van Ravenzwaaij, Don, Provost, Alexander & Brown, Scott D. "A confirmatory approach for integrating neural and behavioral data into a single model" Published in *Journal of Mathematical Psychology*, Vol. 76, p. 131-141, (2017).

Available from: http://dx.doi.org/10.1016/j.jmp.2016.04.005

A Confirmatory Approach for Integrating Neural and Behavioral Data into a Single Model

Don van Ravenzwaaij, Alexander Provost, and Scott D. Brown
University of Newcastle

Word count: 5,800

Correspondence concerning this article should be addressed to:
Don van Ravenzwaaij
University of Newcastle, Department of Psychology
University Drive, Aviation Building, room AVG11
Callaghan NSW 2308, Australia
Ph: (+61) 2–4921–5662
E–mail should be sent to don.vanravenzwaaij@newcastle.edu.au.

Abstract

Recent decades have witnessed amazing advances in both mathematical models of cognition and in the field of cognitive neuroscience. These developments were initially independent of one another, but recently the fields have started to become interested in joining forces. The resulting joint modeling of behavioral and neural data can be difficult, but has proved fruitful. We briefly review different approaches used in decision–making research for linking behavioral and neural data, and also provide an example. Our example provides a tight link between behavioral data and evoked scalp potentials measured during mental rotation. The example model illustrates a powerful hypothesis–driven way of linking such data sets. We demonstrate the use of such a model, provide a model comparison against interesting alternatives, and discuss the conclusions that follow from applying such a joint model.

Keywords: Joint Modeling; Cognitive Neuroscience; Response Time Data; ERP.

Like many areas of scientific enquiry, cognitive psychology began with verbally–specified theories and gradually progressed to quantitative accounts over time. This resulted in mathematical models to describe memory (e.g., Atkinson & Shiffrin, 1968; Raaijmakers & Shiffrin, 1981), categorization (e.g., Nosofsky, 1986; Nosofsky & Palmeri, 1997), speeded

This research was supported by fellowship grants from the Australian Research Council to DVR (DE140101181) and SDB (FT120100244).

and unspeeded decision making (e.g., Ratcliff, 1978; Wagenmakers, 2009), and many other paradigms (for reviews, see Lee & Wagenmakers, 2013; Lewandowsky & Farrell, 2010). A separate and at first mostly unrelated development was the advent of cognitive neuroscience. This field sought to map changes in the brain as they related to cognition, using neural measurements obtained through event–related potentials (ERPs; e.g., Sutton, Braren, Zubin, & John, 1965; Hillyard, Hink, Schwent, & Picton, 1973), the magnetoencephalogram (MEG; e.g., Brenner, Williamson, & Kaufman, 1975), functional magnetic resonance imaging (fMRI; e.g., Belliveau et al., 1991), and single–unit recordings in non–human primates (e.g., Hanes & Schall, 1996; Schall, 2001; Shadlen & Newsome, 1996). As progressively more precise measures of the inner workings of the brain became available, researchers have become increasingly capable at understanding the neural determinants of cognitive processes.

Some research paradigms have well–specified and tractable mathematical models of cognition, and also well–developed methods for neural measurement, for example, simple decision–making and reinforcement learning. Researchers interested in such paradigms started investigating ways to link the neural and behavioral data more carefully. The latest developments include so–called *joint models*, in which data of one kind can inform the model fit of the other kind and vice versa (e.g., Purcell et al., 2010; Turner, Forstmann, et al., 2013; Anderson & Fincham, 2014; Turner, Forstmann, Love, Palmeri, & van Maanen, in press). These accounts aim for the most explicit and careful links, by simultaneously modeling neural recordings and behavioral outputs, allowing both kinds of data to inform model selection and parameter estimation. Joint modeling provides an important theoretical contribution: it allows a researcher to examine common denominators underlying both behavioral data and neural data.

In this paper, we provide an example of how to jointly model behavioral and neural data from simple decision–making. As an illustrative example, we apply a joint model of behavioral responses and EEG recordings to data from an experiment based on the classic Shepard–Metzler mental rotation task (Shepard & Metzler, 1971). However, before describing the model, we review different approaches to linking behavioral and neural data, with a focus on decision–making research.

An important change in the development of decision–making models over the past twenty years has been a steady "tightening" of the link between neural and behavioral data (for reviews and discussion of linking behavioral and neural data, see Teller, 1984). Early models of simple decision–making linked behavioral and neural data loosely, by constraining the development of behavioral models to respect data from neural measurements. For example, the leaky competing accumulator model developed by Usher and McClelland (2001) was structurally constrained to include components supported by neural investigations, such as lateral inhibition between accumulating units, and passive decay of accumulated evidence. These links were included as part of the model development process, and thereafter there was no further attempt to link neural with behavioral data.

Subsequent models tested the links via qualitative comparisons between predictions for corresponding neural and behavioral data sets. This kind of linking was very common in early research into decision–making with fMRI methods, in which predictions were based on the assumption that an experimental manipulation will influence one particular model component, which leads naturally to predictions for the behavioral data, and also for the neural data (via the hypothesized link). Predictions most frequently take the form "in condition A vs. B, behavioral measure X should increase while neural measure Y decreases". Support for the predictions is taken as evidence in favor of the model, including the hypothesized link. As an example, Ho, Brown, and Serences (2009) tested predictions generated from decision–making models via hypothesized neural links. In one part of their study, Ho et al. manipulated the difficulty of a decision-making task and hypothesized that this should result in a change in the speed of evidence accumulation in a sequential sampling model. By examination of the model coupled to a standard model for haemodynamic responses, Ho et al. generated predictions for the blood–oxygen–level dependent (BOLD) response profile within regions that are involved in perceptual decision making. These predictions were compared with data from an fMRI experiment, which lent support to some accounts over others.

Linking via the testing of qualitative hypotheses was later surpassed by quantitative approaches, which provided a tighter link between neural and behavioral data. The most common example of quantitative linking in decision–making models takes parameters of the decision–making model, estimated from behavioral data, and compares them against the parameters of a descriptive model estimated from the neural data. For example, Forstmann et al. (2008) correlated individual subjects' model parameters, estimated from behavioral data, against blood–oxygen–level dependent (BOLD) parameter estimates; subjects with large changes in threshold parameters also showed similarly large changes in BOLD responses.

Most recently, there have been efforts to link neural and behavioral decision–making data even more tightly, by combining both data sets in a single model–based analysis. This approach has culminated in models such as that developed by Purcell et al. (2010) which uses neural measurements as a model input in order to predict both behavioral measurements and a second set of neural measurements. This provides a simultaneous description of neural and behavioral data sets, as well as explicating the links between them. A less detailed, but more general approach was developed by Turner, Forstmann, et al. (2013) and extended by Turner et al. (in press) in this volume. In their method, neural and behavioral models are joined by allowing their parameters to covary. Turner, Forstmann, et al.'s approach is a "joint" model, in the sense that it allows symmetric information flow: behavioral data can influence the neural parameter estimates, and neural data can influence the behavioral parameter estimates. This information flow is achieved via a covariance matrix for the model parameters. This structure allows the identification of covariance between model parameters associated with neural processes and model parameters associated with behavioral processes. However, Turner, Forstmann, et al.'s approach differs from our analyses in its focus. The covariance matrix of Turner, Forstmann, et al.'s approach means that any and all parameters of the behavioral model are allowed to link with any and all parameters of the neural model, although all these links are required to be linear. Our approach is less general, but more pointed, because it requires the specific instantiation of a single, precise link between one parameter of the neural model and one parameter of the behavioral model.[1]

The joint modeling approach of Turner, Forstmann, et al. (2013) is complementary

[1]While it is true that Turner, Forstmann, et al.'s method could, in theory, be restricted to produce our approach (e.g. by setting almost all priors on the covariance matrix components to zero, and by adding in nonlinear parameter link functions) in practice this has not been done.

to the approach we use. For paradigms in which there exist precise hypotheses about the links between neural and behavioral models, our approach offers a straightforward way of instantiating and testing these hypotheses. For paradigms in which this is not the case, Turner, Forstmann, et al.'s approach offers a powerful tool for exploration. What both approaches have in common is that they *jointly* fit the neural and behavioral data, which allows behavioral data to influence parameters on the "neural side" of the model, and vice versa. A joint model in this sense is able to identify a compromise between the two streams of data. This means that, compared to an otherwise–identical model that is fit solely to the behavioral (or neural) data, a joint model will always fit more poorly. Coherently managing the compromise between fitting neural and behavioral data streams is a strength of the joint modeling approach. For example, suppose one was examining a joint model for behavioral and neural data, but was not fitting the model in a "joint" manner. Instead, imagine the model was examined by fitting first to behavioral data alone, and then later evaluating the model by comparing its subsequent predictions for neural effects against the neural data. One problem with this approach arises if the model had two sets of parameters (say, A and B) which both provided very good fits to the behavioral data, but very different fits to the neural data. Suppose that parameter set A provided slightly better behavioral fits, but also terrible neural fits, while parameter set B provided good fits to the neural data. Fitting to the behavioral data alone would lead the researcher to choose parameter set A, and then to reject the model because of the terrible fit to neural data. Joint fitting allows identification of compromise parameters (such as set B) which provide good fits to both data streams.

The two–stage approach to model evaluation, in which the flow of information between the two types of data is mostly one–way, was employed by Purcell et al. (2010) (they used two different neural data streams, only one of which was a fitting target). While we hope that a joint modeling approach has some strengths that the two–stage approach does not, Purcell et al.'s work included important other advantages that have been absent in the joint modeling work to date. For example, Purcell et al.'s approach was used to conduct pointed comparisons between competing hypotheses about both the underlying model structures, and the hypothesized links between neural and behavioral data. While such comparisons are, theoretically, possible in joint modeling approaches, they can be difficult to implement, and have not been investigated to date. The joint model we describe below is an attempt to combine the advantages of the confirmatory approach of Purcell et al. (2010) with the sophisticated estimation approach of Turner, Forstmann, et al. (2013). Similar to Turner, Forstmann, et al.'s approach, we employ a simultaneous estimation procedure. However, our approach is confirmatory in that we test an explicit and pre–specified link between neural and behavioral data. We fit both behavioral and neural data streams at the same time. In the next section, we will describe the behavioral task as well as the two types of data.

Data

The data we use are from an experiment based on the classic Shepard–Metzler *mental rotation task* (Provost, Johnson, Karayanidis, Brown, & Heathcote, 2013). The mental rotation task is a two–alternative forced choice task in which participants are asked to examine a pair of stimuli, one of which is rotated relative to the other. Crucially, participants are asked to indicate as quickly and accurately as possible whether the stimuli are identical ("same") or whether one is different from the other ("different"). For instance, in the left panel of Figure 1, the right stimulus is the same as the left stimulus. On the other hand, in the right panel of Figure 1, the right stimulus is the mirror–image of the left stimulus ("different").

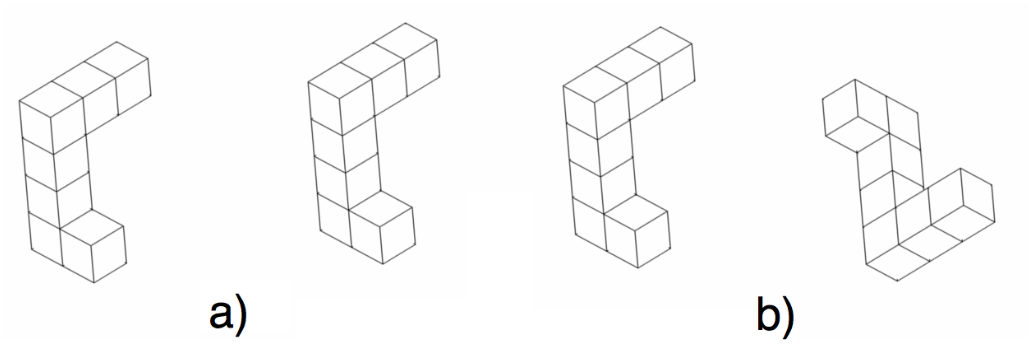

Figure 1. Two sample stimuli from Provost et al. (2013). Left panel: the right stimulus is the same as the left stimulus. Right panel: the right stimulus is different from the left stimulus.

The data we use here is from the first session of the first experiment reported by Provost et al. (2013). The experiment included five conditions that differed in the angle of rotation of the right stimulus: 0°, 45°, 90°, 135°, and 180°. The left stimulus was always identical to the one displayed in Figure 1.

Within each condition, half of the stimuli were "same" and half were "different". The corresponding behavioral data were response times and choices from all conditions, for all participants. The neural data we will consider are mean amplitudes of single trials of the ERP signal corresponding to each trial used in the behavioral analysis. As in Provost et al. (2013), we report ERP effects from the midline parietal electrode site Pz, with a common average reference. Comparing mean amplitudes at Pz we are able to model a specific ERP modulation called "rotation related negativity" (RRN; Heil, 2002; Riečanský & Jagla, 2008), which is considered an index of mental rotation. Specifically, we look at increased mean amplitude negativity associated with increased angular displacement across 8 epochs, from 200 to 1,000ms post stimulus onset in 100ms windows. For more details, please see the methods section of Experiment 1 and Figure 4 of Provost et al. (2013).

The use of sequential accumulator models for the analysis of response time data is not new (e.g., Link & Heath, 1975; Ratcliff, 1978; Wagenmakers, 2009). For these data, we turn to a relatively recent accumulator model: the Linear Ballistic Accumulator (LBA; Brown & Heathcote, 2008). An advantage of the LBA is tractability, as it has an easily–computed closed–form expression for its likelihood. As a result, it is relatively straightforward to expand the model to include a neural component. In the next section, we will introduce the reader to the behavioral and neural components of the model and demonstrate how we combine them into a joint model.

The Modeling

In the first sub–section below, we introduce the behavioral model. In the second sub–section, we introduce the neural modeling and the link between the two elements.

The Behavioral Level: LBA

In the LBA for multi–alternative RT tasks (Brown & Heathcote, 2008), stimulus processing is conceptualized as the accumulation of information over time. A response is initiated when the accumulated evidence reaches a predefined threshold. An illustration for two response options is given in Figure 2.

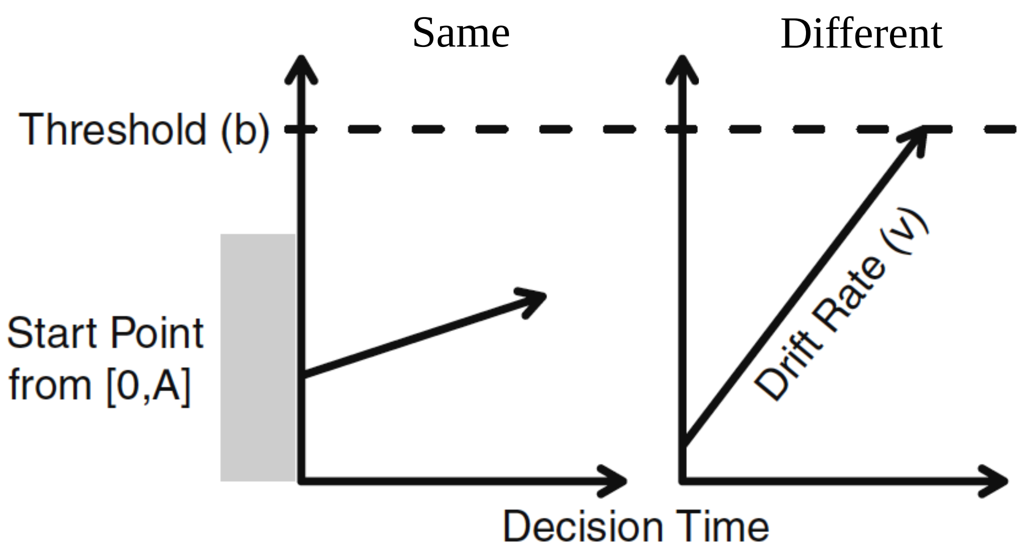

Figure 2. The LBA and its parameters for two response options (for this trial, the "different" response is the correct answer). Evidence accumulation begins at a start point drawn randomly from a uniform distribution with interval [0, A]. Evidence accumulation is governed by drift rate d, drawn across trials from a normal distribution with mean ν and standard deviation s. A response is given as soon as one accumulator reaches threshold b. Observed RT is an additive combination of the time during which evidence is accumulated and non–decision time t_0.

The LBA assumes that the decision process starts from a random point between 0 and A, after which information is accumulated linearly for each response option. The rate of this evidence accumulation is determined by drift rates d_1 and d_2, normally distributed over trials with means ν_1 and ν_2, and common standard deviation s. The distribution of drift rates is truncated at zero to prevent negative accumulation rates. Threshold b determines the speed–accuracy tradeoff; lowering b leads to faster RTs at the cost of a higher error rate.

Together, these parameters generate a distribution of decision times DT. The observed RT, however, also consists of stimulus–nonspecific components such as response preparation and motor execution, which together make up non–decision time t_0. The model assumes that t_0 simply shifts the distribution of DT, such that $RT = DT + t_0$ (Luce, 1986). Hence, the three key components of the LBA are (1) the speed of information processing, quantified by mean drift rate ν; (2) response caution, quantified by boundary separation that averages to $b - A/2$; and (3) non–decision time, quantified by t_0. The LBA has been successfully applied to a number of experimental paradigms including random dot motion tasks, brightness discrimination, consumer choice, and many others (e.g., Rae, Heathcote, Donkin, Averell, & Brown, 2014; Trueblood, Brown, & Heathcote, 2014; Ho, Brown, Abuyo, Ku, & Serences, 2012).

We specified the standard behavioral aspects of the LBA model using 24 parameters per participant for the 20 different response time distributions. The parameters included: one upper range of starting point A parameter, two parameters for threshold b (one each for "same" and "different" responses), ten parameters for both correct drift v_c and for error drift v_e (two stimuli types — "same", "different" — times five angle conditions — 0°, 45°, 90°, 135°, 180°), and one non–decision time parameter t_0. The 20 different response time distributions (and 10 free response probabilities) arose from factorial combination of two stimulus classes (same vs. different) with two response classes (same vs. different) and five rotation angles. This parameterization was chosen because it provides a reasonable compromise between goodness-of-fit and tractability, as demonstrated in the extensive analyses of alternative models for data from a related experiment (Provost & Heathcote, 2015). Importantly, the evidence accumulators of the model have been linked to neural activity in the brain (e.g., Purcell et al., 2010; Gold & Shadlen, 2007). Because of this, mean drift rate ν lends itself naturally to be the driving parameter behind our ERP data.

We used a hierarchical Bayesian implementation of the LBA (Turner, Sederberg, Brown, & Steyvers, 2013). Advantages of the hierarchical Bayesian framework include the ability to fit the LBA to data with relatively few trials, because the model borrows strength from the hierarchical structure. The Bayesian set–up allows for using MCMC sampling, which is an efficient approach to parameter estimation (Gamerman & Lopes, 2006; Gilks, Richardson, & Spiegelhalter, 1996; van Ravenzwaaij, Cassey, & Brown, in press). Starting points for the Markov chains were drawn from the following distributions: $A \sim N(2, 0.2)|(0,)$, both bs $\sim N(1, 0.1)|(0,)$, all ten ν_cs $\sim N(3, 0.3)|(0,)$, all ten ν_es $\sim N(1, 0.1)|(0,)$, and $t_0 \sim N(0.2, 0.02)|(0,)$. In this notation, $\sim N(x, y)|(0,)$ indicates that a parameter is normally distributed with mean x, standard deviation y, and is truncated to positive values only.

The hierarchical set–up prescribes that all individual parameters come from a truncated Gaussian group–level distribution. Thus, for each parameter to be estimated, we estimated a group level mean parameter and a group level standard deviation parameter. Priors for all group level mean parameters were normal distributions, with $A_\mu \sim N(2, 1)|(0,)$, both $b_\mu \sim N(2, 1)|(0,)$, all ten $\nu_{c\mu} \sim N(3, 1)|(0,)$, all ten $\nu_{e\mu} \sim N(1, 1)|(0,)$, and $t_{0\mu} \sim N(0.2, 0.1)|(0,)$. Priors for all group level standard deviation parameters were gamma distributions with a shape and a scale parameter of 1, except for $t_{0\sigma}$ which has a scale parameter of 3. Starting point distributions for group level μ were all identical to starting point distributions for the individual parameters, and starting point distributions for group level σ parameters were derived from starting point distributions for the individual parameters by dividing the mean by 10 and the standard deviation by 2. These prior settings are quite uninformative, and are based on previous experience with parameter estimation for the LBA model. As a result, the specific settings will not have a large influence on the shape of the posterior. For more details on distributional choices for the priors, we refer the reader to Turner, Sederberg, et al. (2013).

For sampling, we used 32 interacting Markov chains, and ran each for 1,000 burn-in iterations followed by 1,000 iterations after convergence. The two tuning parameters of the differential evolution proposal algorithm were set to standard values used in previous work: random perturbations were added to all proposals drawn uniformly from the interval [-.001, .001]; and the scale of the difference added for proposal generation was set to $\gamma = 2.38 \times (2K)^{-0.5}$, where K is the number of parameters per participant (24, in the model described above). The MCMC chains blocked proposals separately for each participant's parameters, and also blocked the group–level parameters in $\{\mu, \sigma\}$ pairs.

Linking to Neural Data

The behavioral model above, based on the LBA, specifies a likelihood function for the response time data which gives the likelihood of observed data conditional on any given set of parameter values. This likelihood function supports all of our statistical analyses. The first step in bringing the neural data into the model is to define a likelihood function for the ERPs. We will assume that the ERP data, within any particular condition for any particular subject, are normally distributed. The next step is to link the parameters of the behavioral LBA model above with the parameters of the assumed normal model for the ERP data. To begin, we assume that the standard deviation of the normal distribution is fixed everywhere, for each subject, and that the mean of the normal distribution is given by an offset parameter (α) plus the drift rate parameter times a scale parameter (β):

$$ERP \sim N(\alpha + v \times \beta, \sigma) \tag{1}$$

The model is graphically displayed in Figure 3. Equation 1 provides a precise instantiation of the linking hypothesis in this joint model. Our very simple hypothesis is that the neural and behavioral data are linked via the drift rate parameter of the model, and that the link is a simple linear function. While simple to specify, this link has complicated implications. For example, the predicted ERP signal will change across conditions whenever drift rate changes – with rotation angle and with same vs. different stimulus pairings, in our experiment. The link also implies particular constraints on the model. For example, the drift rate parameter is forced to accommodate changes in both behavioral and neural data due to changes in rotation angle. The linking parameter serves as a time–sensitive measure of the link between behavioral and neural data. The model does more than just re–describe this link: the model attempts to capture the fact that different rotation angles cause different ERP measurements with a linking function and linking parameters that are identical for all angle conditions. As such, the model accounts for different ERPs across conditions entirely through drift rate.

For the neural data, we estimated one offset parameter α, one standard deviation parameter σ, and eight scale parameters β (one for each 100 ms epoch from 200 up to

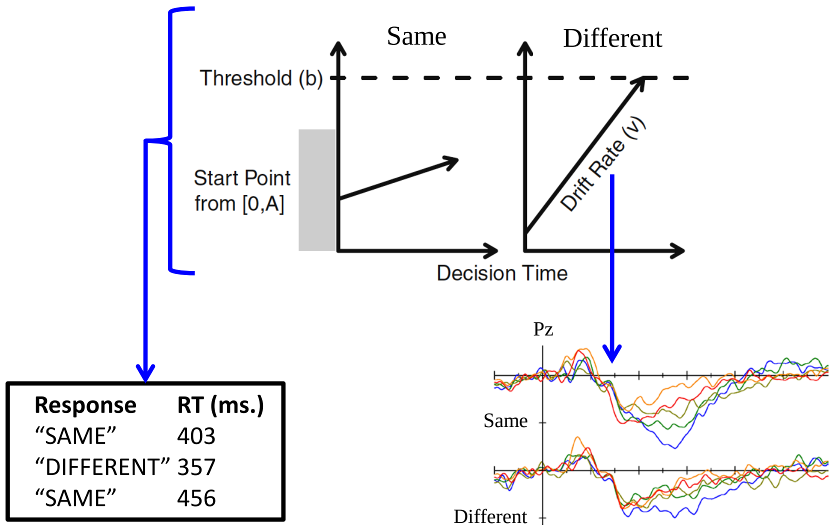

Figure 3. A symbolic display of the joint model. All LBA parameters inform the behavioral response data (bottom–left). The drift rate corresponding to the response given by the participant informs the ERP data (bottom–right).

1,000ms). The eight scale parameters allow investigation of how strongly the ERP signal is linked to cognition across the eight different time windows (200ms–300ms, 300ms–400ms, $\cdots$, 900ms–1,000ms).

Starting points for the linking parameters were drawn from the following distributions: $\alpha \sim N(8, 0.8)|(0,)$, all eight βs $\sim N(1, 0.1)|(0,)$, and $\sigma \sim N(5, 0.5)|(0,)$. Analogous to the LBA parameters, all individual linking parameters were drawn from a truncated Gaussian group–level distribution. Priors for all group level mean parameters are normal distributions, with $\alpha_\mu \sim N(8, 2)|(0,)$, all eight $\beta_\mu \sim N(1, 1)|(0,)$, and $\sigma_\mu \sim N(5, 1)|(0,)$. Priors for all group level standard deviation parameters are gamma distributions with a shape and a scale parameter of 1. Starting point distributions for group level μ were all identical to starting point distributions for the individual parameters, and starting point distributions for group level σ parameters were derived from starting point distributions for the individual parameters by dividing the mean by 10 and the standard deviation by 2.

Data and code for the full model may be found on the web.[2]

[2] `www.donvanravenzwaaij.com/Papers`, and also at `https://osf.io/2r7bv/`.

Model Comparison

It is very difficult to judge model fit in an absolute sense. What constitutes a good fit, how much of a misfit is acceptable? In practice, it is almost always more fruitful to examine comparative goodness-of-fit, and to compare different models. We compare the model described above (henceforth ν*–ERP*) to three competing alternatives:

- t_0*–ERP*: the behavioral parametrization is identical to that of ν*–ERP*, but the linking parameter to the neural data is non–decision time t_0 instead of drift rate ν (see e.g. Pouget et al., 2011, for corroborating evidence).
- *Brev–*t_0*–ERP*: the behavioral parametrization for drift rates and non–decision time are reversed. In this model, we have one ν_c and one ν_e instead of ten each, and we have ten t_0 (one for each stimulus and angle condition) instead of one. Analogous to t_0*–ERP*, the linking parameter to the neural data is non–decision time t_0. Analogous to ν*–ERP*, the linking parameter to the neural data is now free to vary between stimuli types and angle conditions.
- ν*–nonlinear–ERP*: identical to the ν*–ERP* model, but testing a nonlinear link function between the drift rates and the ERP mean parameter. The nonlinear link function we test is the cumulative normal distribution function, which instantiates the hypothesis that scalp potentials might have important ceiling and floor effects. Such effects are plausible for many reasons, for example they may be imposed by physical and physiological limits on the electrical activity and conductivity of the cortex and scalp.

Priors and starting values were analogous for all four models.[3] The models will be compared by visually inspecting the posterior predictives for obvious misfit. Numerically, we compare the models by calculating the Deviance Information Criterion (DIC; Spiegelhalter, Best, Carlin, & van der Linde, 2002), a measure which balances goodness of fit against model complexity. In this sense, DIC is similar to the well–known BIC and AIC measures, but DIC extends these by quantifying model complexity as across–sample variability in model fit rather than simply counting up the number of free parameters. As such, DIC usually assumes a stronger penalty for complexity. Lower values of DIC indicate better support for a model from the data.

Results

ν–ERP

As a first check of model fit, we compared posterior predictive data against the neural and behavioral data in Figure 4. The figure displays data averaged over participants with boxplots representing empirical data and lines representing synthetic data. The left two columns show model correspondence to the .1, .3, .5, .7, and .9 quantiles calculated from correct RTs (green) and error RTs (red). The right two columns show model correspondence to mean ERP amplitudes for each of the eight different time windows (200ms–300ms, 300ms–400ms, $\cdots$, 900ms–1,000ms). The first and third column show model correspondence for "same" stimuli, the second and fourth column show model correspondence for "different"

[3] E.g., all ten t_0 used in the *Brev–*t_0*–ERP* model have the same starting values and prior as the one t_0 used in the ν*–ERP* model; all ten ν_c used in the ν*–ERP* model have the same starting values and prior as the one ν_c used in the *Brev–*t_0*–ERP* model.

stimuli. Rows show model correspondence for different rotation conditions (0°, 45°, 90°, 135°, 180°).

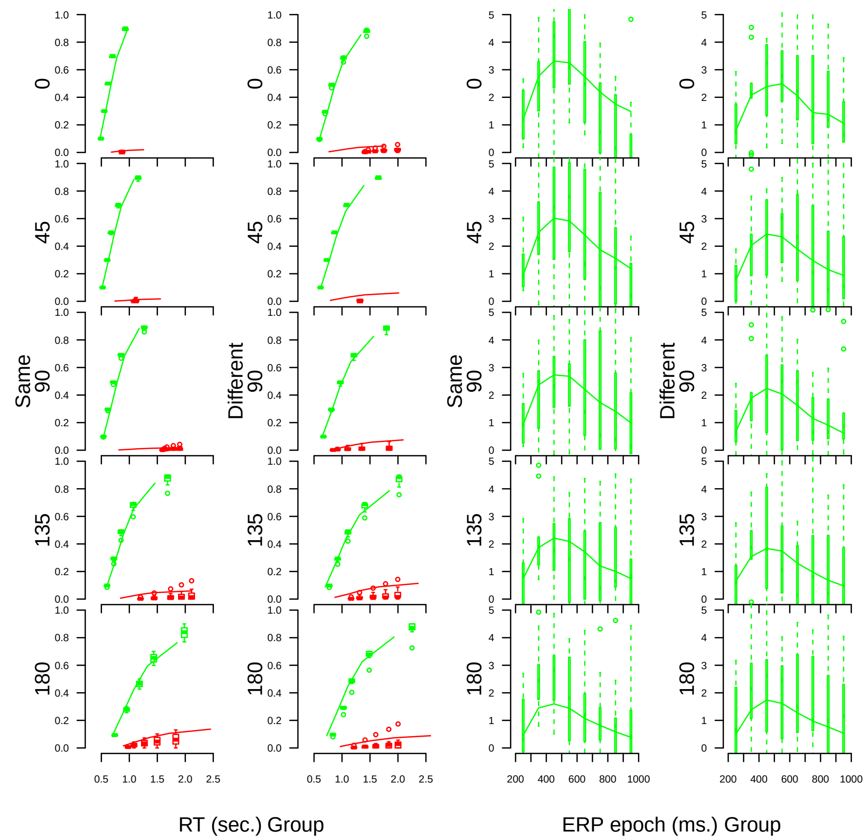

Figure 4. Posterior predictives show that the ν–*ERP* model fit both the behavioral data and the neural data well. Left two columns: proportion correct (y–axis) plotted against RTs (x–axis) for the .1, .3, .5, .7, and .9 quantiles calculated from correct RTs (green) and error RTs (red). Right two columns: mean ERP amplitudes in negative microvolts (y–axis) for eight different time windows (x–axis). For all panels, boxplots represent empirical data and lines represent posterior predictive data.

On the whole, the ν–*ERP* model fit both data sets well, although there is some misfit. The model captures the qualitative changes in RT distributions and percentage correct across same vs. different stimuli, and across the different angles of rotation. There is a tendency for the model to under–predict the accuracy in some conditions, as evidenced by the fact that the green lines are slightly lower than the center of the green boxplots and the red lines slightly higher than the center of the red boxplots. For the neural data, the model seems to capture the ERP distributions over time well. These conclusions about absolute, global model fit are necessarily vague, because of the previously mentioned difficulties in assessing absolute model fit. This is one of the reasons we turn to model comparison below.

The first model comparison we provide is to a behavioral–data–only version of the

ν–*ERP* model. The joint model must necessarily fit more poorly than the behavioral–only model, because the parameters of the joint model are further constrained to accommodate effects in the neural data (in a statistical sense, the behavioural-only model "nests" the behavioral side of the ν–*ERP* model). In order to examine this constraint, we compare the posterior predictives of our joint model fits to posterior predictives of a fit to the behavioral data alone. Parameter settings were as outlined in section "The Behavioral Level: LBA". The posterior predictives for the behavioral-only model can be found in Figure 5.

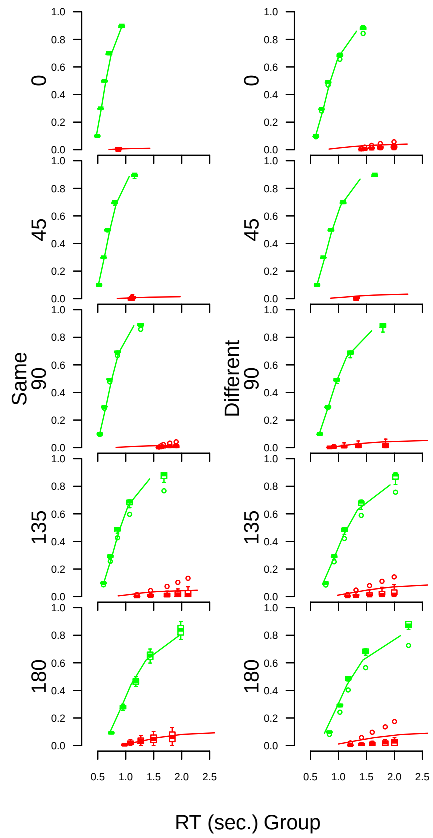

Figure 5. Posterior predictives show that the behavioral-only version of the joint ν–*ERP* model fit the behavioral data well. Displayed are the proportion correct (y–axis) plotted against RTs (x–axis) for the .1, .3, .5, .7, and .9 quantiles calculated from correct RTs (green) and error RTs (red). For both panels, boxplots represent empirical data and lines represent posterior predictive data.

Visual comparison of Figure 4 and Figure 5 shows that the model fit is almost identical. The joint model compromises slightly on the accuracy fit compared to the behavioral–only model, but other than that, the models appear indistinguishable. To provide a statistical comparison, the likelihood of the mean parameters averaged over all participants is -233.93 for the simple model. For the joint model, when selecting the behavioral component of the model and taking the likelihood of those mean parameters averaged over all participants, the value is -255.69, lending further credence to the observation that these models fit the data comparably. It is not appropriate to compare these likelihood values further, e.g. by calculating DIC, because the likelihood of the behavioral data under the joint model does not satisfy the assumptions of those analyses, because of the conditioning on neural data.

We next examine a different, but plausible, candidate for the link between behavioral and neural data: the non–decision time parameter, t_0. We do so by comparing two new models. The first has a behavioral parametrization which is identical to the original ν–*ERP* model, but has a link to the neural data through the t_0 parameter (this model is called t_0–*ERP*). The second model corresponds to the original ν–*ERP* model but with the roles for non–decision time (t_0) and drift rate (ν) reversed (this model is called *Brev*–t_0–*ERP*).

t_0–ERP

Posterior predictive data for the t_0–*ERP* model are shown in Figure 6. Visual inspection of the figure shows that the t_0–*ERP* model fits the behavioral data well, but does not capture the neural data as well as the ν–*ERP* model. This impression is supported by comparison of the DIC values for the two models: ν–*ERP* has an average DIC across participants of 31,880.75, whereas t_0–*ERP* has an average DIC across participants of 31,909.18. Within participants, the ν–*ERP* model was DIC–preferred for 6 out of 9 people.

Brev–t_0–ERP

Posterior predictive data for the *Brev*–t_0–*ERP* model are shown in Figure 7. Visual inspection of the figure shows that the model fits the behavioral data worse than both other models. The *Brev*–t_0–*ERP* model captures the neural data better than the t_0–*ERP* model, but not as good as the ν–*ERP* model. Again, this impression is supported by analysis of DIC values: *Brev*–t_0–*ERP* has an average DIC across participants of 32,039.64, worse than both other models. It is also the model with the poorest DIC out of all models for all nine participants.

ν–nonlinear–ERP

Posterior predictive data for the ν–*nonlinear*–*ERP* are shown in Figure 8. Visual inspection of the figure shows that the ν–*nonlinear*–*ERP* model fits both types of data well, although not as well as the ν–*ERP* model: ν–*ERP* has an average DIC across participants of 31,880.75, whereas ν–*nonlinear*–*ERP* has an average DIC across participants of 31,910.22. Within participants, the ν–*ERP* model was DIC–preferred for 6 out of 9 people. For the remainder of the results section, we will examine the results of ν–*ERP*, the best of the models we have investigated, in more detail.

Central Findings

In sum, we find that of the models we considered, evidence for mean drift rate ν being the linking parameter between behavioral and neural data is strongest. Furthermore, we find that the relationship between mean drift rate ν and the neural data is linear in nature (though for some participants, a nonlinear link provides a better account of the data).

To highlight the central research findings, we will now examine effects across conditions. Summarized data are displayed in Figure 9, with corresponding summaries from the

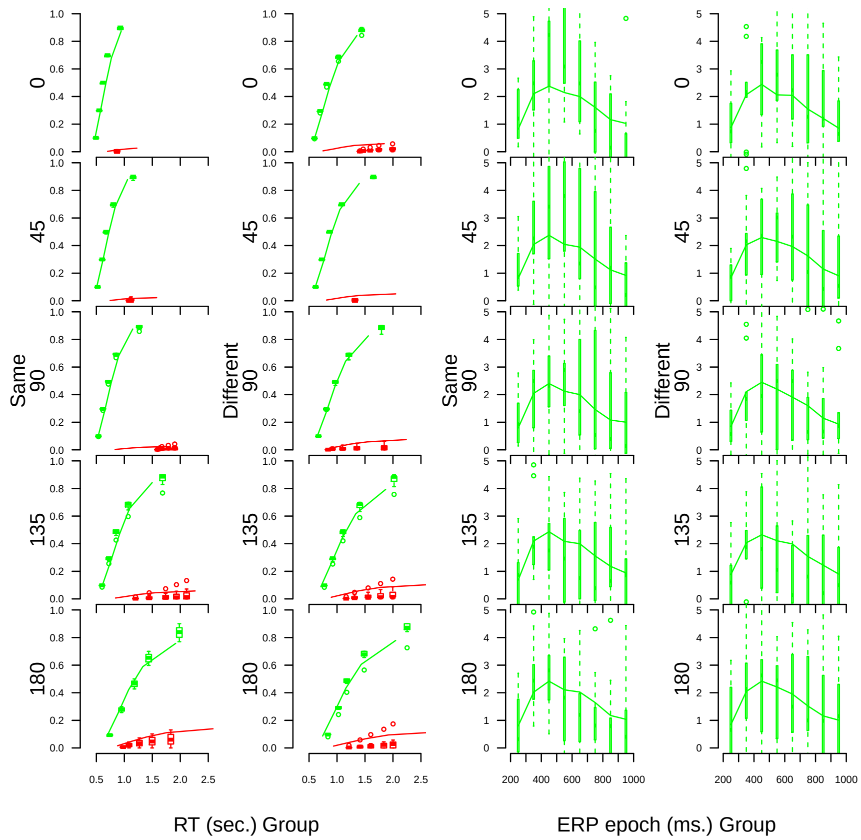

Figure 6. Posterior predictive data show that the t_0–*ERP* model fits the behavioral data well, and the neural data comparatively poorly. Left two columns: proportion correct (y–axis) plotted against RTs (x–axis) for the .1, .3, .5, .7, and .9 quantiles calculated from correct RTs (green) and error RTs (red). Right two columns: mean ERP amplitudes in negative microvolts (y–axis) for eight different time windows (x–axis). For all panels, boxplots represent empirical data and lines represent posterior predictive data.

posterior predictions of the best-supported joint model, ν–*ERP*. The top–left panel displays median RTs for different conditions and stimulus types. RT steadily increases as the rotation angle increases, and also median RT is higher for "different" stimuli than for "same" stimuli. The model captures both of these data patterns very accurately. The bottom–left panel displays mean proportion of correct decisions, separately for different conditions and stimulus types. Accuracy drops as the rotation angle increases, though the trend is less clear than for RTs. The bottom–left panel confirms the earlier observation that the model underestimates some of the accuracies.

The top–right panel displays ERPs for each 100 ms epoch from 200ms up to 1,000ms for "same" stimuli. The amplitude of the ERPs drops as the rotation angle increases. Given the size of the error bars (displayed in the far right of the panel), the mismatch between the data and the model is modest. The bottom–right panel displays ERPs for each 100 ms

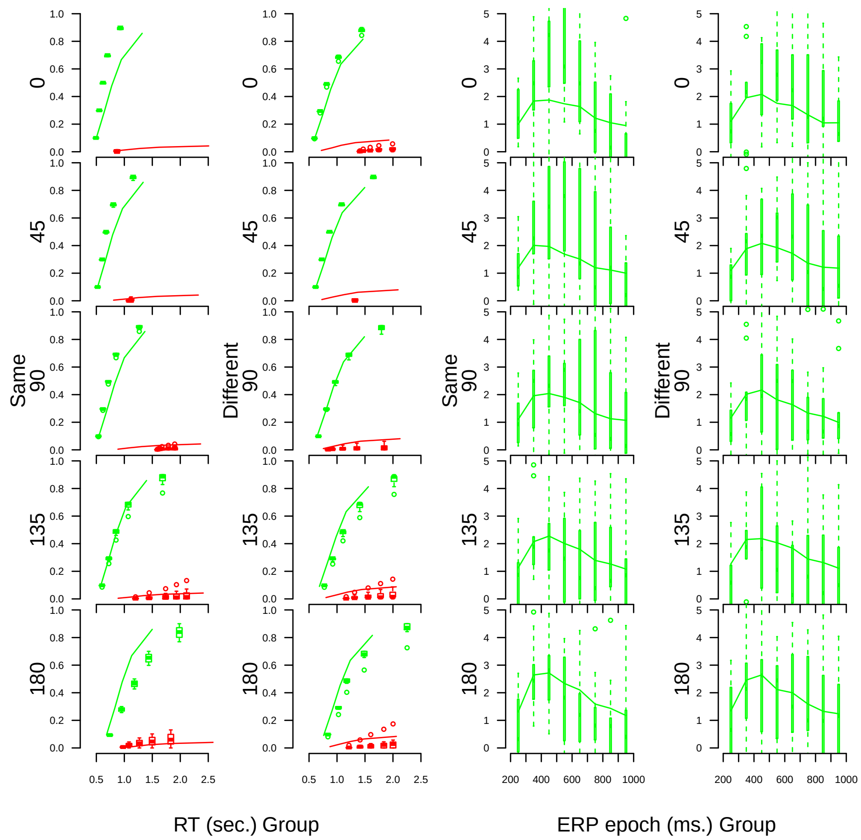

Figure 7. Posterior predictive datas show that the *Brev–t_0–ERP* model fits both the behavioral and the neural data quite poorly. Left two columns: proportion correct (y–axis) plotted against RTs (x–axis) for the .1, .3, .5, .7, and .9 quantiles calculated from correct RTs (green) and error RTs (red). Right two columns: mean ERP amplitudes in negative microvolts (y–axis) for eight different time windows (x–axis). For all panels, boxplots represent empirical data and lines represent posterior predictive data.

epoch from 200 up to 1,000ms for "different" stimuli. Again, the amplitude of the ERPs drops as the rotation factor increases, though interestingly the ERPs for 135° and 180° have reversed order. The model captures the data well, as can be observed by comparing the modest mismatch between data and model to the averaged error bars displayed in the far right of the panel.

The bottom–right panel also includes medians of the posterior distribution over the group level linking parameter (β), with error bars capturing the central 50% of the distribution. The size of linking parameter β follows the amplitude of the rotation–angle effects in the ERP data. To reiterate, estimates of the linking parameter provide time–sensitive measures of the link between behavioral and neural data. For example, at each time window the different rotation angles lead to different ERP measurements (the colored dots vertically spaced), with some time windows showing very little differences between angles and some

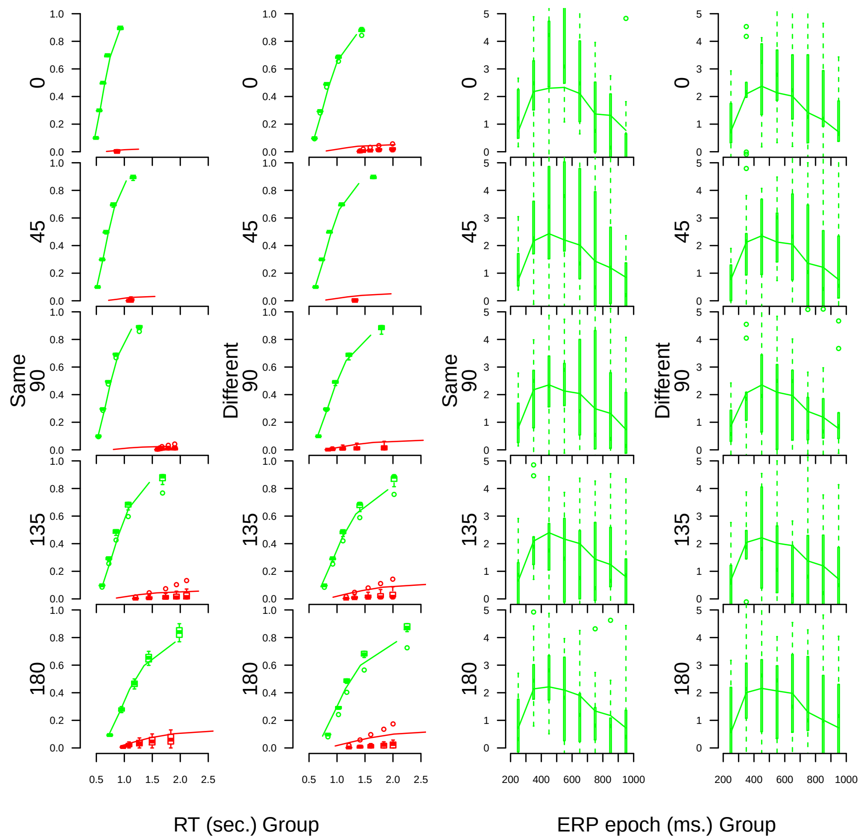

Figure 8. Posterior predictive data show that the *ν–nonlinear–ERP* model fits the behavioral data well, and the neural data comparatively poorly. Left two columns: proportion correct (y–axis) plotted against RTs (x–axis) for the .1, .3, .5, .7, and .9 quantiles calculated from correct RTs (green) and error RTs (red). Right two columns: mean ERP amplitudes in negative microvolts (y–axis) for eight different time windows (x–axis). For all panels, boxplots represent empirical data and lines represent posterior predictive data.

showing very large differences. The model captures these effects, even though the linking function and linking parameters are identical for all angle conditions. This happens because the drift rates are estimated differently for the different angle conditions (and for same vs. different stimulus classes), and these different drift rates influence the ERP predictions via the linking function.

Conclusion

This paper provided an example for cognitive scientists who are interested in investigating the correspondence between neural and behavioral data via building computational models for both data streams. We compared four different models that differ in the parametrization on the behavioral level and in the linking assumptions and showed that drift rate is capable of simultaneously explaining the behavioral data and the neural data.

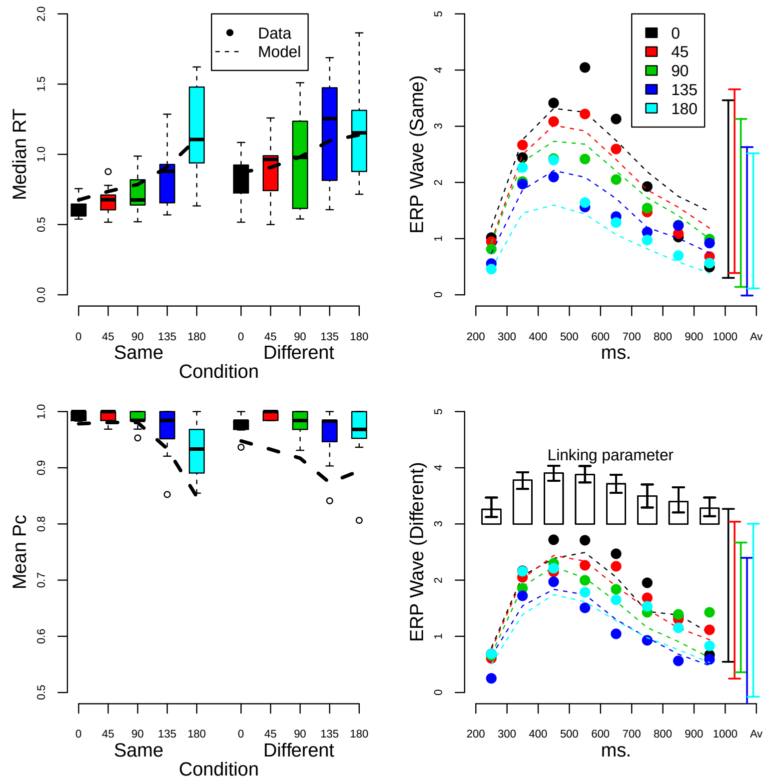

Figure 9. Behavioral and neural data, displayed for each stimulus type and angle condition. Top–left panel: Boxplots display median RT for participants in seconds. Bottom–left panel: Boxplots display proportion correct for participants. Right panels: ERPs in negative microvolts. The far right of both panels display error bars, averaged over all time slots. Inset for the bottom–right panel shows linking parameter β displayed for each time slot with error bars. Error bars in the right panels represent the central 50% of the distribution.

The joint modeling approach that we have used relies on the precise instantiation of hypotheses about the links between parameters related to neural data and parameters related to behavioral data. In addition to specifying which parameters are linked, this approach also requires specification of a particular linking function. The different model versions we investigated differed in these elements, and provided a rigorous framework for investigating important theoretical questions. For example, our model comparisons revealed that, for most participants at least, a simple linear link between drift rate and average ERP amplitude was better than a sigmoidal link. Similarly, an explanation of rotation angle effects in both behavioral and neural data was better when based on drift rates than on non–decision time. It is, of course, entirely conceivable that both drift rate and non–decision time play a role in linking behavioral data to ERP data (recall that DIC preferred the ν–*ERP* model over the t_0–*ERP* model for most, but not all, participants). Provost and Heathcote (2015) explored more sophisticated models using a sequential, two–stage process, to separate the decision process from the mental rotation process. In Provost et al.'s account, the decision processing is delayed by a random amount of time taken to mentally rotate the stimulus, which is equivalent to assuming a random distribution for the non–decision time (t_0) in our model. Provost et al.'s analyses supported models with variable non–decision time processes, and in particular those where the variability increased with the mean. An interesting avenue for future research would be to extend the various model comparisons we have made (above) to models with sophisticated random distributions for non–decision time.

Our approach to linking neural and behavioral data is not unique, and is not necessarily the best for many different situations. Still, we propose that, when possible, researchers should strive for the tightest possible linking, as this provides the greatest opportunity to investigate the underlying linking assumptions. Put differently, joint models that are tightly linked allow us to uncover underlying psychological processes that simultaneously explain behavioral and neural data. Such an approach is arguably more powerful than more loosely linked models, which often do not go beyond correlations between different levels of data. The "when possible" caveat relates to the state–of–the–art in the cognitive modeling and neural modeling of the research field in question. Very tightly–linked models, with explicit and quantitative linking assumptions, are only possible in research fields with tractable quantitative models for both behavioral and neural data.

One of the more interesting implications of joint modeling is trying to relate two streams of data with potentially vastly different scales. In our particular example, we are combining RT data in seconds (from 0 to 7), proportion correct (from 0 to 1), and mean amplitudes at Pz (90% fall in the range -7.5 to 11). Revealing how the different kinds of model misfit interact, and how much influence they have relative to one another, is another of the strengths of the joint modeling approach we have used. Direct comparison of the fits of model that vary in just one component can be very revealing about that component (conditional on the other model components being reasonable, of course). For instance, our models ν–*ERP* and t_0–*ERP* differed only on the neural linking component, whereas t_0–*ERP* and *Brev*–t_0–*ERP* differed only on the behavioral component, and these pairings allowed us to investigate interesting psychological questions.

Our example model demonstrates a very tight link between neural and behavioral data in the field of simple decision–making. This field has seen some excellent interdisciplinary work between neuroscience and psychology (e.g. Purcell et al., 2010). The example model we developed shows how to adapt the LBA model for response time data to incorporate ERP data, recorded during a mental rotation task. The result is a joint model that can simultaneously capture characteristics of data at the behavioral level (response times and choice proportions) and the neural level (ERPs). Approaches like this are very exciting, because they help to reduce barriers between two fields that have operated alongside another rather than together for a long time. We hope that modeling data at different levels with a single set of parameters paves the way to a more integrative cognitive science.

References

Anderson, J. R., & Fincham, J. M. (2014). Discovering the sequential structure of thought. *Cognitive Science*, *38*, 322–352.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. *Psychology of Learning and Motivation*, *2*, 89–195.

Belliveau, J., Kennedy, D., McKinstry, R., Buchbinder, B., Weisskoff, R., Cohen, M., ... Rosen, B. (1991). Functional mapping of the human visual cortex by magnetic resonance imaging. *Science*, *254*, 716–719.

Brenner, D., Williamson, S., & Kaufman, L. (1975). Visually evoked magnetic fields of the human brain. *Science*, *190*, 480–482.

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

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

Gamerman, D., & Lopes, H. F. (2006). *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference.* Boca Raton, FL: Chapman & Hall/CRC.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (Eds.). (1996). *Markov chain Monte Carlo in practice.* Boca Raton (FL): Chapman & Hall/CRC.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, *30*, 535–574.

Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science*, *274*, 427–430.

Heil, M. (2002). The functional significance of ERP effects during mental rotation. *Psychophysiology*, *43*, 401–412.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science*, *182*, 177–180.

Ho, T. C., Brown, S., Abuyo, N. A., Ku, E.-H. J., & Serences, J. T. (2012). Perceptual consequences of feature-based attentional enhancement and suppression. *Journal of vision*, *12*, 1–17.

Ho, T. C., Brown, S., & Serences, J. T. (2009). Domain general mechanisms of perceptual decision making in human cortex. *The Journal of Neuroscience*, *29*, 8675–8687.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course.* Cambridge University Press. Retrieved from `http://books.google.com.au/books?id=Gq6kAgAAQBAJ`

Lewandowsky, S., & Farrell, S. (2010). *Computational modeling in cognition: Principles and practice.* Sage.

Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, *40*, 77–105.

Luce, R. D. (1986). *Response times.* New York: Oxford University Press.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar–based random walk model of speeded classification. *Psychological Review*, *104*, 266–300.

Pouget, P., Logan, G. D., Palmeri, T. J., Boucher, L., Paré, M., & Schall, J. D. (2011). Neural basis of adaptive response time adjustment during saccade countermanding. *The Journal of Neuroscience*, *31*, 12604–12612.

Provost, A., & Heathcote, A. (2015). Titrating decision processes in the mental rotation task. *Psychological Review*, *122*, 735–754.

Provost, A., Johnson, B., Karayanidis, F., Brown, S. D., & Heathcote, A. (2013). Two routes to expertise in mental rotation. *Cognitive Science*, *37*, 1321–1342.

Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., & Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychological Review*, *117*, 1113–1143.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*, 93–134.

Rae, B., Heathcote, A., Donkin, C., Averell, L., & Brown, S. D. (2014). The hare and the tortoise: Emphasizing speed can change the evidence used to make decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 1226–1243.

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

Riečanský, I., & Jagla, F. (2008). Linking performance with brain potentials: Mental rotation-related negativity revisited. *Neuropsychologia*, *46*, 3069–3073.

Schall, J. D. (2001). Neural basis of deciding, choosing, and acting. *Nature Reviews Neuroscience*, *2*, 33–42.

Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: Seeing and deciding. *Proceedings of the National Academy of Sciences*, *93*, 628–633.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three–dimensional objects. *Science*, *171*, 701–703.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B*, *64*, 583–639.

Sutton, S., Braren, M., Zubin, J., & John, E. (1965). Evoked–potential correlates of stimulus uncertainty. *Science*, *150*, 1187–1188.

Teller, D. (1984). Linking propositions. *Vision Research*, *24*, 1233–1246.

Trueblood, J. S., Brown, S. D., & Heathcote, A. (2014). The multiattribute linear ballistic accumulator model of context effects in multialternative choice. *Psychological Review*, *121*, 179–205.

Turner, B. M., Forstmann, B. U., Love, B. C., Palmeri, T. J., & van Maanen, L. (in press). Approaches to analysis in model–based cognitive neuroscience. *Journal of Mathematical Psychology*.

Turner, B. M., Forstmann, B. U., Wagenmakers, E.-J., Brown, S. D., Sederberg, P. B., & Steyvers, M. (2013). A Bayesian framework for simultaneously modeling neural and behavioral data. *Neuroimage*, *72*, 193–206.

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

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

van Ravenzwaaij, D., Cassey, P., & Brown, S. D. (in press). A simple introduction to Markov Chain Monte–Carlo. *Psychonomic Bulletin & Review*.

Wagenmakers, E.-J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology*, *21*, 641–671.