the parameter ρ_2 describes, across subjects, the degree to which $\eta_j^{(2)}$ is related to $\delta_j^{(2)}$. Hence, the parameters ρ_1 and ρ_2 describe how much the effects present in the neural and behavioral data from conditions $V2$ to $V3$ and from $V1$ to $V3$, respectively. Figure 8 shows the estimated posterior distributions for the hyper correlation parameters ρ_1 (left panel), and ρ_2 (right panel). Figure 8 shows that ρ_1 is centered near zero, whereas ρ_2 is centered on positive values. Hence, the posterior estimate of ρ_2 suggests that the additive effects that give rise to higher drift rates in condition $V1$ are correlated with the additive effects of the EEG signal. Hence, across subjects, increases in the EEG signal in condition $V1$ relative to condition $V3$ are predictive of higher drift rates in condition $V3$ (see also Figure 7).

Predictive Modeling Analysis. We also evaluated the generalizability of our results by performing a leave-one-out cross-validation analysis for both models. Although we could have performed a similar analysis on the EEG data, we chose to focus our attention on predicting behavioral data because we felt it was the more likely application of our modeling framework (but see the Trivariate modeling section for predictions of neural data). To do this, we first hid the behavioral data for a single subject (but left the EEG data intact), then fit both the bivariate joint model and the behavioral-data-only model to the remaining data. The fitting process allowed us to estimate the posterior distribution of the model parameters, from which we generated predictions about the behavioral data – a distribution known as the posterior predictive distribution (PPD) – for each model. These predictions

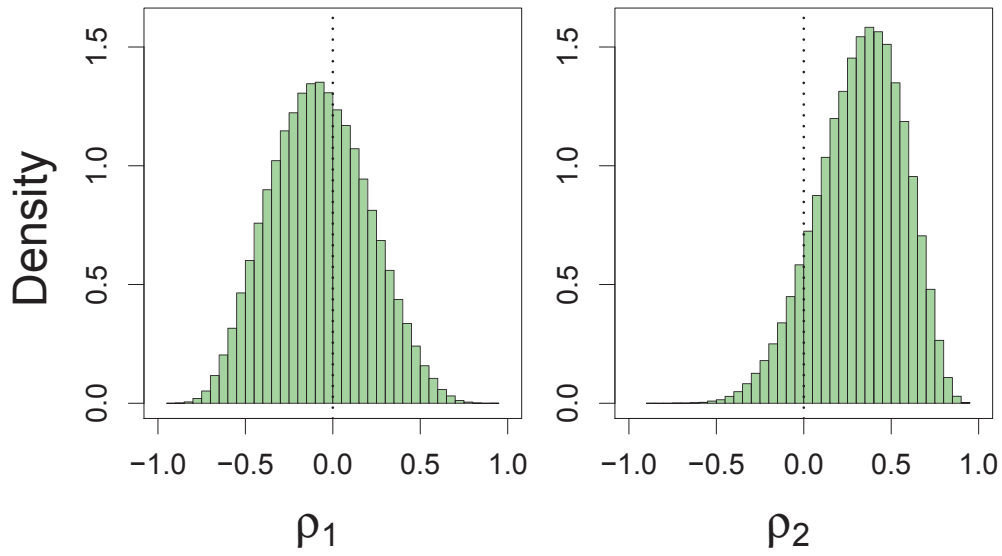

Figure 8: Estimated posterior distributions for the hyper correlation parameters relating the additive shift parameters in the neural and behavioral submodels across the $V3$ and $V2$ conditions (left panel), and the $V3$ and $V1$ conditions (right panel) for the EEG data. In each panel, a reference line is plotted to indicate zero correlation.

were made "out of sample", meaning that the data that was being predicted was not revealed to the model. Importantly, when generating predictions for behavioral data, the model was told which condition to make a prediction for, but was not told anything about the behavioral data themselves. The final step was to evaluate the likelihood of the (hidden) data under the PPD, which gave us a distribution of likelihood values. We took the median of these distributions for each model and plotted them against one another in Figure 9. Figure 9 compares the (log) likelihood values for each subject in our data. A better model fit is acquired by having a higher (log) likelihood value. The green region in Figure 9 represents areas where the bivariate joint model outperforms the behavioral-only model, whereas the orange region in Figure 9 represents areas where the behavioral-only model outperforms the bivariate joint model. Figure 9 shows that for 19 of the 23 subjects in our data, the bivariate joint model outperforms the behavioral-data-only model.

3.3.2. Summary

In this first analysis, we found that the bivariate joint model was able to generalize predictions to drift rate parameters better than a behavioral-data-only model. We found that these generalizations had similar expected values, but had considerably less variance and were better aligned with the patterns we already expected to see from previous experience with subsets of these data (Rodriguez et al., 2014). We then assessed whether or not this additional constraint was actually a good thing by cycling the models through a standard leave-one-out cross-validation test. We found that for 19 out of the 23 subjects in the EEG data subset, the bivariate joint model outperformed the behavioral-data-only model. A large part of this predictive power comes from the enhanced reliability of the model parameters. Having less variance associated with each parameter facilitates performance on cross-validation tests, as we have shown on predictions for single-trial model parameters (Turner et al., 2015b). Taken together, these results suggest that the neural data do indeed provide more insight to decision making in a way that facilitates prediction of behavioral data. Despite these promising findings, it remains an open question of whether this generalization ability works for other measures, such as fMRI.

3.4. Application To fMRI Data

Our second analysis only considers the effects of adding fMRI data to the analysis of behavioral data. To examine this, we focus on the second 23

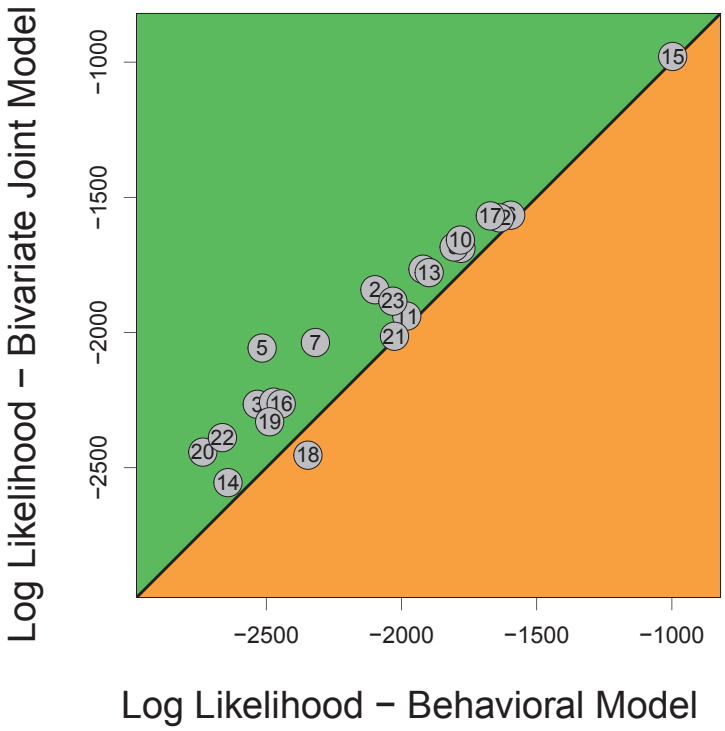

Figure 9: Predictive performance of the joint and behavioral models on the EEG dataset. Each node in the graph represents the models' performance for the subject indicated, where the performance for the joint model is represented on the y-axis and the performance for the behavioral-data-only model is represented on the x-axis. Nodes located in the green region indicate that the performance for the bivariate joint model was better than the behavioral-data-only model. In total, the bivariate joint model performed best for 19 of the 23 subjects.

subjects in our data (i.e., rows 24-46 in Figure 2). These subjects will also be used in the trivariate joint model analysis below. We did not use the data from the last eight subjects (i.e., Subjects 47-54) in our experiment, even though they did provide fMRI data in one condition.

3.4.1. Results

We present our results in an analogous manner as in the EEG application above. First, we present our generative modeling analysis by evaluating the models' ability to generalize the patterns in the data to new subjects where behavioral data are not observed. We then inspect the estimated posterior distributions for the correlation parameters relating the additive shifts in the neural and behavioral measures. Second, we perform another cross-validation test to assess the models' ability to predict behavioral data.

Generative Model Analysis. We assessed the benefits of the bivariate joint model relative to the behavioral-data-only model in an analogous way as in the EEG application above. Specifically, we withheld the behavioral data for Subject 24, but did not withhold Subject 24's neural data. We then estimated parameters from the full hierarchical model, and specifically Subject 24. To accomplish this, we first fit the behavioral-data-only (i.e., a hierarchical LBA) model to the behavioral data, and then fit the bivariate joint model to the full data. The algorithmic details of our model-fitting approach were equivalent to the EEG application above (also see Appendix B).

Figure 10 shows the resulting posterior distributions for Subject 24's main behavioral effect parameters: $\eta_{24}^{(1)}$ (left panel) and $\eta_{24}^{(1)}$ (right panel). In both panels, predictions from the bivariate joint model are shown as blue histograms, whereas predictions from the behavioral-data-only model are shown as orange histograms. Figure 10 shows that while the parameter $\eta_{24}^{(1)}$ is centered on zero, the parameter $\eta_{24}^{(2)}$ has a slightly positive effect, suggesting that as the subjective values of the two alternatives become more dissimilar, the drift rate for the subjectively higher valued alternative increases. Comparing across models, we see that visually both models produce posteriors with similar expected values, the posteriors are again substantially different. For $\eta_{24}^{(1)}$, the means are 0.340 and 0.329 and the standard deviations are 0.999 and 0.512 for the behavioral-data-only and the bivariate joint model, respectively. For $\eta_{24}^{(2)}$, the means are 1.132 and 1.277 and the standard deviations are 1.166 and 0.693 for the behavioral-data-only and the bivariate joint model, respectively. We can also quantitatively evaluate the models' generalizations for the

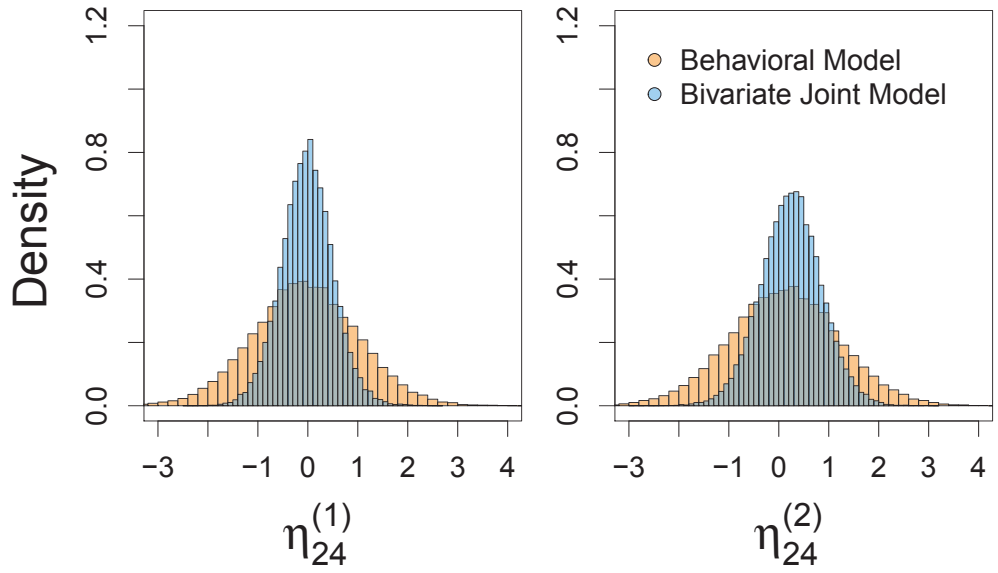

Figure 10: Comparison of the value effect parameters $\eta_{24}^{(1)}$ (left panel), and $\eta_{24}^{(2)}$ (right panel) obtained from the behavioral-data-only model (orange) and the bivariate joint model (blue) for the withheld subject in the fMRI experiment.

probability that $\eta_{24}^{(1)}$ and $\eta_{24}^{(2)}$ are greater than zero. The behavioral-data-only model predicts that $p(\eta_{24}^{(1)} > 0) = 0.638$ and $p(\eta_{24}^{(2)} > 0) = 0.836$, whereas $p(\eta_{24}^{(1)} > 0) = 0.747$ and $p(\eta_{24}^{(2)} > 0) = 0.967$ for the bivariate joint model. Similar to the EEG application above, these results suggest that the bivariate joint model is better able to generalize patterns in the data due to its ability to capture the influence of neural data on the behavioral submodel. In the predictive analysis section below, we will evaluate whether or not this added constraint is a good thing when predicting behavioral data.

As a final generative model evaluation, we can also examine the posterior distributions of the parameters controlling the degree of correlation between the additive effect parameters in the neural and behavioral submodels. Figure 11 shows the correlation parameters ρ_1 (left panel), and ρ_2 (right panel). In the model, ρ_1 controls the degree of correlation between the shift in drift rate parameter $\eta_j^{(1)}$ and the shift in the average single-trial beta values $\delta_j^{(1)}$, whereas ρ_2 controls this same shift in drift rate parameter $\eta_j^{(2)}$ and the shift in the average single-trial values $\delta_j^{(2)}$. Importantly, these correlations are

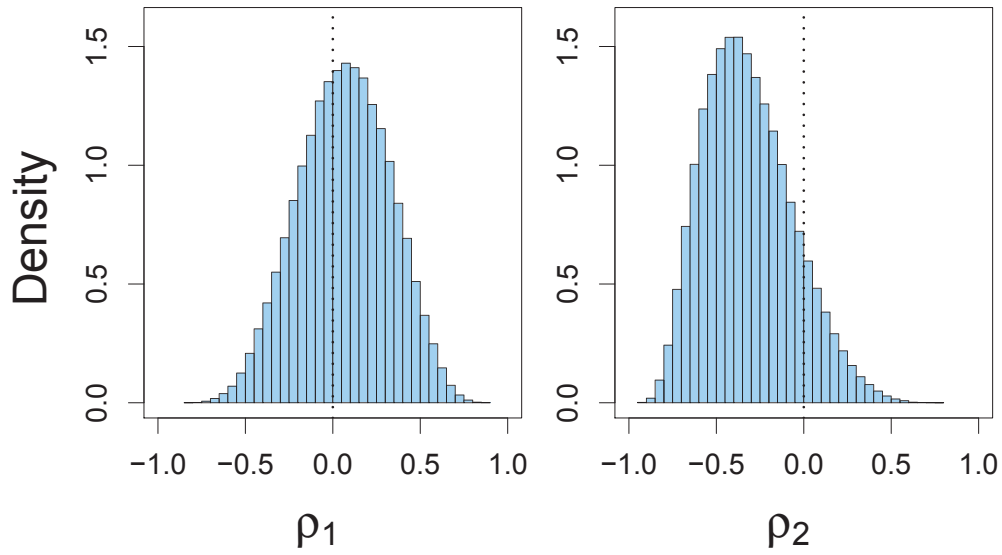

Figure 11: Estimated posterior distributions for the hyper correlation parameters relating the additive shift parameters in the neural and behavioral submodels across the $V3$ and $V2$ conditions (left panel), and the $V3$ and $V1$ conditions for the fMRI data. In each panel, a reference line is plotted to indicate zero correlation.

assessed across subjects, so for any new subject, we can condition on one shift parameter to generate predictions for the other (e.g., see Figure 10. Figure 11 shows that the estimated correlation parameter ρ_1 is centered at zero, reflecting that the information in the neural and behavioral submodel parameters is not strongly related across subjects. However, the estimated posterior distribution for the correlation parameter ρ_2 is centered on negative values off of zero, reflecting shared information between the submodel parameters.

Predictive Modeling Analysis. In the predictive modeling analysis, we used the same procedure described above in the EEG application to carry out a cross-validation test. We first removed the behavioral data for a single subject worth of data. Next, we fit both the bivariate joint model and the behavioral-data-only model to the remaining data set. Third, we generated predictions for the withheld behavioral data. Finally, we evaluated the log likelihood of the withheld data under each predictive distribution and compared the log

likelihood values across models. Figure 12 shows this comparison. The figure shows that the bivariate joint model obtained a higher log likelihood value for each of the 23 subjects' data.

Summary. Similar to the EEG application above, we find that the model predictions for the effects of the behavioral data are in line with our intuition derived from the group data. The shifted effect parameters across the neural and behavioral submodels were strongly correlated between Conditions V1 and V3. The effects present from Condition V3 to V2 were not as pronounced as those observed from Conditions V3 to V1. Together these results suggest that there is shared information between the neural and behavioral measures in this condition, leading to greater constraint on the behavioral submodel parameters (see Figure 11). We then evaluated the utility of this additional constraint via a leave-one-out cross-validation study. We again find that the predictions of the bivariate joint model were appreciably better than the behavioral-data-only model.

These analyses have established the unimodal utility of the neural data in both the EEG and fMRI measures. Both analyses produced better constraint of the behavioral submodel, and more importantly, this constraint lead to more accurate predictions about the behavioral data. However, the question remains as to whether or not the integration of behavior and *all* neural measures can improve on what could be obtained in separate bivariate joint model analyses. In the next section we evaluate this question by extending our bivariate joint model to all three measures simultaneously – a model that we refer to as a trivariate joint model.

4. A Trivariate Joint Model of Value Accumulation

To fit the data, we extended the bivariate joint model used in the previous sections by simply combining the neural submodels presented in the bivariate EEG and bivariate fMRI models above. This creates a model illustrated in Figure 1 that captures behavior, EEG, and fMRI data in one cohesive structure. The technical details of this extension are presented in Appendix C.

4.1. Results

Similar to the bivariate applications to EEG and fMRI data above, we performed both generative and predictive modeling analyses. In addition, in

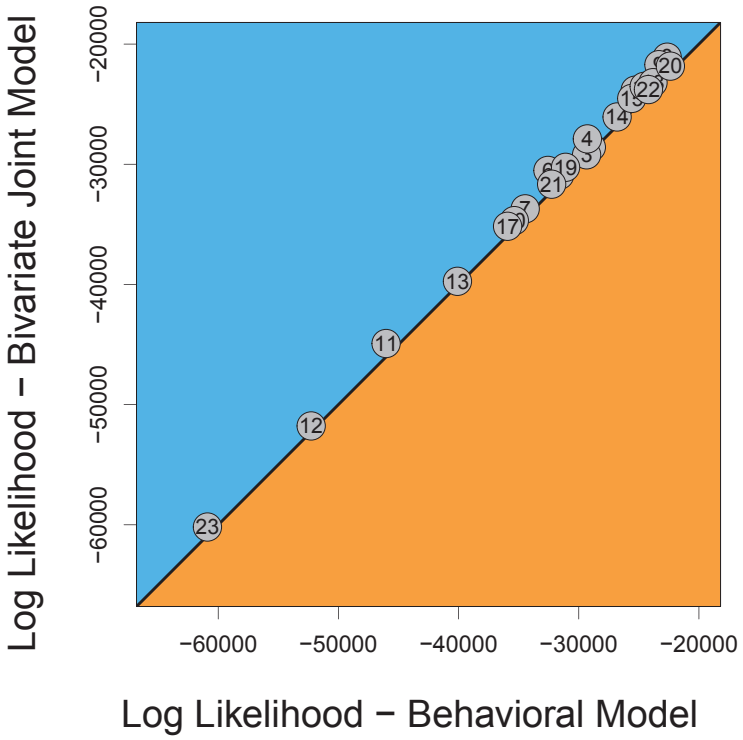

Figure 12: Predictive performance of the joint and behavioral models on the fMRI dataset. Each node in the graph represents the models' performance for the subject indicated, where the performance for the bivariate joint model is represented as the y-axis and the performance for the behavioral model is represented as the x-axis. Nodes located in the blue region indicate that the performance for the bivariate joint model was better than for the behavioral-data-only model. In total, the bivariate joint model performed best for all 23 subjects.

the predictive modeling section below, we show how the trivariate model can be used to generate predictions for neural data conditioned on behavior. We present these analyses in turn.

4.1.1. Generative Model Analysis

Our generative analysis was similar to the bivariate applications above. In our first analysis, we focused on the effects of the neural data on the shift parameters $\eta^{(1)}$ and $\eta^{(2)}$ in the behavioral submodel. Unlike our bivariate applications above, we now have two neural measures we can consider adding to the model, creating four different models to test. The first model only considers the behavioral data, the second model considers only the behavioral and EEG data, the third model considers only the behavioral and fMRI data, and the fourth model considers the behavioral, EEG, and fMRI data, as in Figure 1. For illustration purposes, we again withheld the behavioral data for one subject – Subject 53 – to compare and contrast each model's predictions.[3] In the predictive modeling analysis section below, we will use an extension of this technique in a cross-validation test. Although we withheld Subject 53's behavioral data, the EEG and fMRI data were left intact. Because some of these models are unable to use all components of the neural data, each model's ability to generalize to a new subject's behavioral data depends on the reliability of the usable neural data.

We fit each of the four models using the methods described in the appendices. Figure 13 shows the resulting posterior distributions for Subject 53's main behavioral effect parameters: $\eta_{53}^{(1)}$ (left panel) and $\eta_{53}^{(2)}$ (right panel). In both panels, predictions from the behavioral-data-only, bivariate EEG, bivariate fMRI, and trivariate models are shown as orange, green, blue and red densities, respectively. Figure 13 shows that both $\eta_{53}^{(1)}$ and $\eta_{53}^{(2)}$ are centered on

[3]Unlike the bivariate modeling applications above, the trivariate modeling analyses must deal with a sparsity issue in the number of subjects providing all three measures (i.e., behavioral, EEG, and fMRI). For this reason, Subject 53 was not chosen arbitrarily, but was instead selected on the basis of having the fewest number of trials. Specifically, after data preprocessing, Subject 53 provided 240 observations of joint behavior and EEG data, but only 98 observations of joint behavior and fMRI data, whereas the other 7 subjects provided (on average) 226.14 and 239.71 trials, respectively. This choice makes use of the most amount of data, but should have little impact on the results. For example, choosing another subject might weaken the correlation observed between EEG and fMRI, but only proportional to the difference between the number of data points for this hypothetical subject and the number of data points for Subject 53, which is on average 127.85 trials.

positive values, with $\eta_{53}^{(2)}$ being larger than $\eta_{53}^{(2)}$. These results are consistent with the previous bivariate joint model analyses above. Comparing across models, we see that while all models produce posteriors with similar expected values, the dispersion of those posteriors differs substantially. Namely, the behavioral-data-only model has the widest variance, the bivariate EEG and fMRI models have similar and slightly less variance, and the trivariate model has the smallest variance. We will evaluate whether this pattern in variances across models leads to better predictions of the behavioral data in the predictive modeling analysis section below.

We can also quantitatively evaluate the models' generalizations for the probability that $\eta_{53}^{(1)}$ and $\eta_{53}^{(2)}$ are greater than zero. The behavioral-data-only model predicts that $p(\eta_{53}^{(1)} > 0) = 0.666$ and $p(\eta_{53}^{(2)} > 0) = 0.874$. The bivariate EEG model predicts that $p(\eta_{53}^{(1)} > 0) = 0.766$ and $p(\eta_{53}^{(2)} > 0) = 0.950$. The bivariate fMRI model predicts that $p(\eta_{53}^{(1)} > 0) = 0.796$ and $p(\eta_{53}^{(2)} > 0) = 0.929$. Finally, the trivariate model predicts that $p(\eta_{53}^{(1)} > 0) = 0.839$ and $p(\eta_{53}^{(2)} > 0) = 0.970$. The basic pattern is consistent with the pattern of variance above: the models whose posteriors have less variance predict greater probabilities that the η parameter is greater than zero. Furthermore, the behavioral-data-only and the bivariate models predict greater probabilities that η is greater than zero compared to the bivariate analyses above, suggesting that the additional data is consistent with the subsets considered in the above sections.

The dominant factor contributing to the pattern of predictions observed in Figure 13 is the variance-covariance matrix Σ linking neural and behavioral model parameters. Recall that each shift effect uses a separate variance-covariance matrix, such that $\Sigma^{(1)}$ corresponds to the differences between conditions $V2$ and $V3$, and $\Sigma^{(2)}$ corresponds to the differences between conditions $V1$ and $V3$. In addition, each model's use of Σ varies by the type of neural measure. Specifically, the behavioral-data-only model does not use any neural data, and so it does not use a variance-covariance matrix at all. The bivariate EEG and fMRI models do capture the effects of the shifts between conditions, but in a way that is unique to the type of neural measure used. Hence, the two Σ matrices will be different across the two bivariate models. Finally, the trivariate model has larger Σ matrices to capture the associations between behavioral, EEG, and fMRI data sources. At the outset, we might expect the trivariate model's Σ matrix to resemble components of the bivariate models with corresponding parts. For example, the bivariate

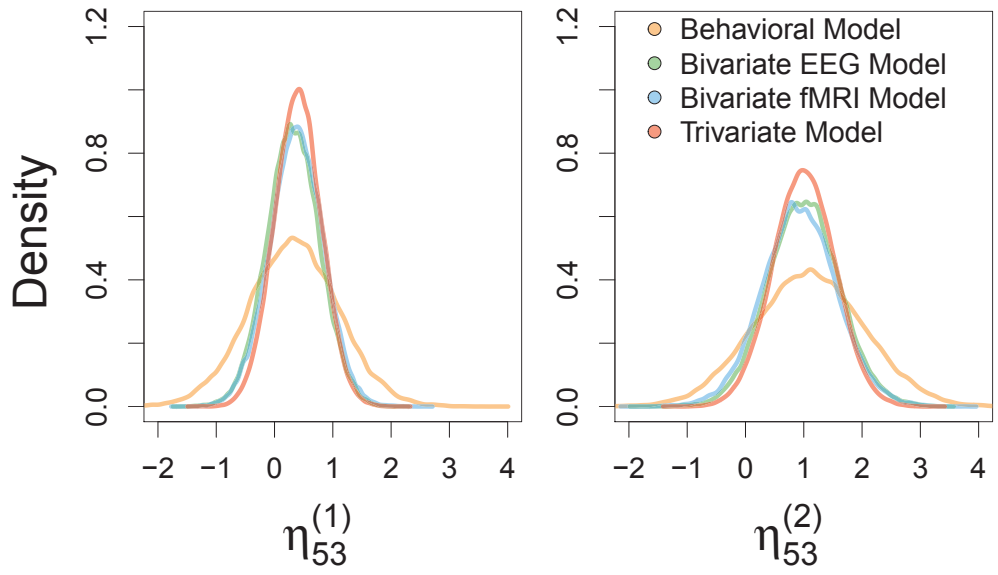

Figure 13: Comparison of the value effect parameters $\eta_{53}^{(1)}$ (left panel), and $\eta_{53}^{(2)}$ (right panel) obtained from the behavioral-data-only model (orange), the bivariate EEG model (green), the bivariate fMRI model (blue), and the trivariate model (red) for the withheld subject (i.e., Subject 53).

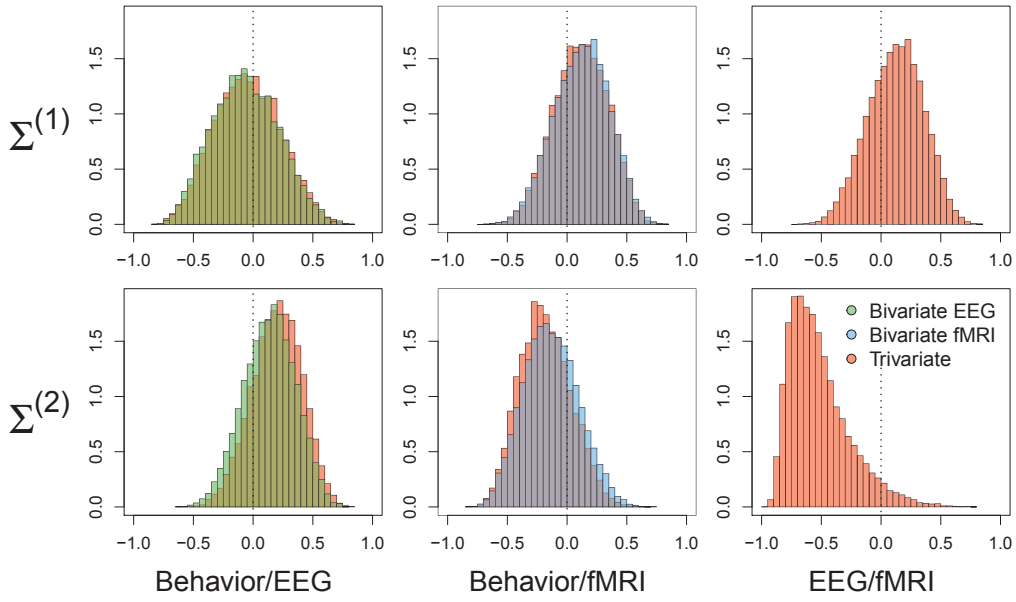

Figure 14: Estimated posterior distributions for the hyper correlation parameters. The top row shows the correlations between the additive shift parameters across the $V3$ and $V2$ conditions, whereas the bottom row shows the correlations between the shift parameters across $V3$ and $V1$. The first column shows the correlations between the behavioral and EEG shift parameters for the bivariate (green) and trivariate (red) joint models. The second column shows the correlations between the behavioral and fMRI shift parameters for the bivariate (blue) and trivariate (red) joint models. The third column shows the correlations between the EEG and fMRI shift parameters for the trivariate joint model. In each panel, a reference line is plotted to indicate zero correlation.

EEG model should have a Σ matrix that is similar to the components of the trivariate model that capture the correlations between behavior and EEG data. To examine this, we can separate out the effect parameters within Σ of the trivariate model, and compare them to similar effect parameters within Σ in the bivariate joint models.

Figure 14 shows the important components of the Σ matrices for the Conditions $V2$ and $V1$ ($\Sigma^{(1)}$; top row), and for the Conditions $V3$ and $V1$ ($\Sigma^{(2)}$; bottom row). Each column represents the corresponding correlation parameters relating the three measures: behavioral and EEG model parameters (left column), behavioral and fMRI model parameters (middle column), and EEG and fMRI model parameters (right column). The estimated correlation

parameters for each of the three joint models are illustrated with histograms of different colors: bivariate EEG as green, bivariate fMRI as blue, and the trivariate model as red. Figure 14 shows that components of the bivariate models do closely align with the corresponding components of the trivariate model; however, the variance of the estimate posterior distribution for the trivariate model is smaller, especially for the $\Sigma^{(2)}$ matrix. Akin to the other comparisons of behavioral-data-only and joint models above, this additional constraint comes from having additional measures in the model that are correlated. The right column of Figure 14 shows that the additive shift parameters in the EEG and fMRI data are uncorrelated across Conditions $V2$ and $V3$, but strongly negatively correlated across Conditions $V1$ and $V3$. Although this relationship is something that we might have logically deduced from two separate bivariate model analyses, only the trivariate model is able to use this information to provide statistical constraint in generating predictions for new data.

4.1.2. Predictive Modeling Analysis

We again performed a leave-one-out cross-validation test to assess the models' ability to generate predictions for behavioral data. To accomplish this, we focused on the eight subjects who provided both EEG and fMRI measures, so that the bivariate and trivariate models could be directly compared. First, we removed all data for one subject (e.g., Subject 47). Second, we fit all four models to the remaining 53 subjects' data. Third, we generate predictions for the withheld subject's behavioral data, conditional on some aspect of the neural data. For example, the bivariate EEG model only takes EEG data as an input, whereas the trivariate model can take both EEG and fMRI data as inputs. Finally, we examined the models' predictive accuracy by calculating the (log) likelihood of the withheld subject's data from the estimated posterior distributions via Monte Carlo procedure (Robert and Casella, 2004). Specifically, we integrated over the posterior distributions to numerically estimate the probability that any given observation for a particular subject was best predicted by a particular model. In the end, we arrived at an estimate of the probability that a particular subject's data was best captured by each model, a value that summed to one across models. Figure 15 shows a ternary plot of these probabilities for the eight subjects in our data. In the figure, we removed the behavioral-data-only model from consideration because (1) we have already established the bivariate model's ability to out predict the behavioral-data-only model, and (2) in this analysis, it

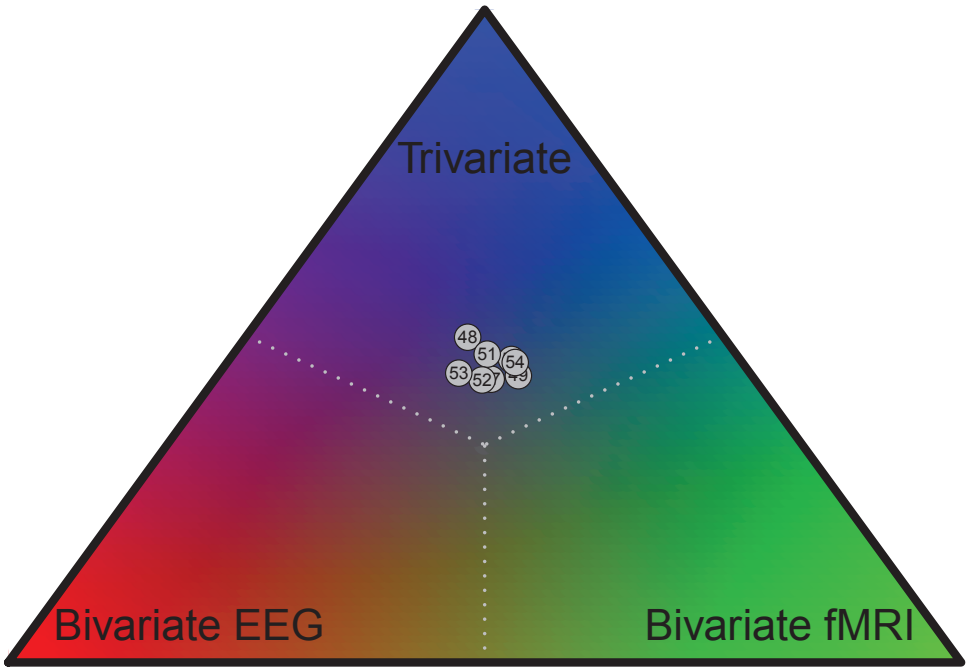

Figure 15: Results of the cross-validation test on the behavioral data from the last eight subjects of the experiment. The three vertices of the ternary plot each designate areas favoring one type of joint model: bivariate EEG (red; bottom left), bivariate fMRI (green; bottom right), and the trivariate model (blue; top). Reference lines are plotted to reflect points of indifference between the models. The nodes correspond to each of the eight subjects in the EEG/fMRI condition (i.e., Subjects 47-54). For each of the subjects, the trivariate model made the most accurate prediction.

performed the worse of the four models.

The reference lines in Figure 15 designate the areas best predicted by a particular model. In addition, the color values illustrate the graded nature of the model selection process. The trivariate model is represented as blue, the bivariate EEG model is represented as red, and the bivariate fMRI model is represented as green. The nodes in the graph mark the probability of a given model, conditional on the behavioral data. The figure shows that for each of the subjects in our data, the trivariate model provides the best predictions than any of the other three models (although the behavioral-data-only model is not shown).

Generalizations for Neural Data. Although Figure 15 shows the trivariate model's predictive abilities for behavioral data relative to the bivariate models, there are other predictive tests we could have performed. As a final test of the utility of our trivariate model, we assessed the model's ability to predict neural data on the basis of behavioral data. To do this, we refit the trivariate model to the full data set, but withheld the neural data for two randomly selected subjects: Subject 4 from the EEG experiment, and Subject 33 from the fMRI experiment. This withholding strategy is again illustrated by the orange blocks in Figure 2. We then generated predictions for the EEG data of Subject 4, and the fMRI data for Subject 33, on the basis of the behavioral data.

Figure 16 shows the predictions for the EEG data made for Subject 4. Panel a shows the PPDs for the difference between value conditions $V1$ and $V3$ (x-axis) against the difference between value conditions $V2$ and $V3$ (y-axis). Samples of the joint PPD are represented with shaded blue dots to show the relative density across the parameter space. Panel a also shows some summary statistics of Subject 4's EEG data, reflected in the "+" symbol: the mean along each dimension is located at the center, and one standard deviation of the PPD along each dimension is represented as the length of the line (i.e., the horizontal line corresponds to the x-axis, and the vertical line corresponds to the y-axis). For comparison, the mean of the PPD is represented as the open circle. Plotting the PPD in this way allows us to interpret the relative differences in activation predicted by the model between the two value conditions. From our previous analyses (also see Figure 6), we might expect the model to predict that raw activation in the EEG signal would be highest in $V1$, and lowest in $V3$, implying that the differences between $V1$ and $V3$ would be positive and larger than the differences between $V2$ and $V3$. Given this intuition, we would expect the largest density of the PPD to be on the gray triangle in Panel a. The PPD, having been centered relative to Condition $V3$, places more density on positive values in both the x and y directions. Further, in the upper right quadrant of Panel a, the PPD has greater density in the half-quadrant representing values where $V1-V3$ is higher than $V2-V3$. Importantly, the model's predictions are also consistent with Subject 4's data in that both the filled circle and the "+" symbol fall in the gray area in Panel a.

Panel b shows how these predictions appear on the scalp. The first column shows the mean of the PPD (represented as the filled circle in Panel a) in the $V2 - V3$ (top) and $V1 - V3$ (bottom) conditions. The second col-

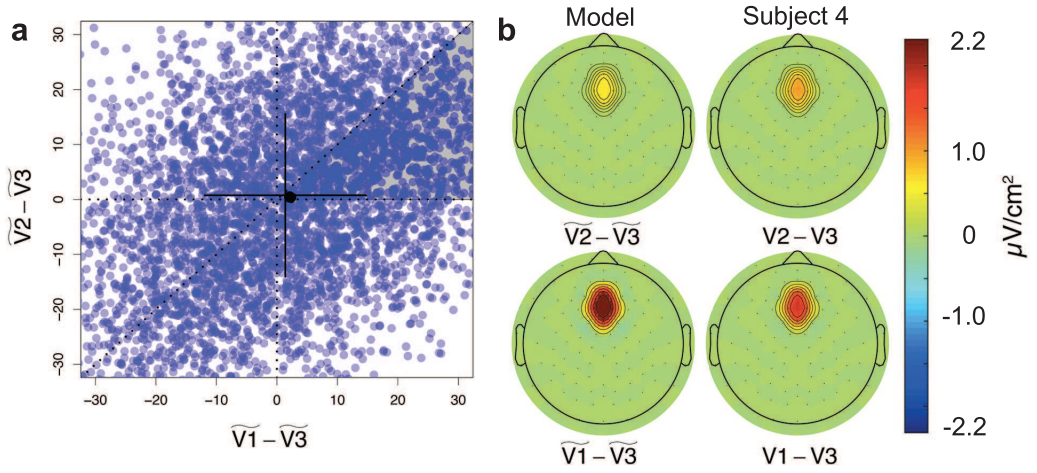

Figure 16: Posterior predictive distributions (PPDs) for the withheld EEG data of Subject 4. Panel a shows the full PPD for the difference between value conditions $V1$ and $V3$ (x-axis) against the difference between value conditions $V2$ and $V3$ (y-axis). The "+" symbol illustrates summary statistics of Subject 4's EEG data: the center designates the mean along each dimension (i.e., the x and y axes), and one standard deviation of the data is represented as the length of the line (i.e., the horizontal line corresponds to the x-axis, and the vertical line corresponds to the y-axis). The mean of the PPD from the model is represented as the filled circle. Reference lines are used to designate important areas relating differences between conditions (see text for details). Panel b illustrates these predictions on the scalp: the means of the PPDs are shown in the first column whereas Subject 4's data are shown in the second column. The topographic plots are colored coded according to the key on the far right side.

umn shows Subject 4's withheld data in the same conditions. In each scalp topography plot, activation is color coded according to the key on the right side. Comparing across panels, we conclude that the model's predictions are entirely consistent with the withheld subject's data.

A parallel analysis was performed for the withheld fMRI data of Subject 33. Similar to Figure 16, Figure 17a shows the PPD for the difference between value conditions $V1$ and $V3$ (x-axis) against the difference between value conditions $V1$ and $V2$ (y-axis). Summary statistics of Subject 33's fMRI data are plotted in Panel a of Figure 17 in the same way they were plotted in Panel a of Figure 16. Given our previous analyses, we would expect the opposite pattern of predictions from the EEG predictions above. That is, we would expect the activation to be highest in Condition $V3$, and lowest in

Condition $V1$, implying that the predictions for $V1 - V3$ would be negative, and more negative than the predictions for $V2 - V3$. Hence, we should expect a high density of the model predictions, and the withheld data to fall in the gray half-quadrant illustrated in Panel a. Figure 17 confirms this prediction and shows that the predictions from the trivariate model are indeed consistent with the withheld data.

Panel b shows the average model predictions from Panel a (left column) beside Subject 33's withheld fMRI data. The degree of activation in the dmFC is color coded according to the key on the far right side. Figure 17 shows that the model predicts that the difference in activation in the dmFC between value conditions $V2$ and $V3$ is smaller in magnitude than the difference in activation between value conditions $V1$ and $V3$. These predictions are in line with the withheld data, which further suggests that the trivariate model uses the inferred relationships between neural activity and the behavioral model to generalize accurately to new subjects.

5. Discussion

In this article, we have shown how the joint modeling framework (Turner et al., 2013a) can be extended to solve the more difficult challenge of integrating multiple neurophysiological measures with a behavioral model. Our results suggest that bivariate models, having either EEG or fMRI, can outperform behavioral-data-only models, and trivariate models, having both EEG and fMRI, can outperform bivariate models. The reason for this pattern of results is due to (1) the interrelationships among the three variables, and (2) the trivariate model's ability to learn the important patterns in the data to facilitate better generalization. While our results certainly support our method for multi-variate extensions, there are a number of important limitations and benefits meriting further discussion.

5.1. Reliance on Preprocessing Methodology

To date, all applications of our joint modeling framework rely on some preprocessing of the neural signal as a way to reduce the dimensionality of these measures. For example, in our processing of the EEG data above, we first computed the average EEG signal for each value condition and visually inspected these curves across time. We then chose a window of time where the EEG signal showed the most differentiation between the three

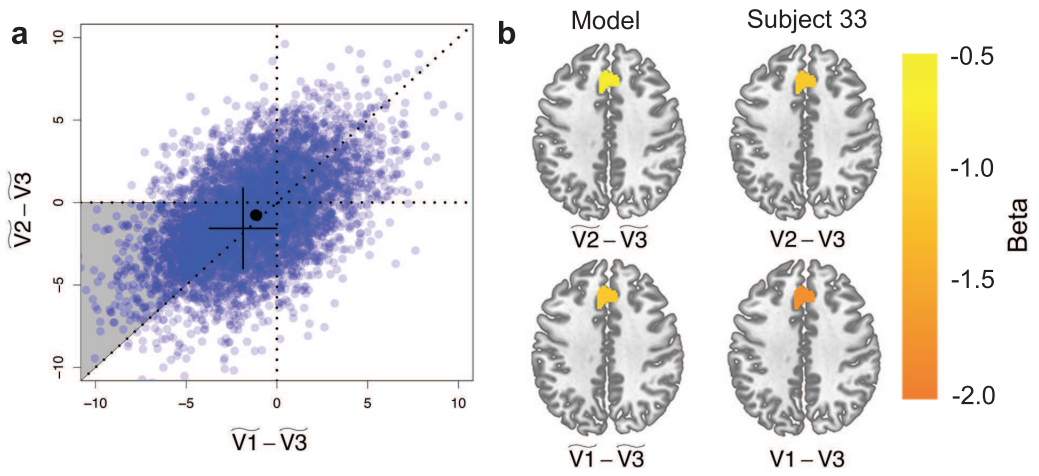

Figure 17: Posterior predictive distributions (PPDs) for the withheld fMRI data of Subject 33. Panel a shows the full PPD for the difference between value conditions $V1$ and $V3$ (x-axis) against the difference between value conditions $V2$ and $V3$ (y-axis). The "+" symbol illustrates summary statistics of Subject 33's fMRI data: the center designates the mean along each dimension (i.e., the x and y axes), and one standard deviation of the data is represented as the length of the line (i.e., the horizontal line corresponds to the x-axis, and the vertical line corresponds to the y-axis). The mean of the PPD from the model is represented as the filled circle. Reference lines are used to designate important areas relating the variables across conditions (see text for details). Panel b illustrates these predictions in the brain: the mean of the PPDs are shown in the first column whereas Subject 33's data are shown in the second column. In each axial slice, the degree of activation is colored coded according to the key on the far right side.

value conditions. This time window represented the period of time we suspected would be the most interesting from a cognitive perspective, but this procedure leaves much to be desired. First, it averages the wave form over time, and second, it averages over only four electrodes out of many. One could easily imagine the analyses performed here with improvements to the neural submodel such as decomposing the neural signal into its temporal and spatial aspects. Future work should endeavor to establish stronger data reduction techniques that are model driven, such as those used by topographic factor analysis (Manning et al., 2014; Gershman et al., 2011).

5.2. Generalization Depends on the Value of Conditioned Variables

Our bivariate and trivariate generative modeling analyses showed that adding neural measures constrains our knowledge about behavioral model parameters (see Figures 7, 10, and 13). In these analyses, constraint was observed through the additional precision (i.e., having less variance) in the posterior estimates, and importantly, this additional constraint facilitated more accurate predictions about withheld behavioral data. The additional constraint depends on a few important factors. First, consider the case of two arbitrary random variables (θ, δ) that have a joint multivariate normal distribution, as in the bivariate joint models presented above, where θ are the linked behavioral submodel parameters, and δ are the linked neural submodel parameters. When provided with neural data, the neural model parameters δ become *estimable*, allowing our model to infer an estimate for δ that comes from both the prior and the likelihood. In our modeling framework, having information about δ informs our understanding about θ. Recall that the distribution of (θ, δ) for all subjects depends on the hyper mean ϕ and the hyper variance-covariance matrix Σ. If we partition these hyperparameters according to what they correspond to in the behavioral and neural submodels, we obtain the following division:

$$\begin{aligned} \phi &= \left[\begin{array}{cc} \phi_1 & \phi_2 \end{array} \right]^{\mathsf{T}} \\ \Sigma &= \left[\begin{array}{cc} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{array} \right]. \end{aligned}$$

Here, parameters ϕ_1 and Σ_{11} correspond to behavioral submodel parameters θ, whereas parameters ϕ_2 and Σ_{22} correspond to neural submodel parameters δ. Given the properties of our linking function, if we know that δ equals a specific value, say $\delta = \delta^*$, then the distribution of θ *conditional* on δ is

$$\theta \big| \delta = \delta^* \sim \mathcal{N}\left(\phi^*, \Sigma^*\right),$$

where

$$\begin{aligned} \phi^* &= \phi_1 + \Sigma_{12}\Sigma_{22}^{-1}(\delta^* - \phi_2) \\ \Sigma^* &= \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}. \end{aligned}$$

The important result is that the variance of the conditional distribution is reduced, specifically it is reduced by $\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. While this amount does not depend on the specific value of δ^*, it clearly depends on δ's variability. Furthermore, as the correlation between the neural and behavioral data approaches zero, the joint model reduces to a univariate model, such that $\phi^* = \phi_1$ and $\Sigma^* = \Sigma_{11}$. The gain of this from a modeling standpoint is that the constraint applied to both θ and δ becomes negligible, and the model effectively learns to ignore the neural data in its estimate of θ. Finally, the mean of the conditional estimate for θ does depend on δ^*, which suggests that the influence and constraint provided by the neural data will vary from one subject to another, and this variation is explained by the properties of that subject's neural data *relative* to the group. Specifically, as a subject's neural data become more unlike the group data (i.e., as $\delta^* - \phi_2$ increases), the constraint applied to the conditional distribution $(\theta \big| \delta = \delta^*)$ increases. These results could have implications for a joint model's performance, and might explain why for some subjects, a joint model performed worse than a behavioral-data-only model (e.g., see Figure 7).

5.3. The Shrinkage Paradox

In the hierarchical models presented in this article, we assumed a hierarchical structure that captured both subject-specific and group-specific effects. One of the major benefits of hierarchical modeling is the "borrowing" of information from one subject (i.e., the lower level of the hierarchy) to another. This happens because the parameters for any given subject are informed by that subject's data, but the pattern of subject-specific parameters is informed through the estimation of group-level parameters. While this may seem like a bottom-up process, it is actually also top-down. For example, when a new subject is added to the data, the model can already generate predictions for that subject's data on the basis of the other subjects in the data set. In this way, the estimation of subject-specific parameters in a hierarchical model consists of a mixture of two quantities: (1) the information learned about the subject from the subject's own data, and (2) what the model predicts for the subject's parameters given the other subjects in the

data set. In statistical terms, these quantities are known as the likelihood function and the prior predictive distribution, respectively.

For the types of experiments we typically see in psychology and neuroscience, the prior predictive distribution can be quite sparse, having density in regions of the parameter space that are not informative for a specific subject. However, as the number of observations increase, the model's reliance on the prior predictive distribution diminishes because the likelihood function begins to dominate what we know about this individual. Said another way, the influence of the likelihood is proportional to the number of observations for that subject. In an analogous way, the influence of the prior predictive distribution is proportional to the number of subjects in the extant data set. With many subjects in the set, the prior predictive distribution will be more reflective of the average subject-to-subject patterns in the set.

The tradeoff dynamics between the prior predictive distribution and the likelihood function are important because they have a direct effect on a model's predictive accuracy. Consider a new subject who has just been added to a small (e.g., 20 subjects) set of data. Further suppose that the data for this subject are relatively sparse, consisting of only a few observations. If the first 20 subjects had been fit hierarchically, the model's prediction for the parameters of this new subject would primarily be generated from the prior predictive density as a result of this subject's sparse data. However, if we had estimated the new subject's parameters *independently* of the hierarchical model, the parameter estimates would only consist of the information in the data themselves, and would not have the information conveyed by the hierarchical model. If we then compared the estimates of this subject's parameters obtained from the hierarchical model to those obtained independently, we should expect some differences. A typical pattern of results is that the independently-obtained estimates are more extreme – meaning they depart from the group average – than the estimates put forth by the hierarchical model (Rouder and Lu, 2005; Rouder et al., 2005). This phenomenon is known as "shrinkage", and is sometimes regarded as a disadvantage of hierarchical modeling.

Shrinkage can be explained simply as a model's tendency to hedge its bets. If we know very little about a new subject, but we know a lot about what subjects typically look like, it would make sense to make predictions about a subject that reflect our most informed base of knowledge. Statistically speaking, this is actually the optimal behavior (Gelman et al., 2004; Christensen et al., 2011). Despite this, it can happen for some subjects that the predic-

tions are inaccurate because they are more reflective of group statistics than subject specific statistics. Models are especially susceptible to these errors when data are completely withheld. For example, in the trivariate modeling section we generate predictions for neural data on the basis of behavioral data alone. The predictions about neural data were based on three things: (1) the behavioral data for those subjects, (2) the prior predictive knowledge about the neural model parameters, and (3) the prior predictive knowledge of the relationship between behavioral and neural model parameters. None of these three things contribute to the likelihood function for the neural model parameters, and as a result, the model generates predictions that are reflective of the group statistics. This was not a problem in our prediction for EEG data (see Figure 16), but for the fMRI data, the predictions for the relative changes in BOLD activity across conditions were not as extreme as they should have been. In Figure 17, this is illustrated in Panel a by the distance of the "+" symbol from the point $(0, 0)$ relative to the filled circle. The interpretation is that the model made less extreme predictions than what was observed in the data. While these predictions are still accurate, it may not generally be the case, and it is important to keep the concept of shrinkage in mind when evaluating the predictive accuracy of joint models.

5.4. Contrasts with Data Fusion Methods

Combining EEG and fMRI measurements is a difficult, but fruitful endeavor. EEG measures have excellent temporal resolution, but have rather limited spatial information, whereas fMRI measures provide good spatial resolution, but lack good temporal resolution. Hence, combining both measures to understand brain function would capitalize on the benefits of each modality, potentially providing new insight to old questions. In addition to the methods of data fusion discussed in the introduction, another method attempting to combine these measures is through simultaneous recordings, where a subject performs a task in an MRI scanner while wearing an EEG net. This approach has yielded new methods for obtaining temporal components that can be associated with clusters of activation determined by fMRI that contribute to the generation of the electrical signal (see Huster et al., 2013, for a review), but does suffer from complications due to artifacts produced from the simultaneous methodology (e.g., Mulert and Lemieux, 2009; Ullsperger and Debener, 2010; Cottereau et al., 2015). By contrast, the method we have proposed here focuses on linking parameters of a cognitive model to both behavior and multiple brain measures. The method proposed

here could also be used with simultaneous data collection, but it need not be, obviating the difficulties of collecting high-qualtiy EEG data inside the scanner. However, our method is not without its own limitations. Specifically, our method assumes the same basic cognitive process underlies the decision making behavior observed in each of these experimental conditions, and that changes across experimental conditions manifests in the neural data in similar ways. There is some evidence that the context of experimental settings influences decision making behavior. Van Maanen et al. (2015) asked subjects to perform the three tasks in two separate conditions, one condition was performed in an fMRI scanner and the other was not. They then analyzed the behavioral data using the LBA model, and compared the parameter estimates across the two conditions for all three experiments. In two of the experiments, the nondecision time parameter was larger in the fMRI condition reflecting slower motor movements, and in one of these experiments, the degree of attentional focus was less in the fMRI condition. Although not reported (but see Rodriguez et al., 2014), we performed a similar analysis comparing parameter estimates across EEG and fMRI blocks in our data. Our model comparisons were not as extensive as in Van Maanen et al. (2015), but we found only marginal differences between the two modalities, and felt assured that our assumptions about the decision making behavior across conditions was justified. We speculated that the differences between our analysis and Van Maanen et al.'s analysis may be due to the subjective nature of the decision process used in our task. Regardless, future research using the joint modeling approach described here will need to properly scrutinize the difference in the decision making behavior across conditions. That is not to say that the decision making behavior needs to be identical across experimental conditions. Context-dependent effects like those observed in Van Maanen et al. (2015) can still be accommodated using our joint modeling approach, but may require additional theoretical overhead in the instantiation of the behavioral (or neural) submodel.

5.5. Neural Data as a Means for Model Selection

In this manuscript, we have only considered a single behavioral (sub)model to capture the behavioral data. However, the extent to which neural data correlate with model parameters is highly contingent on the fidelity of the model itself. One of the primary benefits of this modeling approach is that it is not committed to any particular submodel for either the neural or behavioral data. This lack of commitment manifests as a strength: one can

choose a submodel on the basis of convenience (e.g., mathematical tractability), theoretical endorsement, or simply personal preference. Being able to easily switch between different models also allows for a direct model comparison by way of fit statistics, and prediction performance – a feature of joint modeling that is similar in spirit to other integrative cognitive model comparison methods (Purcell et al., 2010; Mack et al., 2013).

We have speculated elsewhere how using the joint modeling framework can be used to facilitate a more elaborated model comparison analysis on the basis of neural data (cf. Turner, 2015). For example, Turner (2015) showed how tractography measures between the presupplementary motor area and the striatum could be used to better constrain model predictions for new subjects when behavioral data are withheld. Using the same neural measure, Turner (2015) showed how the LBA model and a reduced drift diffusion model compared in fits to both the neural and behavioral data. Ultimately, the LBA model outperformed the reduced drift diffusion model. While we speculate that neural data can provide greater constraint and benchmark tests of cognitive theory, future analyses are needed to better address the interplay of neuroscience and model selection (but see Purcell et al., 2010; Mack et al., 2013; Ditterich, 2010).

5.6. Toward a Common Theoretical Framework

In many ways, the typical experimental design in cognitive neuroscience is similar to those of experimental psychology. First, an experiment is designed with at least two conditions. We then go collect data to fill up the cell blocks in the design. Finally, we perform some statistical test to compare the distribution of data across the cell blocks. Such a procedure can tell us *which* brain regions/areas are sensitive to the experimental manipulation, and even the magnitude of the sensitivity, but this procedure can say nothing about *why* these regions change activity from a mechanistic point of view. Furthermore, while this analytic procedure is essential in statistically testing experimental manipulations within a study, it can sometimes be difficult to aggregate across multiple studies, especially when these studies use different types of neural measures, such as EEG and fMRI. We see this as a limitation, stemming directly from a lack of commitment to a particular theoretical framework. Our joint modeling approach attempts to resolve this limitation. By using a cognitive model, which ostensibly instantiates a cognitive theory, we are better postured to interpret neural function through the lens of a cognitive model. In addition, the joint modeling framework

allows for seamless integration across multiple studies, modalities, and even imbalanced experimental designs (see, e.g., Figure 2).

6. Conclusions

In this article, we have described a method for integrating EEG, fMRI, and behavioral data into one cognitive model. The model assumes the presence of different submodels that capture the modality-specific effects of interest. Important parameters describing changes across experimental conditions are then linked together to facilitate communication across modalities in the data set. Our approach has many advantages, in particular, it is amenable to imbalanced experimental designs including different modalities, and different numbers of subjects per modality. Our method is also well suited for generating predictions for new data, withheld data, or missing data for any modality of interest. Finally, our approach centers on the instantiation of a particular mechanistic cognitive theory, allowing us to interpret brain data through the lens of a cognitive model, and provide a unifying theoretical framework for multi-modal measures of cognition.

7. References

Amano, K., Goda, N., Nishida, S., Ejima, Y., Takeda, T., Ohtani, Y., 2006. Estimation of the timing of human visual perception from magnetoencephalography. Journal of Neuroscience 26, 3981–3991.

Astolfi, L., Cincotti, F., Mattia, D., Salinari, S., Babiloni, C., Basilisco, A., Rossini, P. M., Ding, L., Ni, Y., He, B., Marciani, M. G., Babiloni, F., 2004. Estimation of the effective and functional human cortical connectivity with structural equation modeling and directed transfer function applied to high-resolution EEG. Magnetic Resonance Imaging 22, 1457–1470.

Bai, X., Towle, V. L., He, E. J., He, B., 2007. Evaluation of cortical current density imaging methods using intracranial electrocorticograms and functional MRI. NeuroImage 35, 598–608.

Beckmann, C. F., Smith, S. M., 2005. Tensorial extensions of independent component analysis for multi-subject fMRI analysis. NeuroImage 25, 294–311.

Boehm, U., Van Maanen, L., Forstmann, B., Van Rijn, H., 2014. Trial-by-trial fluctuations in CNV amplitude reflect anticipatory adjustment of response caution. NeuroImage 96, 95–105.

Brown, S., Heathcote, A., 2008. The simplest complete model of choice reaction time: Linear ballistic accumulation. Cognitive Psychology 57, 153–178.

Busemeyer, J., Townsend, J., 1993. Decision Field Theory: A dynamic-cognitive approach to decision making in an uncertain environment. Psychological Review 100, 432–459.

Calhoun, V., Wu, L., Kiehl, K., Eichele, T., Pearlson, G., 2011. Aberrant processing of deviant stimuli in Schizophrenia revealed by fusion of fMRI and EEG data. Acta Neuropsychiatry 22, 127–138.

Calhoun, V. D., Adali, T., 2009. Feature-based fusion of medical imaging data. IEEE Transactions on Information Technology in Biomedicine 13, 711–720.

Calhoun, V. D., Adali, T., Liu, J., 2006. A feature-based approach to combine functional MRI, structural MRI and EEG brain imaging data. Proceedings of the 28th IEEE EMBS Annual International Conference.

Calhoun, V. D., Liu, J., Adali, T., 2009. A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic, and ERP data. NeuroImage 45, 163–172.

Cassey, P., Gaut, G., Steyvers, M., Brown, S., 2015. A generative joint model for spike trains and saccades during perceptual decision making, under review.

Christensen, R., Johnson, W., Branscum, A., Hanson, T. E., 2011. Bayesian Ideas and Data Analysis: An Introduction for Scientists and Statisticians. CRC Press, Taylor and Francis Group, Boca Ranton, FL.

Correa, N. M., Adali, T., Li, Y., Calhoun, V. D., 2010a. Canonical correlation analysis for data fusion and group inferences: Examining applications of medical imaging data. Signal Processing Magazine 27, 39–50.

Correa, N. M., Eichele, T., Adali, T., Li, Y., Calhoun, V. D., 2010b. Multi-set canonical correlation analysis for the fusion of concurrent single trial ERP and functional MRI. NeuroImage 50, 1438–1445.

Correa, N. M., Li, Y., Adali, T., Calhoun, V. D., 2008. Canonical correlation analysis for feature based fusion of biomedical imaging modalities and its application to detection of associative networks in Schizophrenia. Journal of Selected Topics in Signal Processing 2, 998–1007.

Cottereau, B. R., Ales, J. M., Norcia, A. M., 2015. How to use fMRI functional localizers to improve EEG/MEG source estimation. Journal of Neuroscience Methods 250, 64–73.

Dähne, S., Bieβmann, F., Samek, W., Haufe, S., Goltz, D., Gundlach, C., Villringer, A., Fazli, S., Müller, K.-R., 2015. Multivariate machine learning methods for fusing multimodal functional neuroimaging data. Proceedings of the IEEE 103, 1507–1530.

Daw, N. D., 2011. Trial-by-trial data analysis using computational models. Decision making, affect, and learning: Attention and performance XXIII 23, 1.

De Martino, F., Valente, G., de Borst, A. W., Esposito, F., Roebroeck, A., Goebel, R., Formisano, E., 2010. Multimodal imaging: An evaluation of univariate and multivariate methods for simultaneous EEG/fMRI. Magnetic Resonance Imaging 28, 1104–1112.

Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. Journal of Neuroscience Methods 134, 9–21.

Ditterich, J., 2010. A comparison between mechanisms of multi-alternative perceptual decision making: Ability to explain human behavior, predictions for neurophysiology, and relationship with decision theory. Frontiers in Neuroscience 4, 184.

Eichele, T., Calhoun, V. D., Debener, S., 2009. Mining EEG-fMRI using independent component analysis. International Journal of Psychophysiology 75, 53–61.

Eichele, T., Specht, K., Moosmann, M., Jongsma, M. L., Quiroga, R. Q., Nordby, H. e. a., 2005. Assessing the spatiotemporal evolution of neuronal activation with single-trial event-related potentials and functional MRI 102, 17798–17803.

Forstmann, B. U., Anwander, A., Schäfer, A., Neumann, J., Brown, S., Wagenmakers, E.-J., Bogacz, R., Turner, R., 2010. Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. Proceedings of the National Academy of Sciences 107, 15916–15920.

Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y., Ridderinkhof, K. R., Wagenmakers, E.-J., 2008. Striatum and pre-SMA facilitate decision-making under time pressure. Proceedings of the National Academy of Sciences 105, 17538–17542.

Forstmann, B. U., Tittgemeyer, M., Wagenmakers, E.-J., Derrfuss, J., Imperati, D., Brown, S., 2011a. The speed-accuracy tradeoff in the elderly brain: A structural model-based approach. Journal of Neuroscience 31, 17242–17249.

Forstmann, B. U., Wagenmakers, E.-J., 2015. An introduction to model-based cognitive neuroscience. Springer, New York, New York.

Forstmann, B. U., Wagenmakers, E.-J., Eichele, T., Brown, S., Serences, J. T., 2011b. Reciprocal relations between cognitive neuroscience an formal cognitive models: opposites attract? Trends in Cognitive Sciences 15, 272–279.

Franco, A. R., Ling, J., Caprihan, A., Calhoun, V. D., Jung, R. E., Heileman, G. L. e. a., 2008. Multimodal and multi-tissue measures of connectivity revealed by joint independent component analysis. Journal of Selected Topics in Signal Processing 2, 986–997.

Friston, K., Harisson, L., Penny, W., 2003. Dynamic causal modeling. NeuroImage 19, 1273–1302.

Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B., 2004. Bayesian Data Analysis. Chapman and Hall, New York, NY.

Gershman, S. J., Blei, D. M., Pereira, F., Norman, K. A., 2011. A topographic latent source model for fMRI data. Neuroimage 57, 89–100.

Green, D. M., Swets, J. A., 1966. Signal detection theory and psychophysics. Wiley Press, New York.

Hamandi, K., Powell, H. W., Laufs, H., Symms, M. R., Barker, G. J., Parker, G. J., Lemieux, L., Duncan, J. S., 2008. Combined EEG-fMRI and tractography to visualise propagation of epileptic activity. Journal of Neurology Neurosurgery Psychiatry 79, 594–597.

Heathcote, A., Brown, S. D., Wagemakers, E.-J., 2015. An introduction to good practices in cognitive modeling. In: Forstmann, B. U., Wagenmakers, E.-J. (Eds.), An introduction to model-based cognitive neuroscience. Springer, New York, pp. 25–48.

Ho, T., Brown, S., van Maanen, L., Forstmann, B. U., Wagenmakers, E.-J., Serences, J. T., 2012. The optimality of sensory processing during the speed-accuracy tradeoff. J Neurosci 32, 7992–8003.

Huster, R. J., Debener, S., Eichele, T., Herrmann, C. S., 2013. Methods for simultaneous EEG-fMRI: An introductory review. The Journal of Neuroscience 32, 6053–6060.

Kayser, J., Tenke, C. E., 2006. Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: II. Adequacy of low-density estimates. Clinical Neurophysiology 117, 369–380.

Kirby, K. N., 2009. One-year temporal stability of delay-discount rates. pbr 16, 457–462.

Krishnan, A., Williams, L. J., McIntosh, A., Abdi, H., 2015. Partial least squares (PLS) methods for neuroimaging: A tutorial and review. NeuroImage 56, 455–475.

Lee, M. D., Wagenmakers, E.-J., 2013. Bayesian modeling for cognitive science: A practical course. Cambridge University Press.

Lin, F. H., McIntosh, A. R., Agnew, J. A., Eden, G. F., Zeffiro, T. A., Belliveau, J. W., 2003. Multivariate analysis of neuronal interactions in the generalized partial least squares framework: Simulations and empirical studies. NeuroImage 20, 625–642.

Liu, J., Calhoun, V. D., 2007. Parallel independent component analysis for multimodal analysis:Application to fMRI and EEG data. Proceedings of IEEE ISBI conference.

Love, B. C., 2015. The algorithmic level is the bridge between computation and brain. Topics in Cognitive Science 7.

Mack, M. L., Preston, A. R., Love, B. C., 2013. Decoding the brain?s algorithm for categorization from its neural implementation. Current Biology 23, 2023–2027.

Manning, J. R., Ranganath, R., Norman, K. A., Blei, D. A., 2014. Topographic Factor Analysis: A Bayesian model for inferring brain networks from neural data. PLoS One 9, e94914.

Marr, D., 1982. Vision: A Computational Investigation into the Human Representation and Processing of Visual Information. New York: Freeman.

Martinez-Montes, E., Valdes-Sosa, P. A., Miwakeichi, F., Goldman, R. I., Cohen, M. S., 2004. Concurrent EEG/fMRI analysis by multiway partial least squares. NeuroImage 22, 1023–1034.

Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., Forstmann, B. U., 2012. Bias in the brain: A diffusion model analysis of prior probability and potential payoff. Journal of Neuroscience 32, 2335–2343.

Mulert, C., Lemieux, L., 2009. EEG-fMRI: Physiological basis, technique and applications, 1st Edition. Springer, Berlin.

Norman, K., Polyn, S., Detre, G., Haxby, J., 2006. Beyond mind-reading: Multi-voxel pattern analysis of fMRI data. Trends in Cognitive Sciences.

Nunez, M. D., Srinivasan, R., Vandekerckhove, J., 2015. Individual differences in attention influence perceptual decision making. Frontiers in Psychology 8 (18), 1–13.

Palmeri, T., Schall, J., Logan, G., 2015. Neurocognitive modelling of perceptual decisions. In: Busemeyer, J. R., Townsend, J., Wang, Z. J., Eidels, A. (Eds.), Oxford Handbook of Computational and Mathematical Psychology. Oxford University Press.

Plummer, M., Best, N., Cowles, K., Vines, K., March 2006. CODA: Convergence diagnosis and output analysis for MCMC. R News 6 (1), 7–11. URL `http://CRAN.R-project.org/doc/Rnews/`

Purcell, B., Heitz, R., Cohen, J., Schall, J., Logan, G., Palmeri, T., 2010. Neurally-constrained modeling of perceptual decision making. Psychological Review 117, 1113–1143.

Ratcliff, R., 1978. A theory of memory retrieval. Psychological Review 85, 59–108.

Robert, C. P., Casella, G., 2004. Monte Carlo statistical methods. Springer, New York, NY.

Rodriguez, C. A., Norcia, A. M., Cohen, M. X., McClure, S. M., 2015a. Frontoparietal cortex dynamics of value integration in intertemporal choice, manuscript in preparation.

Rodriguez, C. A., Turner, B. M., McClure, S. M., 2014. Intertemporal choice as discounted value accumulation. PLoS ONE 9, e90138.

Rodriguez, C. A., Turner, B. M., Van Zandt, T., McClure, S. M., 2015b. The neural basis of value accumulation in intertemporal choice. European Journal of Neuroscience.

Rouder, J. N., Lu, J., 2005. An introduction to Bayesian hierarchical models with an application in the theory of signal detection. Psychonomic Bulletin and Review 12, 573–604.

Rouder, J. N., Lu, J., Speckman, P., Sun, D., Jiang, Y., 2005. A hierarchical model for estimating response time distributions. Psychonomic Bulletin and Review 12, 195–223.

Shadlen, M. N., Newsome, W. T., 2001. Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. Journal of Neurophysiology 86, 1916–1936.

Shiffrin, R. M., Lee, M. D., Kim, W., Wagenmakers, E.-J., 2008. A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. Cognitive Science 32, 1248–1284.

Sui, J., Adali, T., Yu, Q., Chen, J., Calhoun, V. D., 2012. A review of multivariate methods for multimodal fusion of brain imaging data. Journal of Neuroscience Methods 204, 68–81.

Summerfield, C., Koechlin, E., 2010. Economic value biases uncertain perceptual choices in the parietal and prefrontal cortices. Front Hum Neurosci 4, 208.

Teipel, S. J., Bokde, A. L., Meindl, T., Amaro, E. J., Soldner, J., Reiser, M. F., Herpertz, S. C., M oller, H. J., Hampel, H., 2010. White matter microstructure underlying default mode network connectivity in the human brain. NeuroImage 49, 2021–2032.

Turner, B. M., 2015. Constraining cognitive abstractions through Bayesian modeling. In: Forstmann, B. U., Wagenmakers, E.-J. (Eds.), An introduction to model-based cognitive neuroscience. Springer, New York, pp. 199–220.

Turner, B. M., Forstmann, B. U., Love, B., Palmeri, T., Van Maanen, L., 2015a. Approaches of analysis in model-based cognitive neuroscience, manuscript under review.

Turner, B. M., Forstmann, B. U., Wagenmakers, E.-J., Brown, S. D., Sederberg, P. B., Steyvers, M., 2013a. A bayesian framework for simultaneously modeling neural and behavioral data. NeuroImage 72, 193–206.

Turner, B. M., Sederberg, P. B., Brown, S., Steyvers, M., 2013b. A method for efficiently sampling from distributions with correlated dimensions. Psychological Methods 18, 368–384.

Turner, B. M., Van Maanen, L., Forstmann, B. U., 2015b. Combining cognitive abstractions with neurophysiology: The neural drift diffusion model. Psychological Review 122, 312–336.

Ullsperger, M., Debener, S., 2010. Simultaneous EEG and fMRI: Recording, analysis, and application. Oxford University Press, USA.

Usher, M., McClelland, J. L., 2001. On the time course of perceptual choice: The leaky competing accumulator model. Psychological Review 108, 550–592.

van Maanen, L., Brown, S. D., Eichele, T., Wagenmakers, E.-J., Ho, T., Serences, J., 2011. Neural correlates of trial-to-trial fluctuations in response caution. Journal of Neuroscience 31, 17488–17495.

Van Maanen, L., Forstmann, B., Keuken, M., Wagenmakers, E.-J., Heathcote, A., 2015. The impact of MRI scanner environment on perceptual decision making, in Press.

White, C. N., Congdon, E., Mumford, J. A., Karlsgodt, K. H., Sabb, F. W., Freimer, N. B., London, E. D., Cannon, T. D., Bilder, R. M., Poldrack, R. A., 2014. Decomposing decision components in the stop-signal task: A model-based approach to individual differences in inhibitory control. Journal of Cognitive Neuroscience 26, 1601–1614.

Wilson, T. D., Houston, C. E., Etling, K. M., Brekke, N., 2015. Is model fitting necessary for model-based fMRI? PLoS Computational Biology 11, e1004237.

Xu, L., Pearlson, G., Calhoun, V. D., 2009. Joint source based morphometry identifies linked gray and white matter group differences. NeuroImage 44, 777–789.

Appendix A. Technical Details of the Behavioral Model

While the basic idea of our hierarchical LBA model for intertemporal choice was presented in Rodriguez et al. (2014) (also see Turner et al., 2013b), some of the core assumptions of the model have been changed to facilitate a comparison between the behavioral model and the joint models used in the subsequent sections.

Figure A.18 provides an illustrative diagram of our LBA model of intertemporal choice. To provide a formal description of the model, we denote the RT on the ith trial for the jth subject in the vth value condition as $RT_{i,j,v} \in (0, \infty)$, and the corresponding choice as $C_{i,j,v}$ where $C_{i,j,v} \in \{I, D\}$. I and D are the immediate and delayed rewards respectively. The model assumes that evidence for I and D is accumulated independently in separate accumulators. For a given trial, both accumulators begin with some choice bias, which is provided as independent amounts of starting point evidence $\{a_I, a_D\}$, sampled from a common uniform distribution $\mathcal{U}[0, A_j]$, where A_j is the upper bound of the starting point for Subject j. Evidence then increases through time at rates $\{d_I, d_D\}$, which are sampled from independent normal distributions with means $\{\mu_{v,I,j}, \mu_{v,D,j}\}$. Mean accumulation rates vary across value conditions (i.e., the first index) and subjects (i.e., the third index), whereas the standard deviation σ_j is the same for both immediate (I) and delayed (D) alternatives, but varies across subjects. Therefore, $d_I \sim \mathcal{N}(\mu_{v,I,j}, \sigma_j)$ and $d_D \sim \mathcal{N}(\mu_{v,D,j}, \sigma_j)$. Each accumulator gathers evidence ballistically until the point at which one accumulator reaches the response threshold b_j. The observed response time is the sum of the decision time plus some extra time τ_j attributed to non-comparison and selection processes, such as temporal discounting and motor execution. Letting $\{a_I, a_D\} = \mathbf{a}_{i,j,v}$ and $\{d_I, d_D\} = \mathbf{d}_{i,j,v}$, the response time on Trial i for Subject j is given by

$$RT_{i,j,v} = \min_{\{I,D\}} \left(\frac{b_j - \mathbf{a}_{i,j,v}}{\mathbf{d}_{i,j,v}} \right) + \tau_j, \tag{A.1}$$

and the choice $C_{i,j,v}$ corresponds to the accumulator that reached the response threshold b_j first. The model provides a closed-form and joint account of response time and choice probability across value conditions by specifying "defective" probability density functions (PDFs) for I and D in terms of the parameters just described (see Brown and Heathcote, 2008, for details).

We made two reparameterization choices from our previous application of the model (Rodriguez et al., 2014) to facilitate the hierarchical estimation

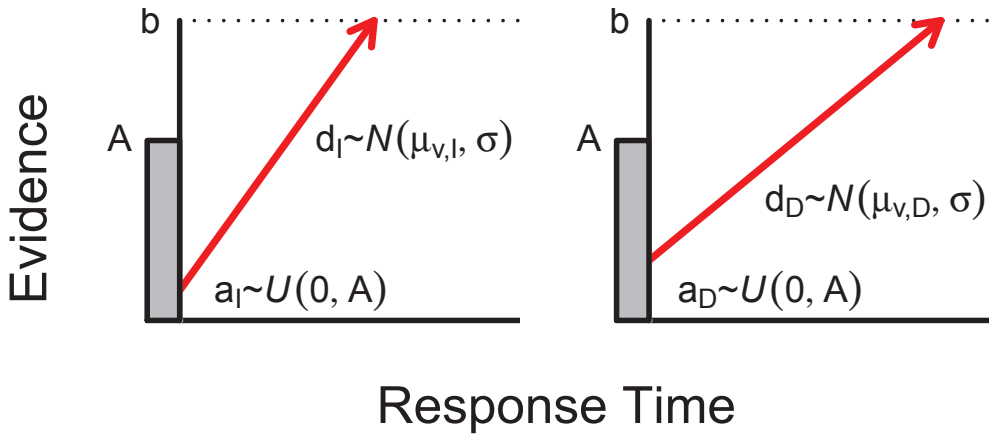

Figure A.18: Our Linear Ballistic Accumulator model for intertemporal choice. Following the presentation of a stimulus and some non-decision time τ, information accumulates ballistically for each alternative. The accumulation processes for both the immediate (left panel) and the delayed (right panel) reward alternatives are independent. Once enough evidence b has been accumulated, a decision is made that corresponds to the accumulator that reaches the threshold first. The model assumes trial-to-trial variation in both starting point and drift rate.

procedure (see below). The first involved modeling the upper boundary of the starting point A_j in terms of the response boundary b_j. The starting point parameter A_j is bounded by 0 and b_j, and this parameter-specific boundary can result in poor estimation efficiency in our sampling procedure. Consequently, we chose to model the proportion of the upper starting point boundary parameter to the response threshold parameter by estimating

$$\alpha_j = A_j / b_j.$$

The parameter α_j is still bounded by 0 and 1, and so in the specification of the model below, we will use the logit transformation to put α_j on a space that has infinite support (i.e., $\alpha_j \in (-\infty, \infty)$), which will improve the accuracy of our estimated posteriors.

Following our folding process described in the methods, our data consisted of three response time distributions, one for each of the value conditions V (see Figure 5). Our second reparametrization involved the drift rate parameters for these response time distributions. Of considerable importance was the difference in the accumulation rates for the most likely alternative (given our experimental manipulation) across these value conditions. To model this effect, we introduced the parameters $\eta_j^{(1)}$, and $\eta_j^{(2)}$ to capture the difference in the rates of accumulation from the $V3$ condition to the $V2$, and from the $V3$ condition to the $V1$ condition, respectively. Formally, we set

$$\begin{aligned}
\mu_{V3,SHV,j} &= \text{logit}^{-1}\left(\nu_j\right), \\
\mu_{V2,SHV,j} &= \text{logit}^{-1}\left(\nu_j + \eta_j^{(1)}\right), \text{ and} \\
\mu_{V1,SHV,j} &= \text{logit}^{-1}\left(\nu_j + \eta_j^{(2)}\right),
\end{aligned}$$

where $\text{logit}^{-1}(x)$ represents the inverse logit transformation of x, ν_j is the "equipreference" drift rate parameter for Subject j, and "SHV" is the subjectively higher valued option. To satisfy scaling conditions of the model, we imposed a constraint such that the drift rates sum to one across each of the value conditions, namely that in condition V, the following should hold:

$$\mu_{V,SHV,j} + \mu_{V,SLV,j} = 1.$$

We estimated LBA model parameters using a hierarchical Bayesian procedure (see Rodriguez et al., 2014; Turner et al., 2013b, for details). To

construct the hierarchical model, we must first make assumptions about the distributions of subject-specific parameters. To do this, we first transformed the parameters so that they had continuous, infinite support (i.e., can take on any real value). In our case, for parameters bounded by zero, we applied a log transformation, and for parameters bounded by zero and one, we applied a logit transformation. Following these transformations, we specified the following priors for the subject-specific parameters:

$$\begin{aligned}
\nu_j &\sim \mathcal{N}(\nu_\mu, \nu_\sigma), \\
\eta_j^{(1)} &\sim \mathcal{N}(\eta_\mu^{(1)}, \eta_\sigma^{(1)}), \\
\eta_j^{(2)} &\sim \mathcal{N}(\eta_\mu^{(2)}, \eta_\sigma^{(2)}), \\
\log(\sigma_j) &\sim \mathcal{N}(\sigma_\mu, \sigma_\sigma), \\
\log(\tau_j^{(k)}) &\sim \mathcal{N}(\tau_\mu^{(k)}, \tau_\sigma^{(k)}), \\
\text{logit}(\alpha_j) &\sim \mathcal{N}(\alpha_\mu, \alpha_\sigma), \text{ and} \\
\log(b_j) &\sim \mathcal{N}(b_\mu, b_\sigma),
\end{aligned}$$

where $k \in \{1, 2, 3\}$ represents the three different folded value conditions, as described above. The nondecision time parameter τ_j was allowed to vary across value conditions because previous modeling results showed that this choice provided the best fit to the data (Rodriguez et al., 2014).

Finally, we complete our model by specifying priors for the hyperparameters. Given the assumptions of normality for the subject-specific parameters, we can establish a conjugate relationship between the prior and posterior distributions at the hyper level. To do this, we specified the following priors for the group-level means:

$$\begin{aligned}
\nu_\mu &\sim \mathcal{N}(0.5, 0.8), \\
\eta_\mu^{(1)} &\sim \mathcal{N}(0, 100), \\
\eta_\mu^{(2)} &\sim \mathcal{N}(0, 100), \\
\sigma_\mu &\sim \mathcal{N}(0.5, 0.8), \\
\tau_\mu^{(k)} &\sim \mathcal{N}(0.75, 0.8), \\
\alpha_\mu &\sim \mathcal{N}(1.5, 0.8), \text{and} \\
b_\mu &\sim \mathcal{N}(1.5, 0.8),
\end{aligned}$$

and the following priors for the group-level standard deviations,

$$\begin{aligned}
\nu_\sigma &\sim \Gamma^{-1}(4, 10), \\
\eta_\sigma^{(1)} &\sim \Gamma^{-1}(4, 10), \\
\eta_\sigma^{(2)} &\sim \Gamma^{-1}(4, 10), \\
\sigma_\sigma &\sim \Gamma^{-1}(4, 10), \\
\tau_\sigma^{(k)} &\sim \Gamma^{-1}(4, 10), \\
\alpha_\sigma &\sim \Gamma^{-1}(4, 10), \text{and} \\
b_\sigma &\sim \Gamma^{-1}(4, 10),
\end{aligned}$$

where $\Gamma^{-1}(a, b)$ denotes the inverse gamma distribution with shape parameter a, and scale parameter b. This particular choice of a and b for the priors produces a skewed distribution placing weight on appropriate ranges of the parameters. Specifically, when $a = 4$ and $b = 10$, the prior has an approximate 95% credible set of (1.14, 9.05), and an expected value of 3.32. These choices reflect our *a priori* beliefs: we did not expect the between-subject variability to be less that 1, and felt that larger values would become increasingly less likely to account for these data.

While our prior selections were informed by other similar modeling applications (see, e.g., Rodriguez et al., 2014; Turner et al., 2013b), we remained somewhat conservative in our choices to avoid undue parameter constraint, because the model structure was considerably different from prior research using the hierarchical version of the LBA model.

Appendix B. Technical Details of the Bivariate Joint Model

The joint model used in both the EEG and fMRI analyses was equivalent, although both the neural and behavioral data changed. Where possible, we kept the behavioral portion of the joint model equal to the behavioral-data-only model, described above. However, given the joint structure of some components of the model, the models are not exactly equivalent in the way they capture the data.

We begin by discussing the neural portion of the joint model. We first processed the neural data to provide a single-trial measure of neural activity for each individual subject (see the main text for details). We denote the neural data on the ith trial for the jth subject in the vth value condition as $N_{i,j,v}$. Because we folded the data, $v \in \{1, 2, 3\}$, where $v = 3$ when $P_D =$

$\{0.5\}$, $v = 2$ when $P_D = \{0.3, 0.7\}$, and $v = 1$ when $P_D = \{0.1, 0.9\}$. Because the neural data for each trial is an average, by the central limit theorem, each $N_{i,j,v}$ is normally distributed. For the $v = 3$ condition, the parameter ϵ_j captures the mean neural activation. Similar to the η parameters in the behavioral model, we modeled the difference in neural activation across value conditions by introducing the variables $\delta_j^{(1)}$ and $\delta_j^{(2)}$ for the $v = 2$ and $v = 1$ value conditions, respectively. We also capture the variability in neural activation with the parameter ζ_j, which was constrained to be the same across each value condition. Formally, we assume the neural data for the jth subject in the vth value condition arise from the following distributions:

$$\begin{aligned} N_{i,j,3} &\sim \mathcal{N}(\epsilon_j, \zeta_j), \\ N_{i,j,2} &\sim \mathcal{N}(\epsilon_j + \delta_j^{(1)}, \zeta_j), \text{ and} \\ N_{i,j,1} &\sim \mathcal{N}(\epsilon_j + \delta_j^{(2)}, \zeta_j). \end{aligned}$$

For each subject, we then assumed the following priors for ϵ_j and ζ_j:

$$\begin{aligned} \epsilon_j &\sim \mathcal{N}(\epsilon_\mu, \epsilon_\sigma), \text{ and} \\ \log(\zeta_j) &\sim \mathcal{N}(\zeta_\mu, \zeta_\sigma). \end{aligned}$$

For the group-level parameters, we specified the following priors for the group-level means:

$$\begin{aligned} \epsilon_\mu &\sim \mathcal{N}(0, 10), \text{ and} \\ \zeta_\mu &\sim \mathcal{N}(0, 10), \end{aligned} \tag{B.1}$$

and the following priors for the group-level standard deviations

$$\begin{aligned} \epsilon_\sigma &\sim \Gamma^{-1}(7, 90), \text{ and} \\ \zeta_\sigma &\sim \Gamma^{-1}(7, 90). \end{aligned}$$

The principle behind joint modeling is to use both neural and behavioral data to enforce constraint on a particular cognitive theory (Turner et al., 2013a; Turner, 2015). In our modeling application, we wish to use the neural signature across value conditions to constrain the differences in the rate of accumulation across these value conditions. Following this motivation, we specify a link between the parameters $\delta_j^{(1)}$ and $\eta_j^{(1)}$, as well as between the

parameters $\delta_j^{(2)}$ and $\eta_j^{(2)}$, by specifying

$$\begin{aligned}\left(\delta_j^{(1)}, \eta_j^{(1)}\right) &\sim \mathcal{N}_2\left(\phi^{(1)}, \Sigma^{(1)}\right), \text{ and} \\ \left(\delta_j^{(2)}, \eta_j^{(2)}\right) &\sim \mathcal{N}_2\left(\phi^{(2)}, \Sigma^{(2)}\right),\end{aligned}$$

where $\mathcal{N}_p(a, b)$ represents the multivariate normal distribution of dimension p with mean vector a and variance-covariance matrix b. Note that $\phi^{(k)} = \left[\delta_\mu^{(k)}, \eta_\mu^{(k)}\right]$, where $\eta_\mu^{(k)}$ is the critical behavioral model group-level parameter by which we draw a comparison to the behavioral-data-only model (see Figures 7 and 10). Specifying the structure in this way is equivalent to the following specification:

$$\left(\delta_j^{(1)}, \eta_j^{(1)}, \delta_j^{(2)}, \eta_j^{(2)}\right) \sim \mathcal{N}_4\left(\phi^*, \Sigma^*\right),$$

where

$$\phi^* = \left[\phi^{(1)}, \phi^{(2)}\right]^\intercal = \left[\delta_\mu^{(1)}, \eta_\mu^{(1)}, \delta_\mu^{(2)}, \eta_\mu^{(2)}\right]^\intercal$$

$$\Sigma^* = \begin{bmatrix} \Sigma^{(1)} & 0 \\ 0 & \Sigma^{(2)} \end{bmatrix} = \begin{bmatrix} \left(\delta_\sigma^{(1)}\right)^2 & \delta_\sigma^{(1)}\eta_\sigma^{(1)}\rho_1 & 0 & 0 \\ \delta_\sigma^{(1)}\eta_\sigma^{(1)}\rho_1 & \left(\eta_\sigma^{(1)}\right)^2 & 0 & 0 \\ 0 & 0 & \left(\delta_\sigma^{(2)}\right)^2 & \delta_\sigma^{(2)}\eta_\sigma^{(2)}\rho_2 \\ 0 & 0 & \delta_\sigma^{(2)}\eta_\sigma^{(2)}\rho_2 & \left(\eta_\sigma^{(2)}\right)^2 \end{bmatrix}$$

where ρ_1 and ρ_2 are correlation parameters. However, as we will see below, with the appropriate selections for the prior distributions, one can establish a conjugate relationship between the prior and posterior of $\phi^{(k)}$ and $\Sigma^{(k)}$, making the estimation more efficient via Gibbs sampling (Robert and Casella, 2004; Gelman et al., 2004). With the exception of $\eta_j^{(1)}$ and $\eta_j^{(2)}$, all other parameter specifications for the behavioral submodel were equivalent to the behavioral-data-only model presented above.

The final step in constructing the joint model is to specify priors for the hyperparameters $\phi^{(k)}$ and $\Sigma^{(k)}$ for $k \in \{1, 2\}$. Following Turner et al. (2015b), we specified a joint prior on $\phi^{(k)}$ and $\Sigma^{(k)}$ so that

$$p\left(\phi^{(k)}, \Sigma^{(k)}\right) = p\left(\phi^{(k)} | \Sigma^{(k)}\right) p\left(\Sigma^{(k)}\right),$$

where

$$\begin{aligned} \phi^{(k)}|\Sigma^{(k)} &\sim \mathcal{N}_2\left(\mu_0, s_0^{-1}\Sigma^{(k)}\right), \text{ and} \\ \Sigma^{(k)} &\sim \mathcal{W}^{-1}(\Phi, d_0), \end{aligned}$$

where $\mathcal{W}^{-1}(a, b)$ denotes the inverse Wishart distribution with dispersion matrix a and degrees of freedom b. We set d_0 equal to the number of linked parameters plus two (i.e., $d_0 = 2 + 2 = 4$), Φ is the identity matrix (i.e., a matrix containing ones on the diagonal, and zeros on the off diagonal) of dimension (2×2), $s_0 = 1/10$, and μ_0 is a vector containing two zeros. These choices were made to establish a conjugate relationship between the prior and posterior, so that analytic expressions could be derived for the conditional distributions of $\phi^{(k)}$ and $\Sigma^{(k)}$, while still specifying uninformative priors.

When estimating the joint posterior distribution of the bivariate model parameters, we used a combination of Gibbs sampling for the group-level (Gelman et al., 2004), and differential evolution with Markov chain Monte Carlo for the subject-level (DE-MCMC; Turner et al., 2013b). For the subject level estimates, we used 24 chains, and obtained 10,000 samples after a burn-in period of 10,000 samples. Thus, our estimates of the joint posterior distributions are based on 240,024 samples. The burn-in period allowed us to converge quickly to the high-density regions of the posterior distribution, while the rest of the samples allowed us to improve the reliability of the estimates. Convergence and overall fidelity of the chains were assessed through visual inspection and the coda package in R (Plummer et al., 2006).

Appendix C. Technical Details of the Trivariate Joint Model

Where possible, the specification of the trivariate model was equivalent to the bivariate model above. The important difference between these two models is the additional data structure, with the trivariate model containing both EEG and fMRI. Using the notation in Figure 1, we denote the EEG data as $E_{i,j,v}$, and the fMRI data as $F_{i,j,v}$. In the model, we super script the parameters ϵ_j, ζ_j, $\delta_j^{(1)}$, and $\delta_j^{(2)}$ with an F for the fMRI data and an E for the EEG data. Hence, we assume the neural data arise from the following

distributions:

$$
\begin{aligned}
F_{i,j,3} &\sim \mathcal{N}\left(\epsilon_j^{(F)}, \zeta_j^{(F)}\right), \\
F_{i,j,2} &\sim \mathcal{N}\left(\epsilon_j^{(F)} + \delta_j^{(1,F)}, \zeta_j^{(F)}\right), \\
F_{i,j,1} &\sim \mathcal{N}\left(\epsilon_j^{(F)} + \delta_j^{(2,F)}, \zeta_j^{(F)}\right), \\
E_{i,j,3} &\sim \mathcal{N}\left(\epsilon_j^{(E)}, \zeta_j^{(E)}\right), \\
E_{i,j,2} &\sim \mathcal{N}\left(\epsilon_j^{(E)} + \delta_j^{(1,E)}, \zeta_j^{(E)}\right), \text{ and} \\
E_{i,j,1} &\sim \mathcal{N}\left(\epsilon_j^{(E)} + \delta_j^{(2,E)}, \zeta_j^{(E)}\right).
\end{aligned}
$$

The prior structure on $\epsilon_j^{(E)}$, $\epsilon_j^{(F)}$, $\zeta_j^{(E)}$, and $\zeta_j^{(F)}$ for the trivariate model were identical to the prior structure on ϵ_j and ζ_j in the bivariate model above. However, in the trivariate model, the linking procedure relating neural and behavioral submodels is slightly different. Namely, we assume

$$
\begin{aligned}
\left(\delta_j^{(1,F)}, \delta_j^{(1,E)}, \eta_j^{(1)}\right) &\sim \mathcal{N}_3\left(\phi^{(1)}, \Sigma^{(1)}\right), \text{ and} \\
\left(\delta_j^{(2,F)}, \delta_j^{(2,E)}, \eta_j^{(2)}\right) &\sim \mathcal{N}_3\left(\phi^{(2)}, \Sigma^{(2)}\right).
\end{aligned}
$$

The prior structure on ϕ and Σ in the trivariate model is identical to the prior structure in the bivariate model above, with the following exceptions: $d_0 = 3 + 2 = 5$, Φ is the identity matrix of dimension (3×3), and μ_0 is a vector containing three zeros.

When estimating the joint posterior distribution of the trivariate model parameters, we ran our sampling algorithm with 24 chains for 20,000 samples after a burnin period of 10,000 samples, and then thinned the samples by retaining every other iteration. This process helped to reduce autocorrelation in the chains, which we suspected would arise due to the high dimensionality of the model. Hence, the estimates of the posterior distribution of the trivariate model were also based on 240,024 samples, to maintain consistency with the bivariate models. Convergence and overall fidelity of the chains were assessed through visual inspection and the coda package in R (Plummer et al., 2006).

Highlights

> We present a method for integrating EEG, fMRI and behavior into a single analysis.

> The method uses neural data to constrain behavioral models.

> The method allows for any behavioral model to be combined with multiple neural measures.