

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 and Vasav Chodisetti for helpful discussions that improved an earlier draft of this manuscript.

¹ **1. Introduction**

² The field of cognitive science is faced with many options for studying how
³ experimentally-derived variables are systematically related to the dynamics
⁴ underlying a cognitive process of interest. To date, much of our understand-
⁵ ing of cognition has been advanced by two dominant, but non-interacting
⁶ groups. The largest group, cognitive neuroscientists, rely on statistical mod-
⁷ els to understand patterns of neural activity. These models are typically
⁸ purely data-mining techniques, and often disregard the computational mech-
⁹ anisms that might detail a cognitive process. The other group, mathematical
¹⁰ psychologists, is strongly motivated by *theoretical* accounts of cognitive pro-
¹¹ cesses, and instantiates these theories by developing formal mathematical
¹² models of cognition. The models often assume a system of computations
¹³ and equations intended to characterize the process assumed to take place in
¹⁴ the brain. To formally test their theory, mathematical psychologists rely on
¹⁵ their model's ability to fit behavioral data. A good fit is thought to reflect
¹⁶ an accurate theory, whereas a bad fit would refute it.

¹⁷ Although both groups are concerned with explaining behavior, they tend
¹⁸ to approach the challenge from different vantage points. Thinking in terms of
¹⁹ Marr (1982)'s levels of analysis, mathematical psychologists tend to focus on
²⁰ the computational and algorithmic levels, whereas cognitive neuroscientists
²¹ focus more on the implementation level. Although progress can be made by
²² maintaining a tight focus on one level, certain opportunities are missed. As
²³ a result of their single-level focus, both approaches suffer from critical limita-
²⁴ tions (Love, 2015), and these limitations have inspired researchers to combine
²⁵ neural and behavioral measures in an integrative fashion. The primary pay-
²⁶ off of this endeavor is the localization of mechanisms assumed to underly the
²⁷ computations supporting task-specific behavior within the brain. The impor-
²⁸ tance of solving the integration problem has spawned several entirely new
²⁹ statistical modeling approaches developed through collaborations between
³⁰ mathematical psychologists and cognitive neuroscientists, collectively form-
³¹ ing a new field often referred to as "model-based cognitive neuroscience" (e.g.,
³² Forstmann and Wagenmakers, 2015; Daw, 2011; Forstmann et al., 2011b; van
³³ Maanen et al., 2011; Turner et al., 2013a; Mack et al., 2013; Boehm et al.,
³⁴ 2014; Love, 2015; Palmeri et al., 2015; Turner et al., 2015b).

³⁵ So far, the field of model-based cognitive neuroscience has established
³⁶ links between both structural measures such as diffusion-weighted imaging
³⁷ (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

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

80 We begin by first describing the joint modeling framework and explaining
81 how it can be extended to multiple neural measures. We then describe
82 an experiment on an intertemporal choice task that was collected for the
83 purposes of testing the framework on imbalanced experimental designs. To
84 provide clarity about our modeling approach, we first describe the exper-
85 imental effects present in the EEG, fMRI, and behavioral data separately.
86 We then illustrate some of the benefits of the joint modeling framework by
87 showing how a “bivariate” model taking into account only one functional
88 measure can be used to boost performance of a behavioral-data-only model.
89 In all of our analyses, we show this boost from both a generative model-
90 ing perspective through analyses of the estimated parameters (i.e., posterior
91 distributions), and a predictive modeling perspective by assessing the mod-
92 els’ performance in a cross-validation test. We then show how the extended
93 “trivariate” model that incorporates both functional neural measures further
94 enhances the performance of the model relative to both bivariate models
95 and the behavioral-data-only model. We close with a discussion of limita-
96 tions and future directions for our approach to integrating data from multiple
97 modalities.

98 *1.1. Joint Modeling*

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

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

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

117 First, the behavioral data B are described in terms of a behavioral model,
118 such as the Linear Ballistic Accumulator (LBA; Brown and Heathcote, 2008)
119 model or classic signal detection theory model (Green and Swets, 1966). Im-
120 portantly, the behavioral model consists of a set of model parameters θ , which
121 are of lower dimensionality than the full behavioral data set B . For example,
122 in the SDT model, θ would consist of parameters such as discriminability (d')
123 and bias (β). Ideally, the chosen behavioral model should consist of mech-
124 anisms that provide insight into how the data arise, and how experimental
125 manipulations predictably affect these mechanisms (a property known as se-
126 lective influence; Heathcote et al., 2015).

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

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

141 The joint modeling framework extends other exploratory type approaches
142 in model-based cognitive neuroscience (see Turner et al., 2015a) by formal-
143 izing these three components into a single model. Furthermore, fitting the
144 joint model via Bayesian statistics allows the model to inherit many of the ad-
145 vantages of Bayesian modeling, which have been discussed by many authors
146 (e.g., Shiffrin et al., 2008; Lee and Wagenmakers, 2013). In the context of
147 model-based cognitive neuroscience, most of these advantages center around
148 flexibility. First, joint models are agnostic – there is no commitment to any
149 particular behavioral model, neural model, or linking function, and similarly,
150 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

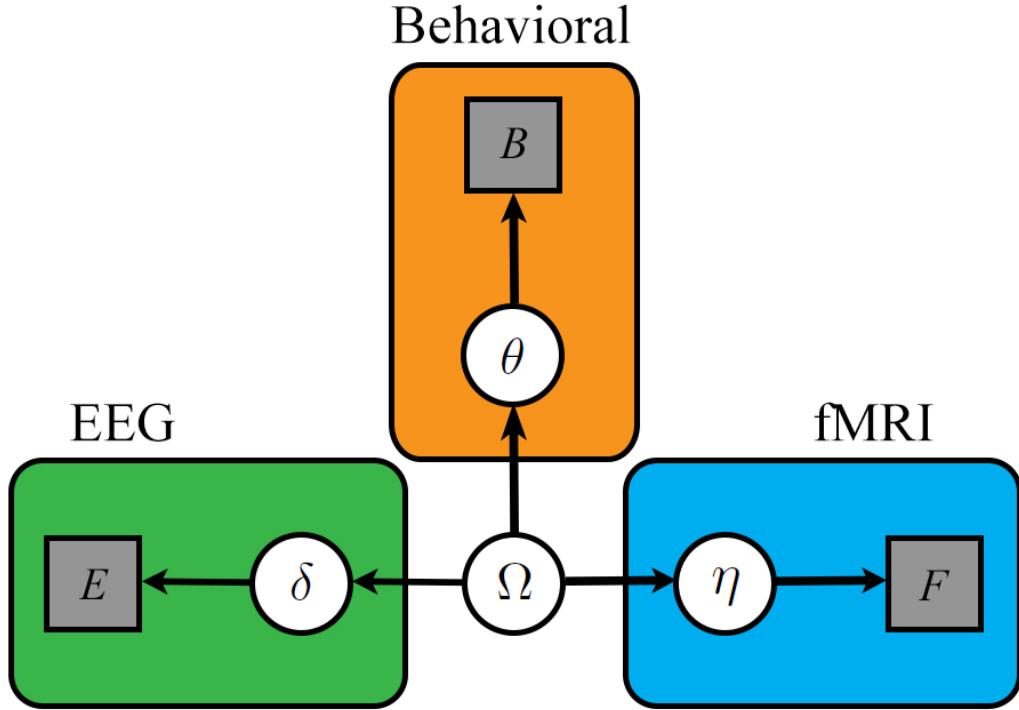


Figure 1: Graphical diagram of models presented in the text. Observable data are represented as gray boxes, whereas unknown (latent) variables are represented as empty circles. The orange plate represents the behavioral data/model, the green plate represents the EEG data/model, and the blue plate represents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

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

205 2. Materials and Methods

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

224 2.1. Data collection

225 2.1.1. Subjects

226 Sixty-one healthy adults participated in this study (32 females, ages 19-
227 46 years, median 24 years). All subjects gave written informed consent.
228 Stanford University’s Institutional Review Board approved the study. Two
229 subjects were excluded because their behavior did not allow us to estimate re-
230 liable temporal discounting parameters. Another five subjects were excluded
231 because of data collection problems. Data from a total of fifty-four subjects
232 were analyzed (30 females, ages 19-46 years, median 25 years).

Subject	Behavioral	EEG	fMRI
1	orange		red
2			red
...		orange	red
23			red
24	orange	red	
25		red	
...		red	orange
46		red	
47	orange	orange	orange
48			
...			
54			

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).

233 *2.1.2. Temporal discounting model and task design*

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

$$V_D = \frac{r}{1 + kt} \quad (1)$$

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

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

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

267 where m accounts for sensitivity to changes in discounted value.

268 We used individually determined values of k and m to generate choices
269 for the second session. At every trial, t was randomly selected from a range
270 of 30-45 days. We then calculated and offered an amount r that would give
271 P_D of 0.1, 0.3, 0.5, 0.7, or 0.9 (Figure 3a-b). The EEG group completed 30
272 trials at every P_D level, except at $P_D=0.5$, for which they completed 60 trials.
273 The fMRI group completed 40 trials at every P_D level, except at $P_D=0.5$, for
274 which they completed 80 trials. Non-uniform trial distributions as a function
275 of P_D were introduced to allow us to study the effects of choice difficulty on
276 EEG and fMRI measures, with equal numbers of trials at each difficulty level.
277 The results of these analyses have been reported elsewhere (Rodriguez et al.,
278 2014, 2015b,a). Trial types were randomized and counterbalanced over two
279 blocks for the EEG group and over four blocks for the fMRI group. We also
280 counterbalanced the mapping between choices and button presses for every
281 subject. During the first half of the second session, approximately half of
282 subjects (17 in EEG, 15 in fMRI) indicated choices of the delayed reward
283 by pressing a button with their left index finger and immediate choices by
284 pressing a different button with their right index finger. The other sub-
285 jects indicated their choices by the inverse left-right mapping. All subjects
286 switched the initial response mapping during the second half of the session.

287 We used a sequential presentation of delay and amount during the sec-
288 ond session (Figure 3c). The rationale for this sequential presentation has
289 been explained elsewhere (Rodriguez et al., 2014, 2015b). We report RT as
290 measured from the onset of the decision period, 1000ms into the trial. The
291 duration of the decision period was fixed at 4000ms. When subjects made
292 choices in less than 4000ms the amount information disappeared and the
293 screen remained blank until 4000ms elapsed. Trial length was thus fixed at
294 5000ms. We discarded any trial in which a response was made in less than
295 200ms or fell outside of the decision period. To optimize experimental time
296 and separability of neural signals across trials for both groups, we introduced
297 a long inter-trial-interval (ITI) for the fMRI sessions (between 4-10s), and a
298 shorter ITI for the EEG sessions (100-350ms). In exchange for participation
299 subjects received \$10 cash for the EEG session or \$20 for the fMRI session,
300 plus an additional amount, determined by their choice in a randomly selected
301 trial.

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

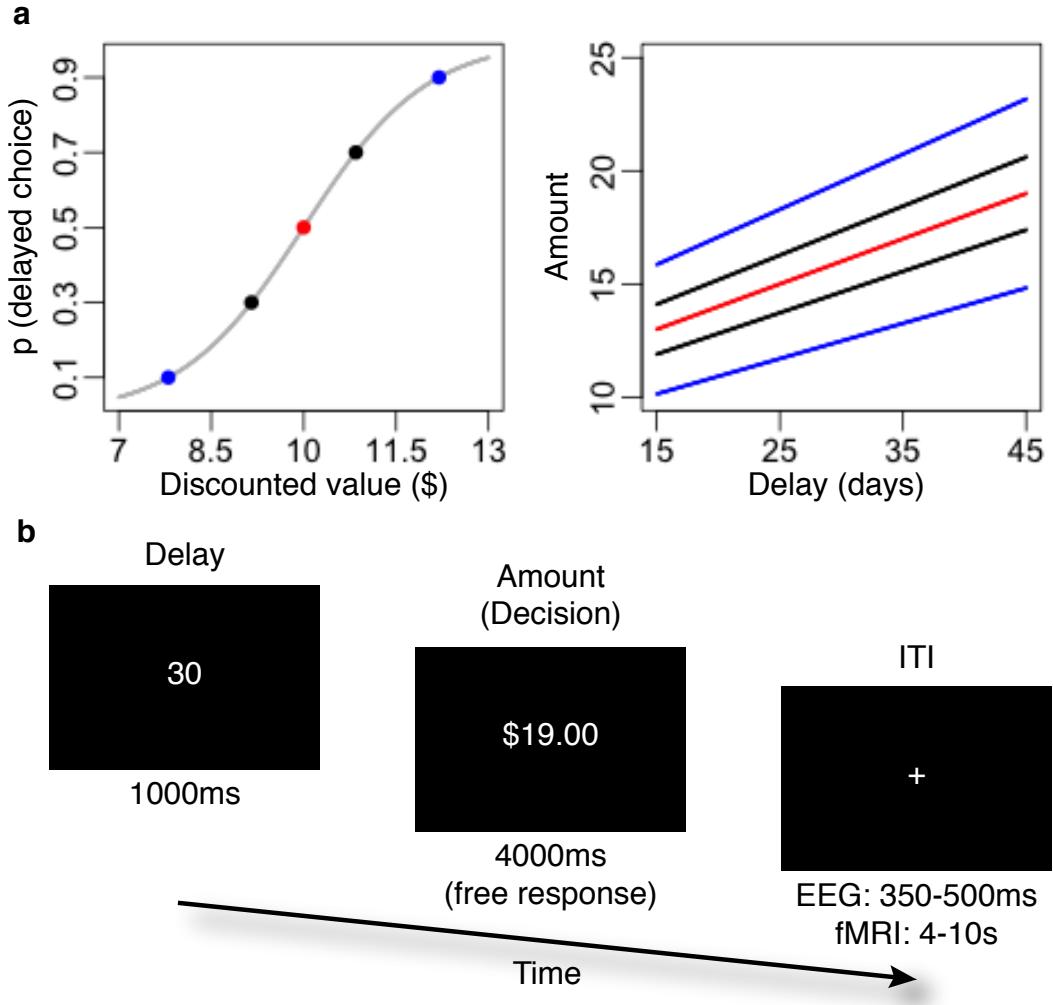


Figure 3: Experimental design: (a) Delayed reward offers corresponded with one of five different levels of discounted value. Each level of discounted value corresponded to one of five probabilities of choosing the delayed reward: 0.1, 0.3, 0.5, 0.7 or 0.9. (b) Every delay could be combined with any of 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.

305 average reference, epoched trials from -1500 to +6500ms around the onset
306 of delay presentation, baseline corrected trials using the average from 0 to
307 5000ms, and band-pass filtered the data at 0.5-200Hz. Trials were visually
308 inspected and rejected if excessive artifacts were present. Normally occurring
309 artifacts were rejected using an independent component analysis algorithm
310 from the EEGLab toolbox (Delorme and Makeig, 2004). Epochs were then
311 transformed to current source density (CSD) using the CSD toolbox (Kayser
312 and Tenke, 2006).

313 fMRI data was collected using a GE Discovery MR750 Scanner. fMRI
314 analyses were conducted on gradient echo T2*-weighted echoplanar func-
315 tional images with BOLD-sensitive contrast (42 transverse slices; TR, 2000
316 ms; TE, 30 ms; 2.9mm isotropic voxels). Slices had no gap between them
317 and were acquired in interleaved order. The slice plane was manually aligned
318 to the anterior-posterior commissure line. The total number of volumes col-
319 lected per subject varied depending on random inter-trial intervals. The first
320 8 seconds (4 volumes) of data contained no stimuli and were discarded to
321 allow for T1 equilibration. In addition to functional data, we collected whole-
322 brain, high-resolution T1-weighted anatomical structural scans (0.9mm isotropic
323 voxels). Image analyses were performed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>).
324 During pre-processing, we first performed slice-timing correction, and re-
325 aligned functional volumes to the first volume. Next, we co-registered the
326 anatomical volume to the realigned functional scans, and performed a seg-
327 mentation of grey and white matter on the anatomical scan. Segmented im-
328 ages were then used to estimate non-linear Montreal Neurological Institute
329 (MNI) normalization parameters for each subject's brain. Normalization pa-
330 rameters estimated from segmented images were used to normalize functional
331 images into MNI space. Finally, normalized functional images were smoothed
332 using a Gaussian kernel of 8 mm full-width at half-maximum.

333 *2.2. Description of the Data*

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

338 *2.2.1. Behavioral Data*

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

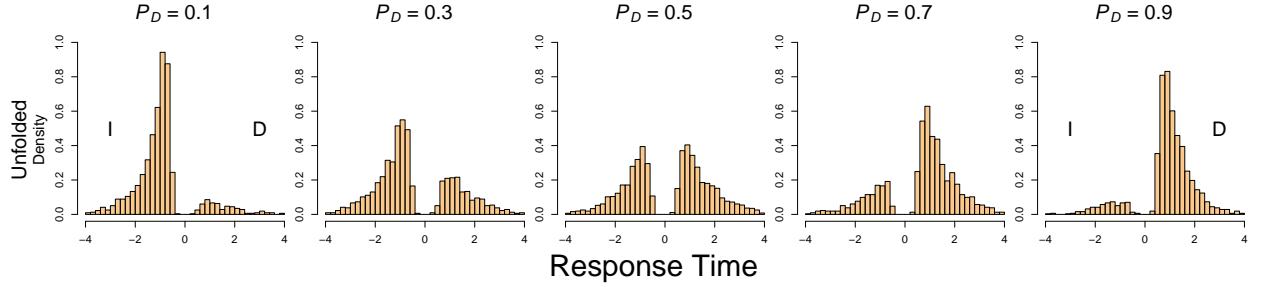


Figure 4: Summary of behavioral data from the experiment. Moving from left to right, each histogram represents the choice response time distribution corresponding to the five levels of probability of delayed reward P_D arranged in increasing order. Response times corresponding to the immediate reward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

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

353 Given the linear relationship to P_D and symmetric properties of the neu-
 354 rophysiological measures reported below, it was convenient to fold our data
 355 symmetrically around $P_D = 0.5$. The folding process created three new
 356 experimental variables: $V1$ consisting of data from $P_D = \{0.1, 0.9\}$, $V2$ con-
 357 sisting of data from $P_D = \{0.3, 0.7\}$, and $V3$ which was equivalent to data
 358 from the previous $P_D = 0.5$ condition. To elaborate, we recoded immediate-
 359 option decisions in the $P_D = 0.1$ condition and delayed-option decisions in the
 360 $P_D = 0.9$ condition to a variable called “subjectively high valued” (SHV),
 361 and delayed-option decisions in the $P_D = 0.1$ condition and immediate-option
 362 decisions in the $P_D = 0.9$ condition to a variable called “subjectively low val-

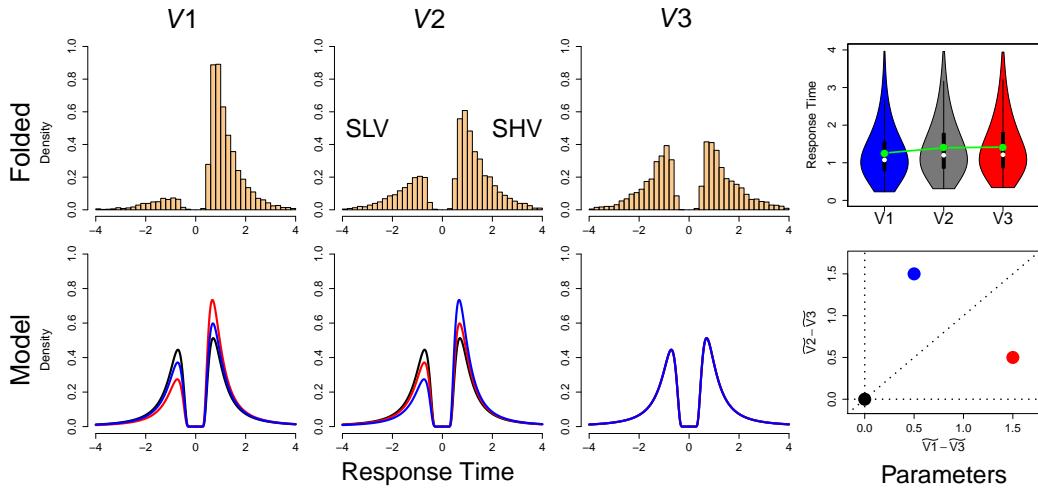


Figure 5: Summary of behavioral data from the experiment. The top row shows the choice response time distributions for the three new conditions $V1$ (first column), $V2$ (second column), and $V3$ (third column) after the folding process (see main text). The response times corresponding to the subjectively 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.

363 ued” (*SLV*). The same mapping was applied to the $P_D = 0.3$ and $P_D = 0.7$
364 conditions, but no mapping was necessary for the $P_D = 0.5$ condition. The
365 top row of Figure 5 shows the response time distributions for the variables
366 $V1$ (first column), $V2$ (second column), and $V3$ (third column), where *SLV*
367 choices are shown on the left and *SHV* choices are shown on the right. The
368 figure shows that as the levels of V increase, preference of the *SHV* option
369 decreases until indifference in condition $V3$. In addition to a relationship
370 between response probability and condition, the top-right panel of Figure
371 5 shows a violin plot of the response time distributions for each condition,
372 collapsed across response probability. The white dots on each violin plot
373 represent the median response time, whereas the green dot represents the
374 mean response time. The figure shows that the response times increase from
375 $V1$ to $V2$, and from $V2$ to $V3$, but to a lesser degree.

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

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

401 V_3 to V_2 , and another parameter $\eta^{(2)}$ additively shifted the model’s drift
 402 rate between Condition V_3 to V_1 . Although Appendix A contains all of the
 403 mathematical details of our implementation, the basic equations of our drift
 404 rate parameterization for the subjectively high valued option are

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

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

410 To illustrate how the LBA model accounts for the folded data, we simu-
 411 lated the model – described in Appendix A – for three different settings of the
 412 shift parameters for each of the three value conditions, shown in the bottom
 413 row of Figure 5. The settings of the shift parameters are represented in the
 414 bottom-right panel of Figure 5, where the x -axis corresponds to $\eta^{(2)}$ (i.e., the
 415 shift from V_3 to V_1), and the y -axis corresponds to the $\eta^{(1)}$ (i.e., the shift
 416 from V_3 to V_2). The first parameter setting – illustrated by the black dot
 417 in the bottom right panel of Figure 5 – was $(0, 0)$. This setting produces no
 418 change in drift rate across any of the value conditions. The first, second, and
 419 third columns show the predicted choice response time distributions from
 420 the LBA model (black lines) for the value conditions V_1 , V_2 , and V_3 , re-
 421 spectively. Here, because no change in drift rate occurred, the model makes
 422 equivalent predictions across value conditions. Other parameter settings are
 423 more interesting. For example, the red dot in the bottom-right panel of Fig-
 424 ure 5 shows the model predictions under the setting $(\eta^{(2)}, \eta^{(1)}) = (1.5, 0.5)$.
 425 Moving from V_3 (third column) to V_2 (second column), we see a marked
 426 increase in the probability of endorsing the *SHV* option. This change is
 427 produced by the 0.5 unit increase (i.e., on the logit scale) in the drift rate for
 428 the *SHV* option. Moving from V_3 (third column) to V_1 (first column), we
 429 see an even larger increase in the probability of endorsing the *SHV* option.
 430 Again, this increase follows directly from the value of $\eta^{(2)} = 1.5$. The third
 431 parameter setting produces response time distributions that are similar to
 432 the second parameter setting, but the magnitude of the effect is in the op-
 433 posite 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 V_1 , V_2 , and V_3 , 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 β s 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 the subjects' response in every trial,

471 plus the onset of the amount presentation for a single trial. Events in all three
472 onset regressors were modeled as impulse delta functions and convolved with
473 the canonical hemodynamic response function (HRF). In addition, the model
474 included six regressors corresponding to the motion parameters estimated
475 during data preprocessing and constants to account for the mean activity
476 within each of the four sessions over which the data were collected. Figure
477 6g shows the average of the single-trial β s in V1 (blue), V2 (black), and V3
478 (red), and this average activity is also illustrated in the brain on the same
479 axial slice given in Figure 6f for V1 (h), V2 (i), and V3 (j). Collectively,
480 Figures g-j show that as the levels of V increase, activation of the dmFC
481 increases, an effect that is opposite in direction to our EEG data. In our
482 analyses below, we will use the single-trial β estimates in the dmFC as input
483 to our joint behavioral-neural model.

484 **3. A Bivariate Joint Model of Value Accumulation**

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

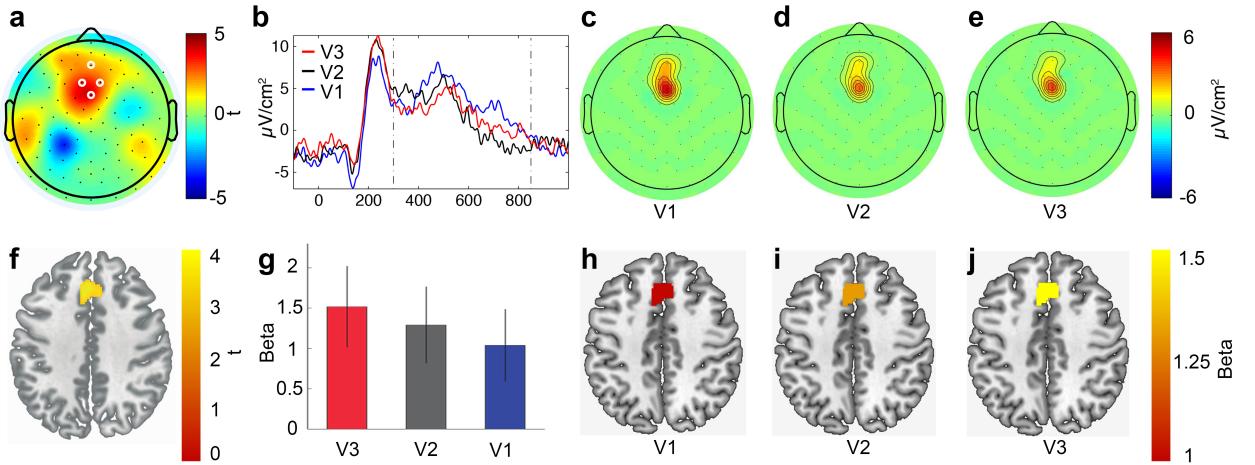


Figure 6: Summary of neural data from the experiment. (a) Topographic plot of the test value statistics relating EEG activity to the levels of P_D in a GLM analysis. The electrodes with the strongest relationship are represented with white circles. (b) Mean stimulus-locked event-related potential (ERP) 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.

504 *3.1. Details of the Model*

505 The joint modeling approach is particularly well-suited for situations in
506 which the measures of interest can be described independently of one an-
507 other, and the research goal is to infer and constrain the model parameters
508 describing each of these measures (see Turner et al., 2015a, for a survey of
509 different statistical strategies for linking neural and behavioral measures).
510 These independent “submodels” are then linked together, comprising an en-
511 tirely new model with different – albeit similar – properties of its constituent
512 parts. Furthermore, the linking process inherently changes the submodels,
513 due to the added constraint provided by the additional measures. That is,
514 data from one submodel affects the parameter estimates of the other sub-
515 model, and vice versa. Perhaps more interesting is that the full model is fit
516 to the full data set in one setting, so elaborate retraining techniques are not
517 necessary in generating flexible predictions (e.g., for missing data). The bi-
518 variate joint model can be described in three parts: (1) the neural submodel
519 which succinctly describes the important statistical properties of the neural
520 data, (2) the behavioral submodel consisting of latent model parameters as-
521 sumed to produce an observer’s set of choice response time data, and (3)
522 a linking structure that binds the two submodels together. We will discuss
523 each of these pieces of the model in turn.

524 *3.2. The Neural Submodel*

525 The first piece of the bivariate joint model is the neural submodel describ-
526 ing the neural data. The choice of a neural submodel is an important one,
527 but in our experience, the most important factor in this choice is reducing
528 the neural data to a (smaller) set of neural model parameters. The neu-
529 ral submodel can be generative, expressing how the neural data arise from
530 a theoretically-oriented process (e.g., Manning et al., 2014; Friston et al.,
531 2003; Gershman et al., 2011), or statistical, describing the statistical prop-
532 erties of the neural data without theoretical considerations (e.g., Norman
533 et al., 2006). One effective strategy within the joint modeling framework for
534 functional data is to average the trial-to-trial neural signal across an impor-
535 tant time window in the decision process (e.g., Turner et al., 2015b). This
536 procedure allows one to easily justify choices of the linking function via the
537 central limit theorem. Specifically, averaging the neural data produces a ran-
538 dom variable that is Gaussian in distribution, making multivariate Gaussian
539 linking functions both justified and convenient (see Appendix B for details).

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

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

572 As described in Appendix B, the bivariate joint model uses a parame-
573 ter to capture the mean neural (i.e., either EEG or fMRI) signal in the V3
574 condition, where the choice options have approximately the same subjective
575 value, as well as additive parameters that capture the change in the mean
576 neural signal from value condition V3 to V2, and from V3 to V1. At the
577 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-

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

628 *3.3. Application To EEG Data*

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

635 *3.3.1. Results*

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

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

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

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

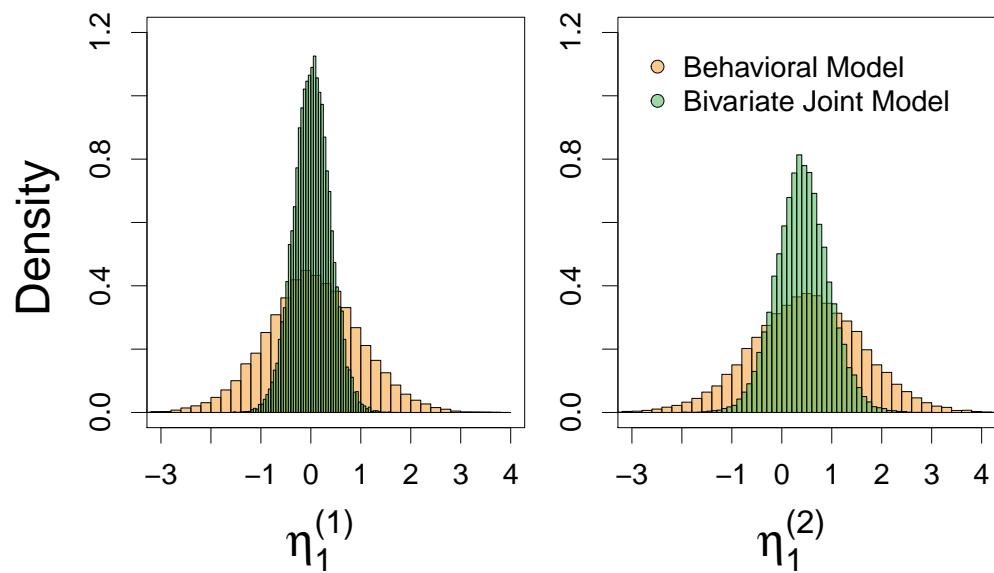


Figure 7: Comparison of the value effect parameters $\eta_1^{(1)}$ (left panel), and $\eta_1^{(2)}$ (right panel) obtained from the behavioral-data-only model (orange) and the bivariate joint model (green) for the withheld subject in the EEG experiment.

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

704 Figure 7 shows that for both models, the prediction for $\eta_1^{(1)}$ is centered
705 on zero, whereas for $\eta_1^{(2)}$, the posteriors are shifted slightly upward. This
706 indicates that both models were capable of using reasonably sparse data (i.e.,
707 just 22 subjects) to generalize predictions for a subject whose data were not
708 observed in a way that is reflective of our previous behavioral analyses on a
709 full data set (consisting of 46 subjects; Rodriguez et al., 2014). However, the
710 important question was whether the bivariate joint model was more sensitive
711 to the effects of value condition due to its ability to exploit the information
712 in the neural signal. To answer this, we only need to compare the orange
713 and green posterior predictions in both panels. Visually, it is clear that
714 the green histograms have less variance, and the locations of the green and
715 orange histograms are approximately equal. For $\eta_1^{(1)}$, the means are 0.393
716 and 0.437 and the standard deviations are 1.005 and 0.500 for the behavioral-
717 data-only and the bivariate joint model, respectively. For $\eta_1^{(2)}$, the means are
718 1.068 and 0.946 and the standard deviations are 1.147 and 0.628 for the
719 behavioral-data-only and the bivariate joint model, respectively. We can

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

720 also evaluate the models' predictions for the probability that $\eta_1^{(1)}$ and $\eta_1^{(2)}$
721 are greater than zero. The behavioral-data-only model predicts that $p(\eta_1^{(1)} >$
722 $0) = 0.657$ and $p(\eta_1^{(2)} > 0) = 0.828$, whereas $p(\eta_1^{(1)} > 0) = 0.819$ and $p(\eta_1^{(2)} >$
723 $0) = 0.937$ for the bivariate joint model. This suggests that the relative
724 generalizability to patterns across conditions is greater in the bivariate joint
725 model, although it is worth emphasizing that this generalization may not
726 result in better predictions for behavioral data. To address that question, we
727 will use predictive modeling techniques in the next section.

728 Another generative modeling analysis would be to inspect the correlation
729 parameters between the neural and behavioral submodels. In the bivariate
730 models, two correlation parameters exist that describe the relationship be-
731 tween the additive shift parameters in the neural and behavioral submodels.
732 Specifically, the parameter ρ_1 describes, across subjects, the degree to which
733 $\eta_j^{(1)}$ is related to $\delta_j^{(1)}$ for all j , and the parameter ρ_2 describes, across sub-
734 jects, the degree to which $\eta_j^{(2)}$ is related to $\delta_j^{(2)}$. Hence, the parameters ρ_1 and
735 ρ_2 describe how much the effects present in the neural and behavioral data
736 from conditions $V2$ to $V3$ and from $V1$ to $V3$, respectively. Figure 8 shows
737 the estimated posterior distributions for the hyper correlation parameters ρ_1
738 (left panel), and ρ_2 (right panel). Figure 8 shows that ρ_1 is centered near
739 zero, whereas ρ_2 is centered on positive values. Hence, the posterior estimate
740 of ρ_2 suggests that the additive effects that give rise to higher drift rates
741 in condition $V1$ are correlated with the additive effects of the EEG signal.
742 Hence, across subjects, increases in the EEG signal in condition $V1$ relative
743 to condition $V3$ are predictive of higher drift rates in condition $V3$ (see also
744 Figure 7).

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

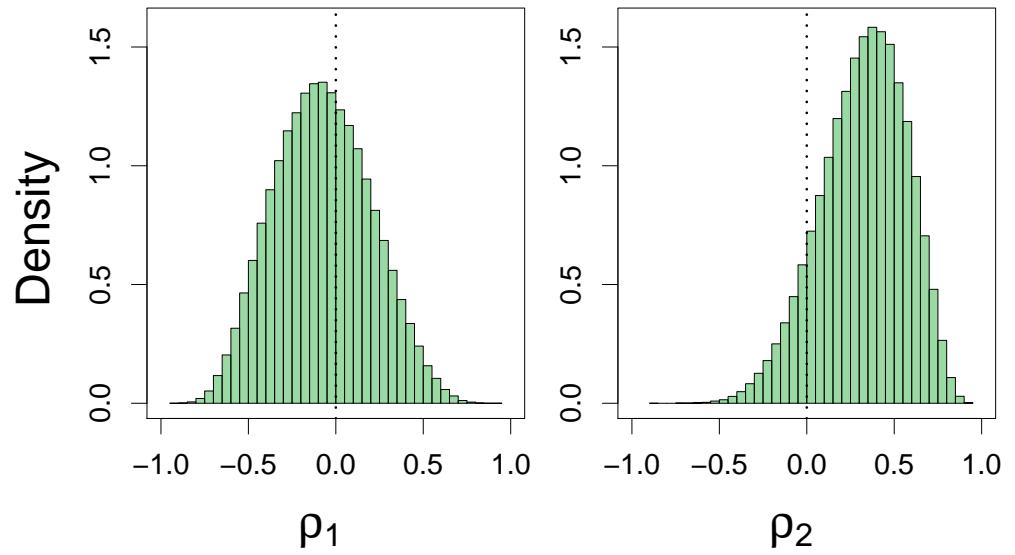


Figure 8: Estimated posterior distributions for the hyper correlation parameters relating the additive shift parameters in the neural and behavioral submodels across the $V3$ and $V2$ conditions (left panel), and the $V3$ and $V1$ conditions (right panel) for the EEG data. In each panel, a reference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

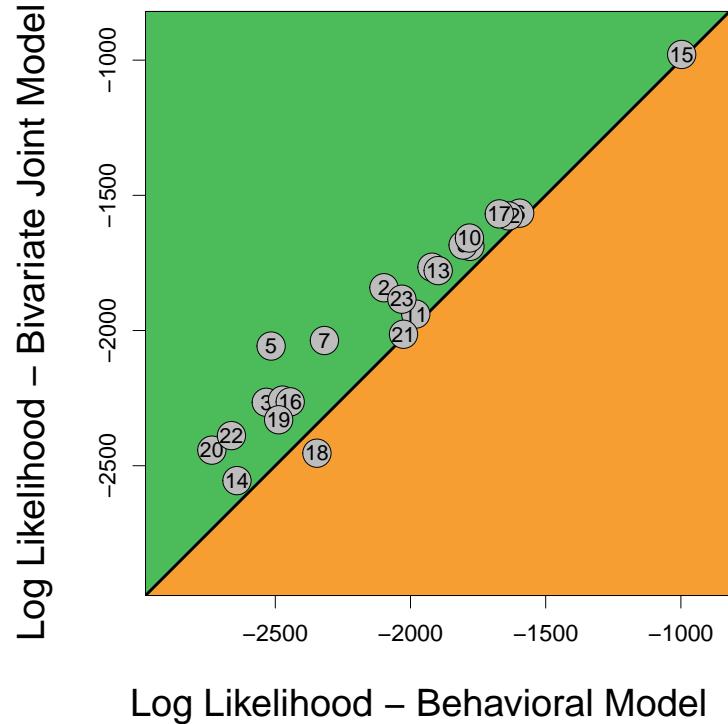


Figure 9: Predictive performance of the joint and behavioral models on the EEG dataset. Each node in the graph represents the models' performance for the subject indicated, where the performance for the joint model is represented on the y -axis and the performance for the behavioral-data-only model is 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.

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

797 *3.4.1. Results*

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

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

814 Figure 10 shows the resulting posterior distributions for Subject 24's main
815 behavioral effect parameters: $\eta_{24}^{(1)}$ (left panel) and $\eta_{24}^{(2)}$ (right panel). In both
816 panels, predictions from the bivariate joint model are shown as blue his-
817 tograms, whereas predictions from the behavioral-data-only model are shown
818 as orange histograms. Figure 10 shows that while the parameter $\eta_{24}^{(1)}$ is cen-
819 tered on zero, the parameter $\eta_{24}^{(2)}$ has a slightly positive effect, suggesting that
820 as the subjective values of the two alternatives become more dissimilar, the
821 drift rate for the subjectively higher valued alternative increases. Compar-
822 ing across models, we see that visually both models produce posteriors with
823 similar expected values, the posteriors are again substantially different. For
824 $\eta_{24}^{(1)}$, the means are 0.340 and 0.329 and the standard deviations are 0.999 and
825 0.512 for the behavioral-data-only and the bivariate joint model, respectively.
826 For $\eta_{24}^{(2)}$, the means are 1.132 and 1.277 and the standard deviations are 1.166
827 and 0.693 for the behavioral-data-only and the bivariate joint model, respec-
828 tively. We can also quantitatively evaluate the models' generalizations for the

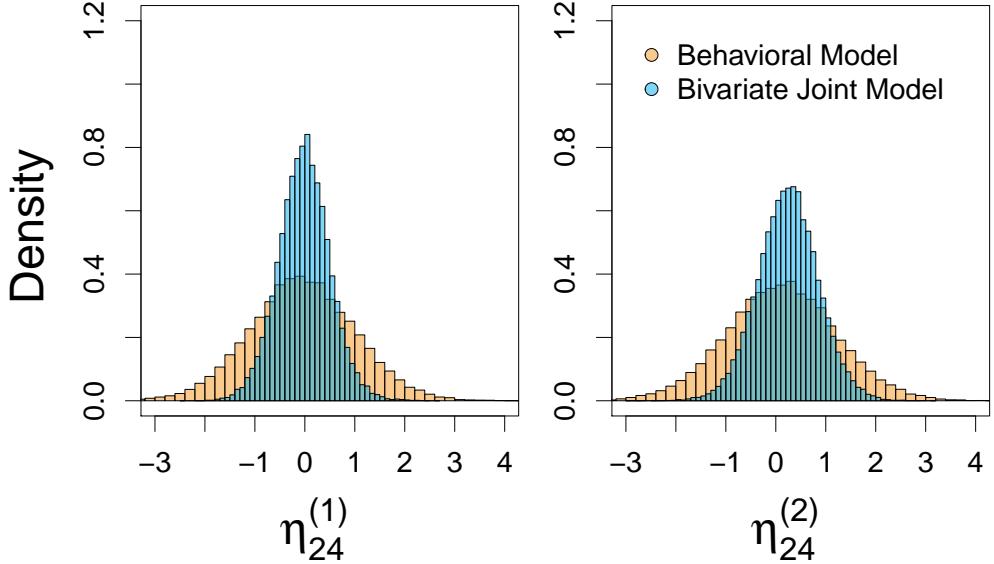


Figure 10: Comparison of the value effect parameters $\eta_{24}^{(1)}$ (left panel), and $\eta_{24}^{(2)}$ (right panel) obtained from the behavioral-data-only model (orange) and the bivariate joint model (blue) for the withheld subject in the fMRI experiment.

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

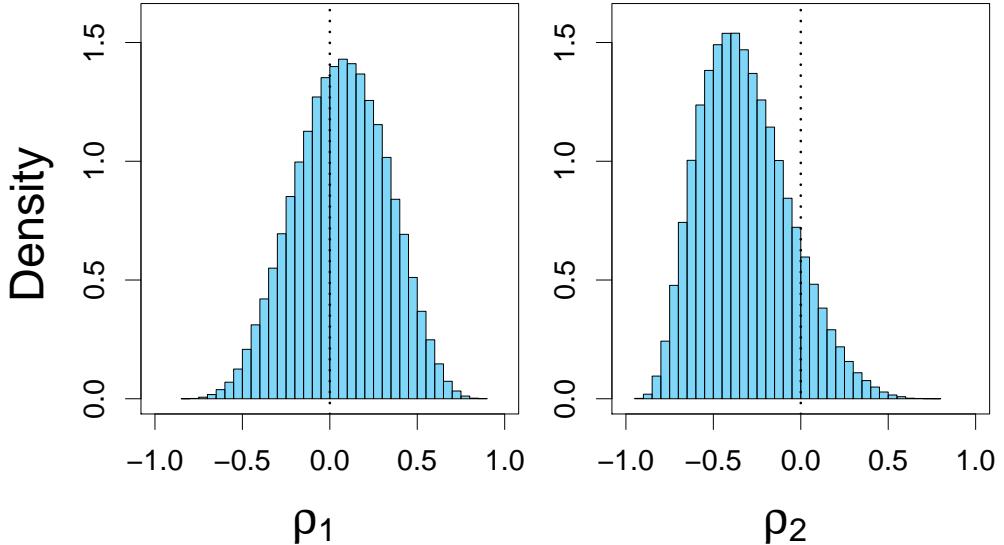


Figure 11: Estimated posterior distributions for the hyper correlation parameters relating the additive shift parameters in the neural and behavioral submodels across the V3 and V2 conditions (left panel), and the V3 and V1 conditions for the fMRI data. In each panel, a reference line is plotted to indicate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

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

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

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

884 4. A Trivariate Joint Model of Value Accumulation

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

891 4.1. Results

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

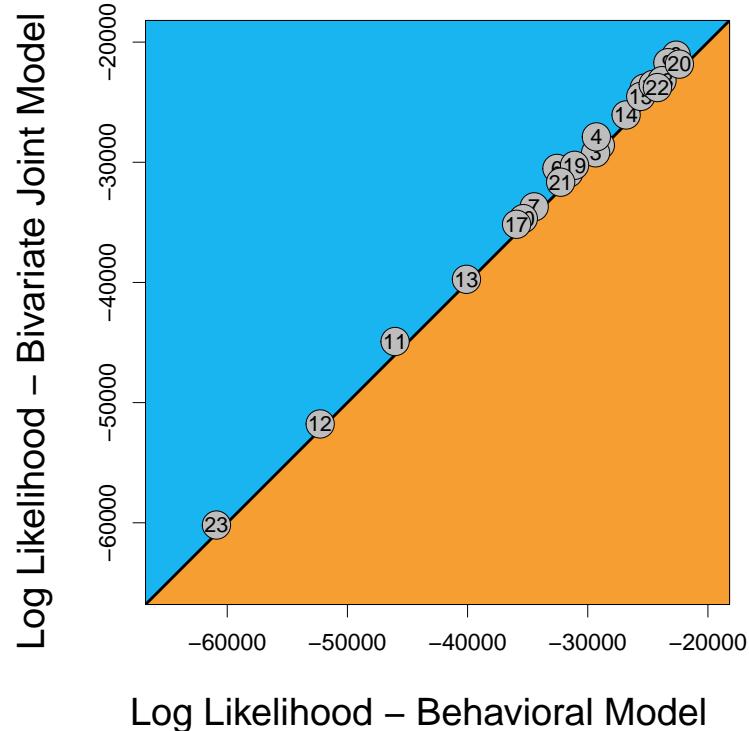


Figure 12: Predictive performance of the joint and behavioral models on the fMRI dataset. Each node in the graph represents the models' performance for the subject indicated, where the performance for the bivariate joint model is represented as the y -axis and the performance for the behavioral model 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.

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

897 *4.1.1. Generative Model Analysis*

898 Our generative analysis was similar to the bivariate applications above.
899 In our first analysis, we focused on the effects of the neural data on the shift
900 parameters $\eta^{(1)}$ and $\eta^{(2)}$ in the behavioral submodel. Unlike our bivariate
901 applications above, we now have two neural measures we can consider adding
902 to the model, creating four different models to test. The first model only
903 considers the behavioral data, the second model considers only the behavioral
904 and EEG data, the third model considers only the behavioral and fMRI data,
905 and the fourth model considers the behavioral, EEG, and fMRI data, as in
906 Figure 1. For illustration purposes, we again withheld the behavioral data for
907 one subject – Subject 53 – to compare and contrast each model’s predictions.³
908 In the predictive modeling analysis section below, we will use an extension
909 of this technique in a cross-validation test. Although we withheld Subject
910 53’s behavioral data, the EEG and fMRI data were left intact. Because some
911 of these models are unable to use all components of the neural data, each
912 model’s ability to generalize to a new subject’s behavioral data depends on
913 the reliability of the usable neural data.

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

³Unlike 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.

920 positive values, with $\eta_{53}^{(2)}$ being larger than $\eta_{53}^{(1)}$. These results are consistent
921 with the previous bivariate joint model analyses above. Comparing across
922 models, we see that while all models produce posteriors with similar expected
923 values, the dispersion of those posteriors differs substantially. Namely, the
924 behavioral-data-only model has the widest variance, the bivariate EEG and
925 fMRI models have similar and slightly less variance, and the trivariate model
926 has the smallest variance. We will evaluate whether this pattern in vari-
927 ances across models leads to better predictions of the behavioral data in the
928 predictive modeling analysis section below.

929 We can also quantitatively evaluate the models' generalizations for the
930 probability that $\eta_{53}^{(1)}$ and $\eta_{53}^{(2)}$ are greater than zero. The behavioral-data-only
931 model predicts that $p(\eta_{53}^{(1)} > 0) = 0.666$ and $p(\eta_{53}^{(2)} > 0) = 0.874$. The bivari-
932 ate EEG model predicts that $p(\eta_{53}^{(1)} > 0) = 0.766$ and $p(\eta_{53}^{(2)} > 0) = 0.950$.
933 The bivariate fMRI model predicts that $p(\eta_{53}^{(1)} > 0) = 0.796$ and $p(\eta_{53}^{(2)} >$
934 $0) = 0.929$. Finally, the trivariate model predicts that $p(\eta_{53}^{(1)} > 0) = 0.839$
935 and $p(\eta_{53}^{(2)} > 0) = 0.970$. The basic pattern is consistent with the pattern
936 of variance above: the models whose posteriors have less variance predict
937 greater probabilities that the η parameter is greater than zero. Furthermore,
938 the behavioral-data-only and the bivariate models predict greater probabili-
939 ties that η is greater than zero compared to the bivariate analyses above,
940 suggesting that the additional data is consistent with the subsets considered
941 in the above sections.

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

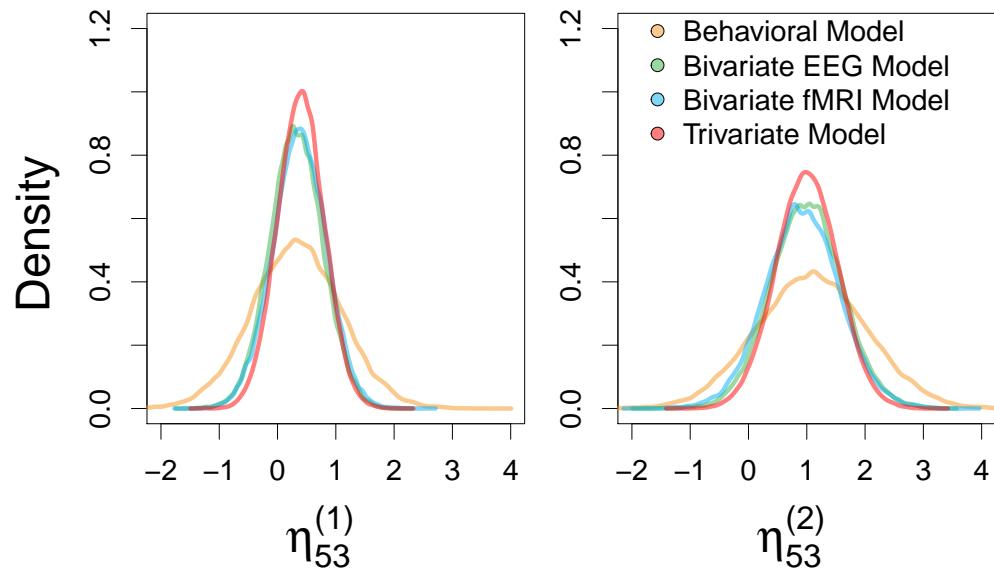


Figure 13: Comparison of the value effect parameters $\eta_{53}^{(1)}$ (left panel), and $\eta_{53}^{(2)}$ (right panel) obtained from the behavioral-data-only model (orange), the bivariate EEG model (green), the bivariate fMRI model (blue), and the trivariate model (red) for the withheld subject (i.e.,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

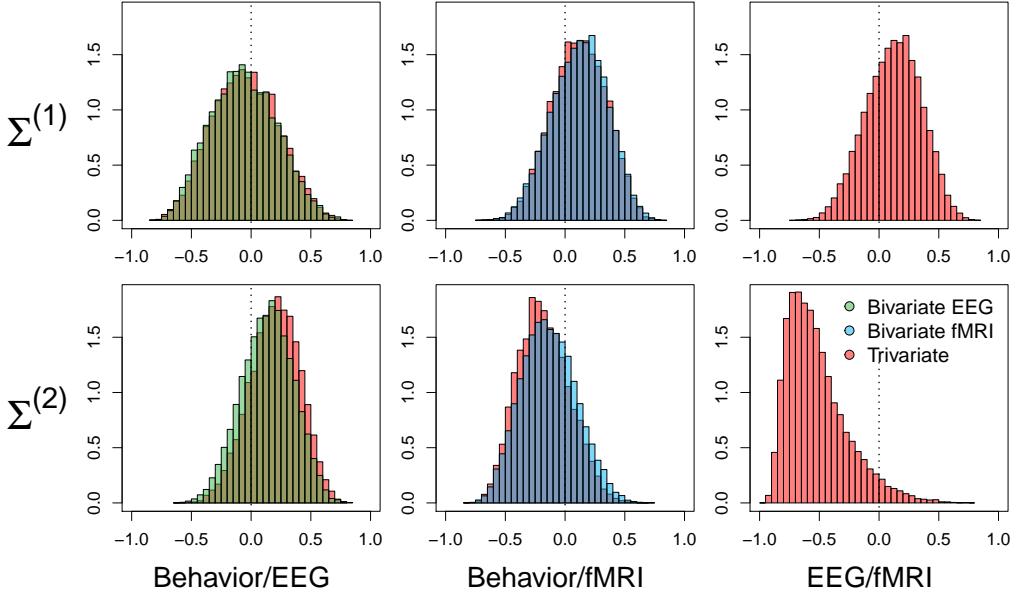


Figure 14: Estimated posterior distributions for the hyper correlation parameters. The top row shows the correlations between the additive shift parameters across the V3 and V2 conditions, whereas the bottom row shows the correlations between the shift parameters across V3 and V1. The first column shows 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.

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

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

968 parameters for each of the three joint models are illustrated with histograms
969 of different colors: bivariate EEG as green, bivariate fMRI as blue, and the
970 trivariate model as red. Figure 14 shows that components of the bivariate
971 models do closely align with the corresponding components of the trivariate
972 model; however, the variance of the estimate posterior distribution for
973 the trivariate model is smaller, especially for the $\Sigma^{(2)}$ matrix. Akin to the
974 other comparisons of behavioral-data-only and joint models above, this ad-
975 ditional constraint comes from having additional measures in the model that
976 are correlated. The right column of Figure 14 shows that the additive shift
977 parameters in the EEG and fMRI data are uncorrelated across Conditions
978 V_2 and V_3 , but strongly negatively correlated across Conditions V_1 and
979 V_3 . Although this relationship is something that we might have logically de-
980 duced from two separate bivariate model analyses, only the trivariate model
981 is able to use this information to provide statistical constraint in generating
982 predictions for new data.

983 *4.1.2. Predictive Modeling Analysis*

984 We again performed a leave-one-out cross-validation test to assess the
985 models' ability to generate predictions for behavioral data. To accomplish
986 this, we focused on the eight subjects who provided both EEG and fMRI
987 measures, so that the bivariate and trivariate models could be directly com-
988 pared. First, we removed all data for one subject (e.g., Subject 47). Second,
989 we fit all four models to the remaining 53 subjects' data. Third, we generate
990 predictions for the withheld subject's behavioral data, conditional on some
991 aspect of the neural data. For example, the bivariate EEG model only takes
992 EEG data as an input, whereas the trivariate model can take both EEG and
993 fMRI data as inputs. Finally, we examined the models' predictive accuracy
994 by calculating the (log) likelihood of the withheld subject's data from the
995 estimated posterior distributions via Monte Carlo procedure (Robert and
996 Casella, 2004). Specifically, we integrated over the posterior distributions to
997 numerically estimate the probability that any given observation for a partic-
998 ular subject was best predicted by a particular model. In the end, we arrived
999 at an estimate of the probability that a particular subject's data was best
1000 captured by each model, a value that summed to one across models. Figure
1001 15 shows a ternary plot of these probabilities for the eight subjects in our
1002 data. In the figure, we removed the behavioral-data-only model from consid-
1003 eration because (1) we have already established the bivariate model's ability
1004 to out predict the behavioral-data-only model, and (2) in this analysis, it

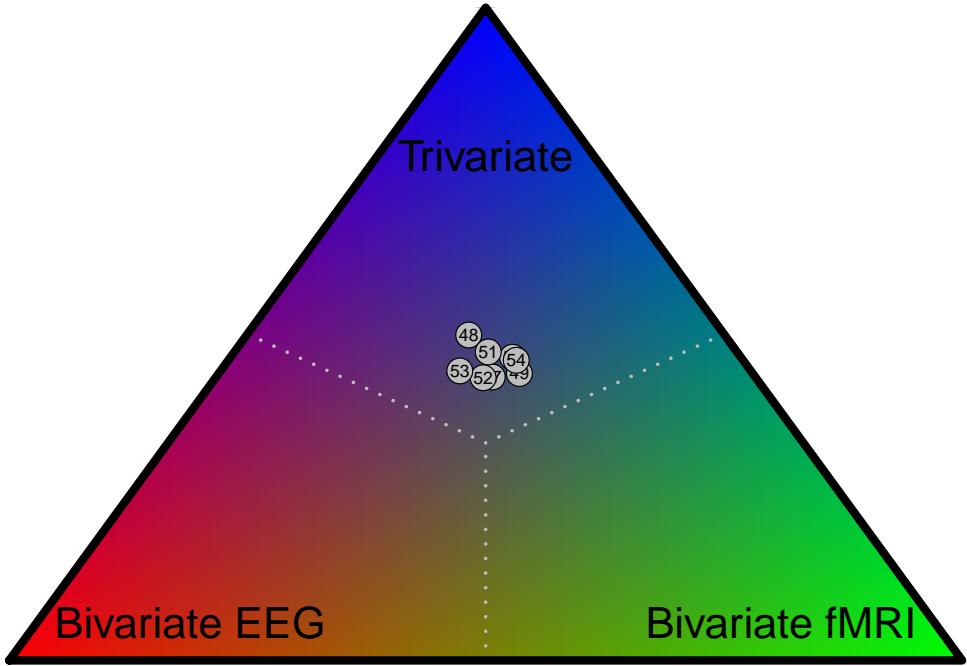


Figure 15: Results of the cross-validation test on the behavioral data from the last eight subjects of the experiment. The three vertices of the ternary plot each designate areas favoring one type of joint model: bivariate EEG (red; bottom left), bivariate fMRI (green; bottom right), and the trivariate 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.

1005 performed the worse of the four models.

1006 The reference lines in Figure 15 designate the areas best predicted by a
 1007 particular model. In addition, the color values illustrate the graded nature of
 1008 the model selection process. The trivariate model is represented as blue, the
 1009 bivariate EEG model is represented as red, and the bivariate fMRI model is
 1010 represented as green. The nodes in the graph mark the probability of a given
 1011 model, conditional on the behavioral data. The figure shows that for each of
 1012 the subjects in our data, the trivariate model provides the best predictions
 1013 than any of the other three models (although the behavioral-data-only model
 1014 is not shown).

1015 *Generalizations for Neural Data.* Although Figure 15 shows the trivariate
1016 model’s predictive abilities for behavioral data relative to the bivariate mod-
1017 els, there are other predictive tests we could have performed. As a final
1018 test of the utility of our trivariate model, we assessed the model’s ability
1019 to predict neural data on the basis of behavioral data. To do this, we refit
1020 the trivariate model to the full data set, but withheld the neural data for
1021 two randomly selected subjects: Subject 4 from the EEG experiment, and
1022 Subject 33 from the fMRI experiment. This withholding strategy is again
1023 illustrated by the orange blocks in Figure 2. We then generated predictions
1024 for the EEG data of Subject 4, and the fMRI data for Subject 33, on the
1025 basis of the behavioral data.

1026 Figure 16 shows the predictions for the EEG data made for Subject 4.
1027 Panel a shows the PPDs for the difference between value conditions $V1$ and
1028 $V3$ (x -axis) against the difference between value conditions $V2$ and $V3$ (y -
1029 axis). Samples of the joint PPD are represented with shaded blue dots to
1030 show the relative density across the parameter space. Panel a also shows some
1031 summary statistics of Subject 4’s EEG data, reflected in the “+” symbol:
1032 the mean along each dimension is located at the center, and one standard
1033 deviation of the PPD along each dimension is represented as the length of
1034 the line (i.e., the horizontal line corresponds to the x -axis, and the vertical
1035 line corresponds to the y -axis). For comparison, the mean of the PPD is
1036 represented as the open circle. Plotting the PPD in this way allows us to
1037 interpret the relative differences in activation predicted by the model between
1038 the two value conditions. From our previous analyses (also see Figure 6), we
1039 might expect the model to predict that raw activation in the EEG signal
1040 would be highest in $V1$, and lowest in $V3$, implying that the differences
1041 between $V1$ and $V3$ would be positive and larger than the differences between
1042 $V2$ and $V3$. Given this intuition, we would expect the largest density of the
1043 PPD to be on the gray triangle in Panel a. The PPD, having been centered
1044 relative to Condition $V3$, places more density on positive values in both the
1045 x and y directions. Further, in the upper right quadrant of Panel a, the PPD
1046 has greater density in the half-quadrant representing values where $V1 - V3$ is
1047 higher than $V2 - V3$. Importantly, the model’s predictions are also consistent
1048 with Subject 4’s data in that both the filled circle and the “+” symbol fall
1049 in the gray area in Panel a.

1050 Panel b shows how these predictions appear on the scalp. The first col-
1051 umn shows the mean of the PPD (represented as the filled circle in Panel a)
1052 in the $V2 - V3$ (top) and $V1 - V3$ (bottom) conditions. The second col-

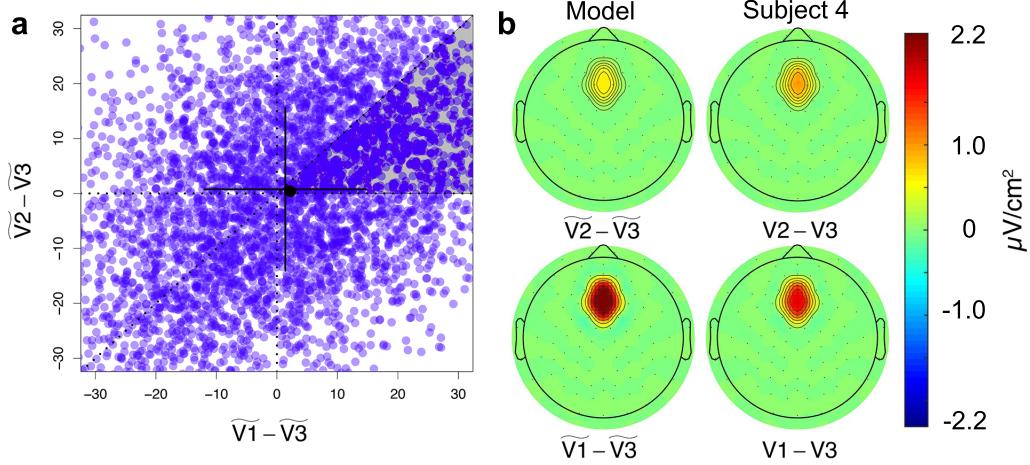


Figure 16: Posterior predictive distributions (PPDs) for the withheld EEG data of Subject 4. Panel a shows the full PPD for the difference between value conditions $V1$ and $V3$ (x -axis) against the difference between value conditions $V2$ and $V3$ (y -axis). The “+” symbol illustrates summary statistics 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.

1053 umn shows Subject 4’s withheld data in the same conditions. In each scalp
 1054 topography plot, activation is color coded according to the key on the right
 1055 side. Comparing across panels, we conclude that the model’s predictions are
 1056 entirely consistent with the withheld subject’s data.

1057 A parallel analysis was performed for the withheld fMRI data of Subject
 1058 33. Similar to Figure 16, Figure 17a shows the PPD for the difference between
 1059 value conditions $V1$ and $V3$ (x -axis) against the difference between value
 1060 conditions $V1$ and $V2$ (y -axis). Summary statistics of Subject 33’s fMRI
 1061 data are plotted in Panel a of Figure 17 in the same way they were plotted
 1062 in Panel a of Figure 16. Given our previous analyses, we would expect the
 1063 opposite pattern of predictions from the EEG predictions above. That is,
 1064 we would expect the activation to be highest in Condition $V3$, and lowest in

1065 Condition V_1 , implying that the predictions for $V_1 - V_3$ would be negative,
1066 and more negative than the predictions for $V_2 - V_3$. Hence, we should expect
1067 a high density of the model predictions, and the withheld data to fall in the
1068 gray half-quadrant illustrated in Panel a. Figure 17 confirms this prediction
1069 and shows that the predictions from the trivariate model are indeed consistent
1070 with the withheld data.

1071 Panel b shows the average model predictions from Panel a (left column)
1072 beside Subject 33’s withheld fMRI data. The degree of activation in the
1073 dmFC is color coded according to the key on the far right side. Figure
1074 17 shows that the model predicts that the difference in activation in the
1075 dmFC between value conditions V_2 and V_3 is smaller in magnitude than
1076 the difference in activation between value conditions V_1 and V_3 . These
1077 predictions are in line with the withheld data, which further suggests that
1078 the trivariate model uses the inferred relationships between neural activity
1079 and the behavioral model to generalize accurately to new subjects.

1080 5. Discussion

1081 In this article, we have shown how the joint modeling framework (Turner
1082 et al., 2013a) can be extended to solve the more difficult challenge of inte-
1083 grating multiple neurophysiological measures with a behavioral model. Our
1084 results suggest that bivariate models, having either EEG or fMRI, can out-
1085 perform behavioral-data-only models, and trivariate models, having both
1086 EEG and fMRI, can outperform bivariate models. The reason for this pat-
1087 tern of results is due to (1) the interrelationships among the three variables,
1088 and (2) the trivariate model’s ability to learn the important patterns in the
1089 data to facilitate better generalization. While our results certainly support
1090 our method for multi-variate extensions, there are a number of important
1091 limitations and benefits meriting further discussion.

1092 5.1. Reliance on Preprocessing Methodology

1093 To date, all applications of our joint modeling framework rely on some
1094 preprocessing of the neural signal as a way to reduce the dimensionality of
1095 these measures. For example, in our processing of the EEG data above, we
1096 first computed the average EEG signal for each value condition and visu-
1097 ally inspected these curves across time. We then chose a window of time
1098 where the EEG signal showed the most differentiation between the three

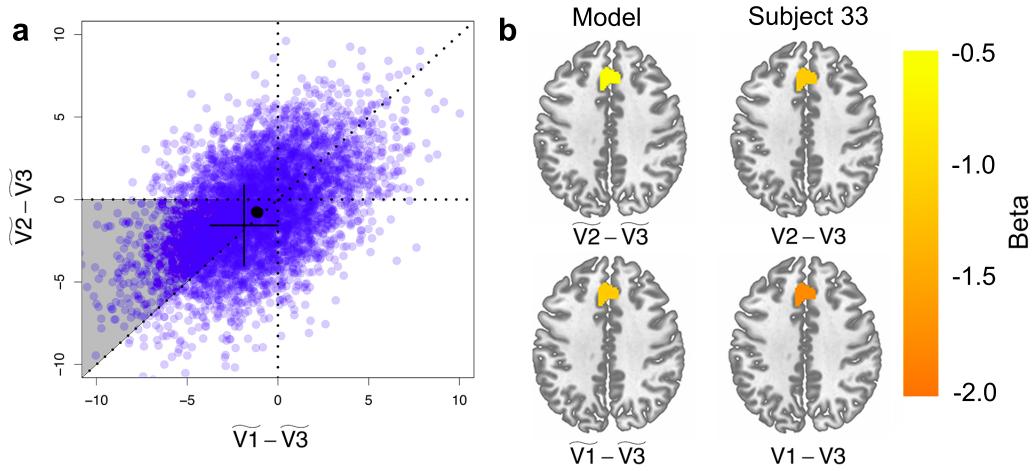


Figure 17: Posterior predictive distributions (PPDs) for the withheld fMRI data of Subject 33. Panel a shows the full PPD for the difference between value conditions V_1 and V_3 (x -axis) against the difference between value conditions V_2 and V_3 (y -axis). The “+” symbol illustrates summary statistics 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.

1099 value conditions. This time window represented the period of time we sus-
 1100 pected would be the most interesting from a cognitive perspective, but this
 1101 procedure leaves much to be desired. First, it averages the wave form over
 1102 time, and second, it averages over only four electrodes out of many. One
 1103 could easily imagine the analyses performed here with improvements to the
 1104 neural submodel such as decomposing the neural signal into its temporal and
 1105 spatial aspects. Future work should endeavor to establish stronger data re-
 1106 duction techniques that are model driven, such as those used by topographic
 1107 factor analysis (Manning et al., 2014; Gershman et al., 2011).

1108 5.2. Generalization Depends on the Value of Conditioned Variables

1109 Our bivariate and trivariate generative modeling analyses showed that
 1110 adding neural measures constrains our knowledge about behavioral model
 1111 parameters (see Figures 7, 10, and 13). In these analyses, constraint was
 1112 observed through the additional precision (i.e., having less variance) in the
 1113 posterior estimates, and importantly, this additional constraint facilitated
 1114 more accurate predictions about withheld behavioral data. The additional
 1115 constraint depends on a few important factors. First, consider the case of
 1116 two arbitrary random variables (θ, δ) that have a joint multivariate normal
 1117 distribution, as in the bivariate joint models presented above, where θ are the
 1118 linked behavioral submodel parameters, and δ are the linked neural submodel
 1119 parameters. When provided with neural data, the neural model parameters
 1120 δ become *estimable*, allowing our model to infer an estimate for δ that comes
 1121 from both the prior and the likelihood. In our modeling framework, having
 1122 information about δ informs our understanding about θ . Recall that the
 1123 distribution of (θ, δ) for all subjects depends on the hyper mean ϕ and the
 1124 hyper variance-covariance matrix Σ . If we partition these hyperparameters
 1125 according to what they correspond to in the behavioral and neural submodels,
 1126 we obtain the following division:

$$\begin{aligned}\phi &= [\phi_1 \ \phi_2]^\top \\ \Sigma &= \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}.\end{aligned}$$

1127 Here, parameters ϕ_1 and Σ_{11} correspond to behavioral submodel parameters
 1128 θ , whereas parameters ϕ_2 and Σ_{22} correspond to neural submodel parameters
 1129 δ . Given the properties of our linking function, if we know that δ equals a
 1130 specific value, say $\delta = \delta^*$, then the distribution of θ *conditional* on δ is

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

1131 where

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

1132 The important result is that the variance of the conditional distribution is
1133 reduced, specifically it is reduced by $\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. While this amount does
1134 not depend on the specific value of δ^* , it clearly depends on δ 's variability.
1135 Furthermore, as the correlation between the neural and behavioral data
1136 approaches zero, the joint model reduces to a univariate model, such that
1137 $\phi^* = \phi_1$ and $\Sigma^* = \Sigma_{11}$. The gain of this from a modeling standpoint is that
1138 the constraint applied to both θ and δ becomes negligible, and the model
1139 effectively learns to ignore the neural data in its estimate of θ . Finally, the
1140 mean of the conditional estimate for θ does depend on δ^* , which suggests
1141 that the influence and constraint provided by the neural data will vary from
1142 one subject to another, and this variation is explained by the properties of
1143 that subject's neural data *relative* to the group. Specifically, as a subject's
1144 neural data become more unlike the group data (i.e., as $\delta^* - \phi_2$ increases),
1145 the constraint applied to the conditional distribution ($\theta|\delta = \delta^*$) increases.
1146 These results could have implications for a joint model's performance, and
1147 might explain why for some subjects, a joint model performed worse than a
1148 behavioral-data-only model (e.g., see Figure 7).

1149 *5.3. The Shrinkage Paradox*

1150 In the hierarchical models presented in this article, we assumed a hi-
1151 erarchical structure that captured both subject-specific and group-specific
1152 effects. One of the major benefits of hierarchical modeling is the “borow-
1153 ing” of information from one subject (i.e., the lower level of the hierarchy)
1154 to another. This happens because the parameters for any given subject are
1155 informed by that subject's data, but the pattern of subject-specific param-
1156 eters is informed through the estimation of group-level parameters. While
1157 this may seem like a bottom-up process, it is actually also top-down. For
1158 example, when a new subject is added to the data, the model can already
1159 generate predictions for that subject's data on the basis of the other subjects
1160 in the data set. In this way, the estimation of subject-specific parameters in
1161 a hierarchical model consists of a mixture of two quantities: (1) the infor-
1162 mation learned about the subject from the subject's own data, and (2) what the
1163 model predicts for the subject's parameters given the other subjects in the

1164 data set. In statistical terms, these quantities are known as the likelihood
1165 function and the prior predictive distribution, respectively.

1166 For the types of experiments we typically see in psychology and neuro-
1167 science, the prior predictive distribution can be quite sparse, having density
1168 in regions of the parameter space that are not informative for a specific sub-
1169 ject. However, as the number of observations increase, the model’s reliance
1170 on the prior predictive distribution diminishes because the likelihood function
1171 begins to dominate what we know about this individual. Said another way,
1172 the influence of the likelihood is proportional to the number of observations
1173 for that subject. In an analogous way, the influence of the prior predictive
1174 distribution is proportional to the number of subjects in the extant data set.
1175 With many subjects in the set, the prior predictive distribution will be more
1176 reflective of the average subject-to-subject patterns in the set.

1177 The tradeoff dynamics between the prior predictive distribution and the
1178 likelihood function are important because they have a direct effect on a
1179 model’s predictive accuracy. Consider a new subject who has just been added
1180 to a small (e.g., 20 subjects) set of data. Further suppose that the data for
1181 this subject are relatively sparse, consisting of only a few observations. If
1182 the first 20 subjects had been fit hierarchically, the model’s prediction for
1183 the parameters of this new subject would primarily be generated from the
1184 prior predictive density as a result of this subject’s sparse data. However, if
1185 we had estimated the new subject’s parameters *independently* of the hierar-
1186 chical model, the parameter estimates would only consist of the information
1187 in the data themselves, and would not have the information conveyed by
1188 the hierarchical model. If we then compared the estimates of this subject’s
1189 parameters obtained from the hierarchical model to those obtained indepen-
1190 dently, we should expect some differences. A typical pattern of results is
1191 that the independently-obtained estimates are more extreme – meaning they
1192 depart from the group average – than the estimates put forth by the hierar-
1193 chical model (Rouder and Lu, 2005; Rouder et al., 2005). This phenomenon
1194 is known as “shrinkage”, and is sometimes regarded as a disadvantage of
1195 hierarchical modeling.

1196 Shrinkage can be explained simply as a model’s tendency to hedge its bets.
1197 If we know very little about a new subject, but we know a lot about what
1198 subjects typically look like, it would make sense to make predictions about a
1199 subject that reflect our most informed base of knowledge. Statistically speak-
1200 ing, this is actually the optimal behavior (Gelman et al., 2004; Christensen
1201 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; Cottreau 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-quality 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

₁₃₁₃ allows for seamless integration across multiple studies, modalities, and even
₁₃₁₄ imbalanced experimental designs (see, e.g., Figure 2).

₁₃₁₅ **6. Conclusions**

₁₃₁₆ In this article, we have described a method for integrating EEG, fMRI,
₁₃₁₇ and behavioral data into one cognitive model. The model assumes the pres-
₁₃₁₈ ence of different submodels that capture the modality-specific effects of inter-
₁₃₁₉ est. Important parameters describing changes across experimental conditions
₁₃₂₀ are then linked together to facilitate communication across modalities in the
₁₃₂₁ data set. Our approach has many advantages, in particular, it is amenable
₁₃₂₂ to imbalanced experimental designs including different modalities, and dif-
₁₃₂₃ ferent numbers of subjects per modality. Our method is also well suited for
₁₃₂₄ generating predictions for new data, withheld data, or missing data for any
₁₃₂₅ modality of interest. Finally, our approach centers on the instantiation of a
₁₃₂₆ particular mechanistic cognitive theory, allowing us to interpret brain data
₁₃₂₇ through the lens of a cognitive model, and provide a unifying theoretical
₁₃₂₈ framework for multi-modal measures of cognition.

1329 **7. References**

- 1330 Amano, K., Goda, N., Nishida, S., Ejima, Y., Takeda, T., Ohtani, Y., 2006.
1331 Estimation of the timing of human visual perception from magnetoen-
1332 cephalography. *Journal of Neuroscience* 26, 3981–3991.
- 1333 Astolfi, L., Cincotti, F., Mattia, D., Salinari, S., Babiloni, C., Basilisco, A.,
1334 Rossini, P. M., Ding, L., Ni, Y., He, B., Marciani, M. G., Babiloni, F.,
1335 2004. Estimation of the effective and functional human cortical connec-
1336 tivity with structural equation modeling and directed transfer function
1337 applied to high-resolution EEG. *Magnetic Resonance Imaging* 22, 1457–
1338 1470.
- 1339 Bai, X., Towle, V. L., He, E. J., He, B., 2007. Evaluation of cortical cur-
1340 rent density imaging methods using intracranial electrocorticograms and
1341 functional MRI. *NeuroImage* 35, 598–608.
- 1342 Beckmann, C. F., Smith, S. M., 2005. Tensorial extensions of independent
1343 component analysis for multi-subject fMRI analysis. *NeuroImage* 25, 294–
1344 311.
- 1345 Boehm, U., Van Maanen, L., Forstmann, B., Van Rijn, H., 2014. Trial-
1346 by-trial fluctuations in CNV amplitude reflect anticipatory adjustment of
1347 response caution. *NeuroImage* 96, 95–105.
- 1348 Brown, S., Heathcote, A., 2008. The simplest complete model of choice re-
1349 action time: Linear ballistic accumulation. *Cognitive Psychology* 57, 153–
1350 178.
- 1351 Busemeyer, J., Townsend, J., 1993. Decision Field Theory: A dynamic-
1352 cognitive approach to decision making in an uncertain environment. *Psy-
1353 chological Review* 100, 432–459.
- 1354 Calhoun, V., Wu, L., Kiehl, K., Eichele, T., Pearson, G., 2011. Aberrant
1355 processing of deviant stimuli in Schizophrenia revealed by fusion of fMRI
1356 and EEG data. *Acta Neuropsychiatry* 22, 127–138.
- 1357 Calhoun, V. D., Adali, T., 2009. Feature-based fusion of medical imaging
1358 data. *IEEE Transactions on Information Technology in Biomedicine* 13,
1359 711–720.

- 1360 Calhoun, V. D., Adali, T., Liu, J., 2006. A feature-based approach to combine
1361 functional MRI, structural MRI and EEG brain imaging data. Proceedings
1362 of the 28th IEEE EMBS Annual International Conference.
- 1363 Calhoun, V. D., Liu, J., Adali, T., 2009. A review of group ICA for fMRI data
1364 and ICA for joint inference of imaging, genetic, and ERP data. NeuroImage
1365 45, 163–172.
- 1366 Cassey, P., Gaut, G., Steyvers, M., Brown, S., 2015. A generative joint model
1367 for spike trains and saccades during perceptual decision making, under
1368 review.
- 1369 Christensen, R., Johnson, W., Branscum, A., Hanson, T. E., 2011. Bayesian
1370 Ideas and Data Analysis: An Introduction for Scientists and Statisticians.
1371 CRC Press, Taylor and Francis Group, Boca Ranton, FL.
- 1372 Correa, N. M., Adali, T., Li, Y., Calhoun, V. D., 2010a. Canonical correlation
1373 analysis for data fusion and group inferences: Examining applications of
1374 medical imaging data. Signal Processing Magazine 27, 39–50.
- 1375 Correa, N. M., Eichele, T., Adali, T., Li, Y., Calhoun, V. D., 2010b. Multi-
1376 set canonical correlation analysis for the fusion of concurrent single trial
1377 ERP and functional MRI. NeuroImage 50, 1438–1445.
- 1378 Correa, N. M., Li, Y., Adali, T., Calhoun, V. D., 2008. Canonical correlation
1379 analysis for feature based fusion of biomedical imaging modalities and its
1380 application to detection of associative networks in Schizophrenia. Journal
1381 of Selected Topics in Signal Processing 2, 998–1007.
- 1382 Cottereau, B. R., Ales, J. M., Norcia, A. M., 2015. How to use fMRI func-
1383 tional localizers to improve EEG/MEG source estimation. Journal of Neu-
1384 roscience Methods 250, 64–73.
- 1385 Dähne, S., Bießmann, F., Samek, W., Haufe, S., Goltz, D., Gundlach, C.,
1386 Villringer, A., Fazli, S., Müller, K.-R., 2015. Multivariate machine learning
1387 methods for fusing multimodal functional neuroimaging data. Proceedings
1388 of the IEEE 103, 1507–1530.
- 1389 Daw, N. D., 2011. Trial-by-trial data analysis using computational models.
1390 Decision making, affect, and learning: Attention and performance XXIII
1391 23, 1.

- 1392 De Martino, F., Valente, G., de Borst, A. W., Esposito, F., Roebroeck, A.,
1393 Goebel, R., Formisano, E., 2010. Multimodal imaging: An evaluation of
1394 univariate and multivariate methods for simultaneous EEG/fMRI. Magnetic Resonance Imaging 28, 1104–1112.
- 1395
- 1396 Delorme, A., Makeig, S., 2004. EEGLAB: An open source toolbox for analysis
1397 of single-trial EEG dynamics including independent component analysis.
1398 Journal of Neuroscience Methods 134, 9–21.
- 1399 Ditterich, J., 2010. A comparison between mechanisms of multi-alternative
1400 perceptual decision making: Ability to explain human behavior, predictions
1401 for neurophysiology, and relationship with decision theory. Frontiers in
1402 Neuroscience 4, 184.
- 1403 Eichele, T., Calhoun, V. D., Debener, S., 2009. Mining EEG-fMRI using in-
1404 dependent component analysis. International Journal of Psychophysiology
1405 75, 53–61.
- 1406 Eichele, T., Specht, K., Moosmann, M., Jongsma, M. L., Quiroga, R. Q.,
1407 Nordby, H. e. a., 2005. Assessing the spatiotemporal evolution of neuronal
1408 activation with single-trial event-related potentials and functional MRI
1409 102, 17798–17803.
- 1410 Forstmann, B. U., Anwander, A., Schäfer, A., Neumann, J., Brown, S., Wa-
1411 genmakers, E.-J., Bogacz, R., Turner, R., 2010. Cortico-striatal connec-
1412 tions predict control over speed and accuracy in perceptual decision mak-
1413 ing. Proceedings of the National Academy of Sciences 107, 15916–15920.
- 1414 Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., von Cramon, D. Y.,
1415 Ridderinkhof, K. R., Wagenmakers, E.-J., 2008. Striatum and pre-SMA
1416 facilitate decision-making under time pressure. Proceedings of the National
1417 Academy of Sciences 105, 17538–17542.
- 1418 Forstmann, B. U., Tittgemeyer, M., Wagenmakers, E.-J., Derrfuss, J., Im-
1419 perati, D., Brown, S., 2011a. The speed-accuracy tradeoff in the elderly
1420 brain: A structural model-based approach. Journal of Neuroscience 31,
1421 17242–17249.
- 1422 Forstmann, B. U., Wagenmakers, E.-J., 2015. An introduction to model-
1423 based cognitive neuroscience. Springer, New York, New York.

- 1424 Forstmann, B. U., Wagenmakers, E.-J., Eichele, T., Brown, S., Serences,
1425 J. T., 2011b. Reciprocal relations between cognitive neuroscience and formal
1426 cognitive models: opposites attract? *Trends in Cognitive Sciences* 15, 272–
1427 279.
- 1428 Franco, A. R., Ling, J., Caprihan, A., Calhoun, V. D., Jung, R. E., Heileman,
1429 G. L. et al., 2008. Multimodal and multi-tissue measures of connectivity re-
1430 vealed by joint independent component analysis. *Journal of Selected Topics*
1431 in Signal Processing
- 2, 986–997.
- 1432 Friston, K., Harisson, L., Penny, W., 2003. Dynamic causal modeling. *NeuroImage* 19, 1273–1302.
- 1433
- 1434 Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B., 2004. Bayesian Data
1435 Analysis. Chapman and Hall, New York, NY.
- 1436 Gershman, S. J., Blei, D. M., Pereira, F., Norman, K. A., 2011. A topographic
1437 latent source model for fMRI data. *Neuroimage* 57, 89–100.
- 1438 Green, D. M., Swets, J. A., 1966. Signal detection theory and psychophysics.
1439 Wiley Press, New York.
- 1440 Hamandi, K., Powell, H. W., Laufs, H., Symms, M. R., Barker, G. J., Parker,
1441 G. J., Lemieux, L., Duncan, J. S., 2008. Combined EEG-fMRI and tractog-
1442 raphy to visualise propagation of epileptic activity. *Journal of Neurology*
1443 *Neurosurgery Psychiatry* 79, 594–597.
- 1444 Heathcote, A., Brown, S. D., Wagenmakers, E.-J., 2015. An introduction to
1445 good practices in cognitive modeling. In: Forstmann, B. U., Wagenmak-
1446 ers, E.-J. (Eds.), *An introduction to model-based cognitive neuroscience*.
1447 Springer, New York, pp. 25–48.
- 1448 Ho, T., Brown, S., van Maanen, L., Forstmann, B. U., Wagenmakers, E.-
1449 J., Serences, J. T., 2012. The optimality of sensory processing during the
1450 speed-accuracy tradeoff. *J Neurosci* 32, 7992–8003.
- 1451 Huster, R. J., Debener, S., Eichele, T., Herrmann, C. S., 2013. Methods for
1452 simultaneous EEG-fMRI: An introductory review. *The Journal of Neuro-
1453 science* 32, 6053–6060.

- 1454 Kayser, J., Tenke, C. E., 2006. Principal components analysis of Laplacian
1455 waveforms as a generic method for identifying ERP generator patterns: II.
1456 Adequacy of low-density estimates. *Clinical Neurophysiology* 117, 369–380.
- 1457 Kirby, K. N., 2009. One-year temporal stability of delay-discount rates. *pbr*
1458 16, 457–462.
- 1459 Krishnan, A., Williams, L. J., McIntosh, A., Abdi, H., 2015. Partial least
1460 squares (PLS) methods for neuroimaging: A tutorial and review. *NeuroImage* 56, 455–475.
- 1462 Lee, M. D., Wagenmakers, E.-J., 2013. Bayesian modeling for cognitive sci-
1463 ence: A practical course. Cambridge University Press.
- 1464 Lin, F. H., McIntosh, A. R., Agnew, J. A., Eden, G. F., Zeffiro, T. A.,
1465 Belliveau, J. W., 2003. Multivariate analysis of neuronal interactions in
1466 the generalized partial least squares framework: Simulations and empirical
1467 studies. *NeuroImage* 20, 625–642.
- 1468 Liu, J., Calhoun, V. D., 2007. Parallel independent component analysis for
1469 multimodal analysis: Application to fMRI and EEG data. *Proceedings of*
1470 IEEE ISBI conference.
- 1471 Love, B. C., 2015. The algorithmic level is the bridge between computation
1472 and brain. *Topics in Cognitive Science* 7.
- 1473 Mack, M. L., Preston, A. R., Love, B. C., 2013. Decoding the brain's algo-
1474 rithm for categorization from its neural implementation. *Current Biology*
1475 23, 2023–2027.
- 1476 Manning, J. R., Ranganath, R., Norman, K. A., Blei, D. A., 2014. Topo-
1477 graphic Factor Analysis: A Bayesian model for inferring brain networks
1478 from neural data. *PLoS One* 9, e94914.
- 1479 Marr, D., 1982. Vision: A Computational Investigation into the Human Rep-
1480 resentation and Processing of Visual Information. New York: Freeman.
- 1481 Martinez-Montes, E., Valdes-Sosa, P. A., Miwakeichi, F., Goldman, R. I.,
1482 Cohen, M. S., 2004. Concurrent EEG/fMRI analysis by multiway partial
1483 least squares. *NeuroImage* 22, 1023–1034.

- 1484 Mulder, M. J., Wagenmakers, E.-J., Ratcliff, R., Boekel, W., Forstmann,
1485 B. U., 2012. Bias in the brain: A diffusion model analysis of prior proba-
1486 bility and potential payoff. *Journal of Neuroscience* 32, 2335–2343.
- 1487 Mulert, C., Lemieux, L., 2009. EEG-fMRI: Physiological basis, technique
1488 and applications, 1st Edition. Springer, Berlin.
- 1489 Norman, K., Polyn, S., Detre, G., Haxby, J., 2006. Beyond mind-reading:
1490 Multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*.
- 1491 Nunez, M. D., Srinivasan, R., Vandekerckhove, J., 2015. Individual differ-
1492 ences in attention influence perceptual decision making. *Frontiers in Psy-
1493 chology* 8 (18), 1–13.
- 1494 Palmeri, T., Schall, J., Logan, G., 2015. Neurocognitive modelling of percep-
1495 tual decisions. In: Busemeyer, J. R., Townsend, J., Wang, Z. J., Eidels, A.
1496 (Eds.), *Oxford Handbook of Computational and Mathematical Psychol-
1497 ogy*. Oxford University Press.
- 1498 Plummer, M., Best, N., Cowles, K., Vines, K., March 2006. CODA: Conver-
1499 gence diagnosis and output analysis for MCMC. *R News* 6 (1), 7–11.
1500 URL <http://CRAN.R-project.org/doc/Rnews/>
- 1501 Purcell, B., Heitz, R., Cohen, J., Schall, J., Logan, G., Palmeri, T., 2010.
1502 Neurally-constrained modeling of perceptual decision making. *Psycholog-
1503 ical Review* 117, 1113–1143.
- 1504 Ratcliff, R., 1978. A theory of memory retrieval. *Psychological Review* 85,
1505 59–108.
- 1506 Robert, C. P., Casella, G., 2004. Monte Carlo statistical methods. Springer,
1507 New York, NY.
- 1508 Rodriguez, C. A., Norcia, A. M., Cohen, M. X., McClure, S. M., 2015a. Fron-
1509 toparietal cortex dynamics of value integration in intertemporal choice,
1510 manuscript in preparation.
- 1511 Rodriguez, C. A., Turner, B. M., McClure, S. M., 2014. Intertemporal choice
1512 as discounted value accumulation. *PLoS ONE* 9, e90138.

- 1513 Rodriguez, C. A., Turner, B. M., Van Zandt, T., McClure, S. M., 2015b.
1514 The neural basis of value accumulation in intertemporal choice. European
1515 Journal of Neuroscience.
- 1516 Rouder, J. N., Lu, J., 2005. An introduction to Bayesian hierarchical models
1517 with an application in the theory of signal detection. Psychonomic Bulletin
1518 and Review 12, 573–604.
- 1519 Rouder, J. N., Lu, J., Speckman, P., Sun, D., Jiang, Y., 2005. A hierarchical
1520 model for estimating response time distributions. Psychonomic Bulletin
1521 and Review 12, 195–223.
- 1522 Shadlen, M. N., Newsome, W. T., 2001. Neural basis of a perceptual deci-
1523 sion in the parietal cortex (area LIP) of the rhesus monkey. Journal of
1524 Neurophysiology 86, 1916–1936.
- 1525 Shiffrin, R. M., Lee, M. D., Kim, W., Wagenmakers, E.-J., 2008. A survey
1526 of model evaluation approaches with a tutorial on hierarchical Bayesian
1527 methods. Cognitive Science 32, 1248–1284.
- 1528 Sui, J., Adali, T., Yu, Q., Chen, J., Calhoun, V. D., 2012. A review of
1529 multivariate methods for multimodal fusion of brain imaging data. Journal
1530 of Neuroscience Methods 204, 68–81.
- 1531 Summerfield, C., Koechlin, E., 2010. Economic value biases uncertain per-
1532 ceptual choices in the parietal and prefrontal cortices. Front Hum Neurosci
1533 4, 208.
- 1534 Teipel, S. J., Bokde, A. L., Meindl, T., Amaro, E. J., Soldner, J., Reiser,
1535 M. F., Herpertz, S. C., Moller, H. J., Hampel, H., 2010. White matter
1536 microstructure underlying default mode network connectivity in the human
1537 brain. NeuroImage 49, 2021–2032.
- 1538 Turner, B. M., 2015. Constraining cognitive abstractions through Bayesian
1539 modeling. In: Forstmann, B. U., Wagenmakers, E.-J. (Eds.), An intro-
1540 duction to model-based cognitive neuroscience. Springer, New York, pp.
1541 199–220.
- 1542 Turner, B. M., Forstmann, B. U., Love, B., Palmeri, T., Van Maanen,
1543 L., 2015a. Approaches of analysis in model-based cognitive neuroscience,
1544 manuscript under review.

- 1545 Turner, B. M., Forstmann, B. U., Wagenmakers, E.-J., Brown, S. D., Seder-
1546 berg, P. B., Steyvers, M., 2013a. A bayesian framework for simultaneously
1547 modeling neural and behavioral data. *NeuroImage* 72, 193–206.
- 1548 Turner, B. M., Sederberg, P. B., Brown, S., Steyvers, M., 2013b. A method
1549 for efficiently sampling from distributions with correlated dimensions. *Psy-
1550 chological Methods* 18, 368–384.
- 1551 Turner, B. M., Van Maanen, L., Forstmann, B. U., 2015b. Combining cogni-
1552 tive abstractions with neurophysiology: The neural drift diffusion model.
1553 *Psychological Review* 122, 312–336.
- 1554 Ullsperger, M., Debener, S., 2010. Simultaneous EEG and fMRI: Recording,
1555 analysis, and application. Oxford University Press, USA.
- 1556 Usher, M., McClelland, J. L., 2001. On the time course of perceptual choice:
1557 The leaky competing accumulator model. *Psychological Review* 108, 550–
1558 592.
- 1559 van Maanen, L., Brown, S. D., Eichele, T., Wagenmakers, E.-J., Ho, T.,
1560 Serences, J., 2011. Neural correlates of trial-to-trial fluctuations in response
1561 caution. *Journal of Neuroscience* 31, 17488–17495.
- 1562 Van Maanen, L., Forstmann, B., Keuken, M., Wagenmakers, E.-J., Heath-
1563 cote, A., 2015. The impact of MRI scanner environment on perceptual
1564 decision making, in Press.
- 1565 White, C. N., Congdon, E., Mumford, J. A., Karlsgodt, K. H., Sabb, F. W.,
1566 Freimer, N. B., London, E. D., Cannon, T. D., Bilder, R. M., Poldrack,
1567 R. A., 2014. Decomposing decision components in the stop-signal task:
1568 A model-based approach to individual differences in inhibitory control.
1569 *Journal of Cognitive Neuroscience* 26, 1601–1614.
- 1570 Wilson, T. D., Houston, C. E., Etling, K. M., Brekke, N., 2015. Is model
1571 fitting necessary for model-based fMRI? *PLoS Computational Biology* 11,
1572 e1004237.
- 1573 Xu, L., Pearson, G., Calhoun, V. D., 2009. Joint source based morphometry
1574 identifies linked gray and white matter group differences. *NeuroImage* 44,
1575 777–789.

1576 **Appendix A. Technical Details of the Behavioral Model**

1577 While the basic idea of our hierarchical LBA model for intertemporal
 1578 choice was presented in Rodriguez et al. (2014) (also see Turner et al., 2013b),
 1579 some of the core assumptions of the model have been changed to facilitate a
 1580 comparison between the behavioral model and the joint models used in the
 1581 subsequent sections.

1582 Figure A.18 provides an illustrative diagram of our LBA model of in-
 1583 tertemporal choice. To provide a formal description of the model, we denote
 1584 the RT on the i th trial for the j th subject in the v th value condition as
 1585 $RT_{i,j,v} \in (0, \infty)$, and the corresponding choice as $C_{i,j,v}$ where $C_{i,j,v} \in \{I, D\}$.
 1586 I and D are the immediate and delayed rewards respectively. The model as-
 1587 sumes that evidence for I and D is accumulated independently in separate ac-
 1588 cumulators. For a given trial, both accumulators begin with some choice bias,
 1589 which is provided as independent amounts of starting point evidence $\{a_I, a_D\}$,
 1590 sampled from a common uniform distribution $\mathcal{U}[0, A_j]$, where A_j is the upper
 1591 bound of the starting point for Subject j . Evidence then increases through
 1592 time at rates $\{d_I, d_D\}$, which are sampled from independent normal distri-
 1593 butions with means $\{\mu_{v,I,j}, \mu_{v,D,j}\}$. Mean accumulation rates vary across value
 1594 conditions (i.e., the first index) and subjects (i.e., the third index), whereas
 1595 the standard deviation σ_j is the same for both immediate (I) and delayed
 1596 (D) alternatives, but varies across subjects. Therefore, $d_I \sim \mathcal{N}(\mu_{v,I,j}, \sigma_j)$ and
 1597 $d_D \sim \mathcal{N}(\mu_{v,D,j}, \sigma_j)$. Each accumulator gathers evidence ballistically until the
 1598 point at which one accumulator reaches the response threshold b_j . The ob-
 1599 served response time is the sum of the decision time plus some extra time τ_j
 1600 attributed to non-comparison and selection processes, such as temporal dis-
 1601 counting and motor execution. Letting $\{a_I, a_D\} = \mathbf{a}_{i,j,v}$ and $\{d_I, d_D\} = \mathbf{d}_{i,j,v}$,
 1602 the response time on Trial i for Subject j is given by

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

1603 and the choice $C_{i,j,v}$ corresponds to the accumulator that reached the response
 1604 threshold b_j first. The model provides a closed-form and joint account of
 1605 response time and choice probability across value conditions by specifying
 1606 “defective” probability density functions (PDFs) for I and D in terms of the
 1607 parameters just described (see Brown and Heathcote, 2008, for details).

1608 We made two reparameterization choices from our previous application
 1609 of the model (Rodriguez et al., 2014) to facilitate the hierarchical estimation

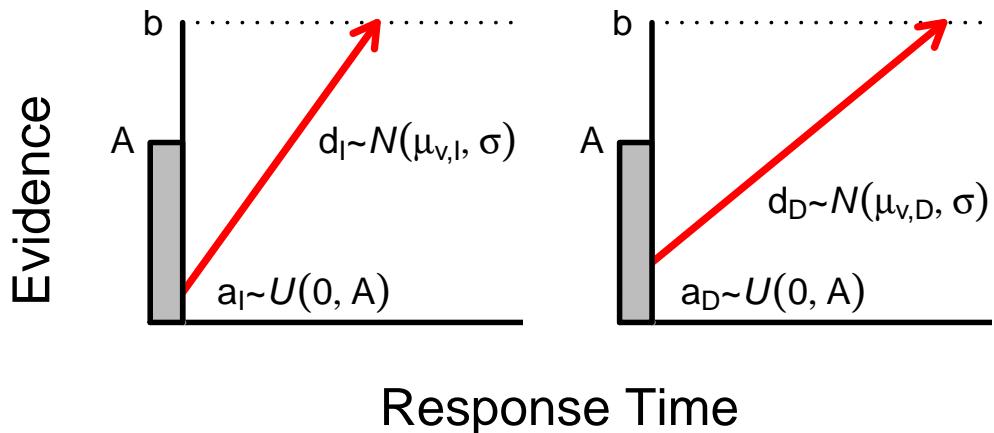


Figure A.18: Our Linear Ballistic Accumulator model for intertemporal choice. Following the presentation of a stimulus and some non-decision time τ , information accumulates ballistically for each alternative. The accumulation processes for both the immediate (left panel) and the delayed (right panel) 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.

1610 procedure (see below). The first involved modeling the upper boundary of
 1611 the starting point A_j in terms of the response boundary b_j . The starting
 1612 point parameter A_j is bounded by 0 and b_j , and this parameter-specific
 1613 boundary can result in poor estimation efficiency in our sampling procedure.
 1614 Consequently, we chose to model the proportion of the upper starting point
 1615 boundary parameter to the response threshold parameter by estimating

$$\alpha_j = A_j/b_j.$$

1616 The parameter α_j is still bounded by 0 and 1, and so in the specification
 1617 of the model below, we will use the logit transformation to put α_j on a
 1618 space that has infinite support (i.e., $\alpha_j \in (-\infty, \infty)$), which will improve the
 1619 accuracy of our estimated posteriors.

1620 Following our folding process described in the methods, our data consisted
 1621 of three response time distributions, one for each of the value conditions V
 1622 (see Figure 5). Our second reparametrization involved the drift rate param-
 1623 eters for these response time distributions. Of considerable importance was
 1624 the difference in the accumulation rates for the most likely alternative (given
 1625 our experimental manipulation) across these value conditions. To model this
 1626 effect, we introduced the parameters $\eta_j^{(1)}$, and $\eta_j^{(2)}$ to capture the difference
 1627 in the rates of accumulation from the $V3$ condition to the $V2$, and from the
 1628 $V3$ condition to the $V1$ condition, respectively. Formally, we set

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

1629 where $\text{logit}^{-1}(x)$ represents the inverse logit transformation of x , ν_j is the
 1630 “equipreference” drift rate parameter for Subject j , and “SHV” is the sub-
 1631 jectively higher valued option. To satisfy scaling conditions of the model, we
 1632 imposed a constraint such that the drift rates sum to one across each of the
 1633 value conditions, namely that in condition V , the following should hold:

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

1634 We estimated LBA model parameters using a hierarchical Bayesian pro-
 1635 cedure (see Rodriguez et al., 2014; Turner et al., 2013b, for details). To

1636 construct the hierarchical model, we must first make assumptions about the
 1637 distributions of subject-specific parameters. To do this, we first transformed
 1638 the parameters so that they had continuous, infinite support (i.e., can take
 1639 on any real value). In our case, for parameters bounded by zero, we applied
 1640 a log transformation, and for parameters bounded by zero and one, we ap-
 1641 plied a logit transformation. Following these transformations, we specified
 1642 the following priors for the subject-specific parameters:

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

1643 where $k \in \{1, 2, 3\}$ represents the three different folded value conditions, as
 1644 described above. The nondecision time parameter τ_j was allowed to vary
 1645 across value conditions because previous modeling results showed that this
 1646 choice provided the best fit to the data (Rodriguez et al., 2014).

1647 Finally, we complete our model by specifying priors for the hyperpara-
 1648 meters. Given the assumptions of normality for the subject-specific parameters,
 1649 we can establish a conjugate relationship between the prior and posterior dis-
 1650 tributions at the hyper level. To do this, we specified the following priors for
 1651 the group-level means:

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

1652 and the following priors for the group-level standard deviations,

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

1653 where $\Gamma^{-1}(a, b)$ denotes the inverse gamma distribution with shape parameter 1654 a , and scale parameter b . This particular choice of a and b for the 1655 priors produces a skewed distribution placing weight on appropriate ranges 1656 of the parameters. Specifically, when $a = 4$ and $b = 10$, the prior has an 1657 approximate 95% credible set of $(1.14, 9.05)$, and an expected value of 3.32. 1658 These choices reflect our *a priori* beliefs: we did not expect the between- 1659 subject variability to be less than 1, and felt that larger values would become 1660 increasingly less likely to account for these data.

1661 While our prior selections were informed by other similar modeling applications 1662 (see, e.g., Rodriguez et al., 2014; Turner et al., 2013b), we remained 1663 somewhat conservative in our choices to avoid undue parameter constraint, 1664 because the model structure was considerably different from prior research 1665 using the hierarchical version of the LBA model.

1666 Appendix B. Technical Details of the Bivariate Joint Model

1667 The joint model used in both the EEG and fMRI analyses was equivalent, 1668 although both the neural and behavioral data changed. Where possible, we 1669 kept the behavioral portion of the joint model equal to the behavioral-data- 1670 only model, described above. However, given the joint structure of some 1671 components of the model, the models are not exactly equivalent in the way 1672 they capture the data.

1673 We begin by discussing the neural portion of the joint model. We first 1674 processed the neural data to provide a single-trial measure of neural activity 1675 for each individual subject (see the main text for details). We denote the 1676 neural data on the i th trial for the j th subject in the v th value condition as 1677 $N_{i,j,v}$. Because we folded the data, $v \in \{1, 2, 3\}$, where $v = 3$ when $P_D =$

1678 $\{0.5\}$, $v = 2$ when $P_D = \{0.3, 0.7\}$, and $v = 1$ when $P_D = \{0.1, 0.9\}$. Because
 1679 the neural data for each trial is an average, by the central limit theorem,
 1680 each $N_{i,j,v}$ is normally distributed. For the $v = 3$ condition, the parameter
 1681 ϵ_j captures the mean neural activation. Similar to the η parameters in the
 1682 behavioral model, we modeled the difference in neural activation across value
 1683 conditions by introducing the variables $\delta_j^{(1)}$ and $\delta_j^{(2)}$ for the $v = 2$ and $v =$
 1684 1 value conditions, respectively. We also capture the variability in neural
 1685 activation with the parameter ζ_j , which was constrained to be the same
 1686 across each value condition. Formally, we assume the neural data for the j th
 1687 subject in the v th value condition arise from the following distributions:

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

1688 For each subject, we then assumed the following priors for ϵ_j and ζ_j :

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

1689 For the group-level parameters, we specified the following priors for the
 1690 group-level means:

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

1691 and the following priors for the group-level standard deviations

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

1692 The principle behind joint modeling is to use both neural and behavioral
 1693 data to enforce constraint on a particular cognitive theory (Turner et al.,
 1694 2013a; Turner, 2015). In our modeling application, we wish to use the neural
 1695 signature across value conditions to constrain the differences in the rate of
 1696 accumulation across these value conditions. Following this motivation, we
 1697 specify a link between the parameters $\delta_j^{(1)}$ and $\eta_j^{(1)}$, as well as between the

¹⁶⁹⁸ parameters $\delta_j^{(2)}$ and $\eta_j^{(2)}$, by specifying

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

¹⁶⁹⁹ 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
¹⁷⁰⁰ $\phi^{(k)} = [\delta_\mu^{(k)}, \eta_\mu^{(k)}]$, where $\eta_\mu^{(k)}$ is the critical behavioral model group-level parameter
¹⁷⁰¹ by which we draw a comparison to the behavioral-data-only model
¹⁷⁰² (see Figures 7 and 10). Specifying the structure in this way is equivalent to
¹⁷⁰³ the following specification:
¹⁷⁰⁴

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

¹⁷⁰⁵ where

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

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

¹⁷⁰⁶
¹⁷⁰⁷ where ρ_1 and ρ_2 are correlation parameters. However, as we will see below,
¹⁷⁰⁸ with the appropriate selections for the prior distributions, one can establish
¹⁷⁰⁹ a conjugate relationship between the prior and posterior of $\phi^{(k)}$ and $\Sigma^{(k)}$,
¹⁷¹⁰ making the estimation more efficient via Gibbs sampling (Robert and Casella,
¹⁷¹¹ 2004; Gelman et al., 2004). With the exception of $\eta_j^{(1)}$ and $\eta_j^{(2)}$, all other
¹⁷¹² parameter specifications for the behavioral submodel were equivalent to the
¹⁷¹³ behavioral-data-only model presented above.

¹⁷¹⁴ The final step in constructing the joint model is to specify priors for the
¹⁷¹⁵ hyperparameters $\phi^{(k)}$ and $\Sigma^{(k)}$ for $k \in \{1, 2\}$. Following Turner et al. (2015b),
¹⁷¹⁶ we specified a joint prior on $\phi^{(k)}$ and $\Sigma^{(k)}$ so that

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

1717 where

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

1718 where $\mathcal{W}^{-1}(a, b)$ denotes the inverse Wishart distribution with dispersion
1719 matrix a and degrees of freedom b . We set d_0 equal to the number of linked
1720 parameters plus two (i.e., $d_0 = 2 + 2 = 4$), Φ is the identity matrix (i.e.,
1721 a matrix containing ones on the diagonal, and zeros on the off diagonal) of
1722 dimension (2×2) , $s_0 = 1/10$, and μ_0 is a vector containing two zeros. These
1723 choices were made to establish a conjugate relationship between the prior and
1724 posterior, so that analytic expressions could be derived for the conditional
1725 distributions of $\phi^{(k)}$ and $\Sigma^{(k)}$, while still specifying uninformative priors.

1726 When estimating the joint posterior distribution of the bivariate model
1727 parameters, we used a combination of Gibbs sampling for the group-level
1728 (Gelman et al., 2004), and differential evolution with Markov chain Monte
1729 Carlo for the subject-level (DE-MCMC; Turner et al., 2013b). For the sub-
1730 ject level estimates, we used 24 chains, and obtained 10,000 samples after a
1731 burn-in period of 10,000 samples. Thus, our estimates of the joint posterior
1732 distributions are based on 240,024 samples. The burn-in period allowed us
1733 to converge quickly to the high-density regions of the posterior distribution,
1734 while the rest of the samples allowed us to improve the reliability of the esti-
1735 mates. Convergence and overall fidelity of the chains were assessed through
1736 visual inspection and the coda package in R (Plummer et al., 2006).

1737 Appendix C. Technical Details of the Trivariate Joint Model

1738 Where possible, the specification of the trivariate model was equivalent
1739 to the bivariate model above. The important difference between these two
1740 models is the additional data structure, with the trivariate model containing
1741 both EEG and fMRI. Using the notation in Figure 1, we denote the EEG
1742 data as $E_{i,j,v}$, and the fMRI data as $F_{i,j,v}$. In the model, we super script the
1743 parameters ϵ_j , ζ_j , $\delta_j^{(1)}$, and $\delta_j^{(2)}$ with an F for the fMRI data and an E for
1744 the EEG data. Hence, we assume the neural data arise from the following

1745 distributions:

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

1746 The prior structure on $\epsilon_j^{(E)}$, $\epsilon_j^{(F)}$, $\zeta_j^{(E)}$, and $\zeta_j^{(F)}$ for the trivariate model were
1747 identical to the prior structure on ϵ_j and ζ_j in the bivariate model above.
1748 However, in the trivariate model, the linking procedure relating neural and
1749 behavioral submodels is slightly different. Namely, we assume

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

1750 The prior structure on ϕ and Σ in the trivariate model is identical to the
1751 prior structure in the bivariate model above, with the following exceptions:
1752 $d_0 = 3 + 2 = 5$, Φ is the identity matrix of dimension (3×3) , and μ_0 is a
1753 vector containing three zeros.

1754 When estimating the joint posterior distribution of the trivariate model
1755 parameters, we ran our sampling algorithm with 24 chains for 20,000 samples
1756 after a burnin period of 10,000 samples, and then thinned the samples by re-
1757 taining every other iteration. This process helped to reduce autocorrelation
1758 in the chains, which we suspected would arise due to the high dimensionality
1759 of the model. Hence, the estimates of the posterior distribution of the trivari-
1760 ate model were also based on 240,024 samples, to maintain consistency with
1761 the bivariate models. Convergence and overall fidelity of the chains were as-
1762 sessed through visual inspection and the coda package in R (Plummer et al.,
1763 2006).