

Comparing models of contrast gain using psychophysical experiments

Christopher DiMattina

Computational Perception Laboratory,
Department of Psychology, Florida Gulf Coast University,
Fort Myers, FL, USA



In a wide variety of neural systems, neurons tuned to a primary dimension of interest often have responses that are modulated in a multiplicative manner by other features such as stimulus intensity or contrast. In this methodological study, we present a demonstration that it is possible to use psychophysical experiments to compare competing hypotheses of multiplicative gain modulation in a neural population, using the specific example of contrast gain modulation in orientation-tuned visual neurons. We demonstrate that fitting biologically interpretable models to psychophysical data yields physiologically accurate estimates of contrast tuning parameters and allows us to compare competing hypotheses of contrast tuning. We demonstrate a powerful methodology for comparing competing neural models using adaptively generated psychophysical stimuli and demonstrate that such stimuli can be highly effective for distinguishing qualitatively similar hypotheses. We relate our work to the growing body of literature that uses fits of neural models to behavioral data to gain insight into neural coding and suggest directions for future research.

Introduction

A large body of experimental and theoretical work has quantitatively analyzed the relationship between neural codes, perception, and perceptual decisions (Nienborg, Cohen, & Cumming, 2012; Parker & Newsome, 1998; Romo & de Lafuente, 2013). Typically, these studies use physiological data to explain behavior by correlating neural performance with behavioral performance (e.g., Britten, Shadlen, Newsome, & Movshon, 1992; Cohen & Newsome, 2009; Egger & Britten, 2013; Vogels & Orban, 1990; L. Wang, Narayan, Graña, Shamir, & Sen, 2007) or by using the responses of a neural population to predict behavior (e.g., Bollimunta, Totten, & Ditterich, 2012; Kiani, Cueva, Reppas, & Newsome, 2014). However, in recent

years, an ever-growing body of literature (reviewed in the Discussion) has taken a complementary approach by making use of behavioral data or theoretically optimal performance on well-defined behavioral tasks to inform and connect with models of neural encoding. This work has demonstrated that quantitatively characterizing behavioral data using neurally plausible models can yield insight into sensory receptive field properties (e.g., Burge & Geisler, 2014, 2015; W. S. Geisler, Najemnik, & Ing, 2009; Neri & Levi, 2006; Yamins et al., 2014), pooling of neural population responses (e.g., Goris, Putzeys, Wagemans, & Wichmann, 2013; Morgenstern & Elder, 2012), attentional modulation (e.g., Murray, Sekuler, & Bennett, 2003; Neri, 2004; Pestilli, Carrasco, Heeger, & Gardner, 2011; Pestilli, Ling, & Carrasco, 2009), perceptual learning (e.g., Petrov, Doshier, & Lu, 2005), and near-optimal performance in perceptual tasks (e.g., Ma, Navalpakkam, Beck, Van Den Berg, & Pouget, 2011; Qamar et al., 2013).

In this paper, we extend this growing body of literature by presenting a general methodology for using data obtained in psychophysical experiments to characterize contrast gain modulation in sensory neural populations. Although we focus on contrast gain in early vision, many sensory neural populations tuned to parameters of primary interest (tactile orientation, auditory frequency, etc.) also exhibit response modulation by stimulus amplitude or contrast (Barbour & Wang, 2003; Bensmaia, Denchev, Dammann, Craig, & Hsiao, 2008; Kiang, 1965; Muniak, Ray, Hsiao, Dammann, & Bensmaia, 2007; Sachs & Abbas, 1974; Sadagopan & Wang, 2008). We apply this methodology in real psychophysical experiments to analyze a simple model of orientation decoding from a population of contrast- and orientation-tuned neurons in order to demonstrate how psychophysical data may be used to (a) accurately recover neural encoding model parameters and (b) compare competing hypotheses of neural encoding. In particular, we demonstrate that we can

Citation: DiMattina, C. (2016). Comparing models of contrast gain using psychophysical experiments. *Journal of Vision*, 16(9):1, 1–18, doi:10.1167/16.9.1.

doi: 10.1167/16.9.1

Received November 14, 2015; published July 1, 2016

ISSN 1534-7362

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.



Downloaded From: <http://jov.arvojournals.org/pdfaccess.aspx?url=/data/Journals/JOV/935414/> on 07/05/2016

use psychophysical data to correctly infer the physiologically measured values of contrast gain function parameters in visual neurons (Albrecht & Hamilton, 1982). We further demonstrate experimentally that adaptive stimulus optimization methods that have recently gained traction in brain and cognitive science (e.g., Cavagnaro, Myung, Pitt, & Kujala, 2010; DiMattina, 2015; DiMattina & Zhang, 2013; Lewi, Butera, & Paninski, 2009; Myung, Cavagnaro, & Pitt, 2013; Paninski, Pillow, & Lewi, 2007; Z. Wang & Simoncelli, 2008) can be used to find psychophysical stimuli during the course of the experimental session, which are optimized for distinguishing competing hypotheses of neural coding. We find that presenting stimuli adaptively optimized for model comparison may in some cases be very helpful for discriminating between qualitatively similar hypotheses of neural encoding. We discuss the limitations of the present methodology and suggest interesting directions for future research. We believe that with further developments of biologically motivated approaches to modeling psychophysical data, psychophysical experiments can more directly inform investigations of neural encoding.

Methods and results

Defining biologically interpretable psychometric models

Here we present for didactic purposes a derivation of the psychometric function that makes explicit the fact that perceptual behavior is ultimately dependent on the parameters of the sensory neural population used to guide that behavior.

For a population of N neurons, a *neural encoding* model $P(\mathbf{r}|\mathbf{s}, \theta)$ specifies the probability of observing neural responses $\mathbf{r} = (r_1, \dots, r_N)^T$ as a function of stimulus parameters $\mathbf{s}$ and neuronal population parameters θ (Borst & Theunissen, 1999; Paninski et al., 2007). Perhaps the simplest possible neural encoding model is a set of tuning curves specifying the expected firing rate of each neuron in the population as a function of the sensory variable $\mathbf{s}$, for instance, the orientation-tuning curves shown in Figure 1a. In this case, the population parameters θ would represent the properties of this set of tuning curves, for instance, the centers $\mu_1, \dots, \mu_N$, tuning curve width σ , and amplitude A . Similarly, a *neural decoding* model $P(\mathbf{s}|\mathbf{r}, \omega)$ specifies the probability of a stimulus $\mathbf{s}$ being present as a function of the observed neural responses $\mathbf{r}$ and possibly additional parameters ω (Paninski et al., 2007).

We define a *behavioral decoding* model $P(b|\mathbf{r}, \omega)$ as specifying the probability of a behavioral response b as a function of neural responses $\mathbf{r}$ as well as additional decoding parameters ω . In this formulation, the stochastic neural responses $\mathbf{r}$, which is the output of the neural encoding model, serves as the input to the behavioral decoding model, as illustrated in Figure 1b and c. The behavioral decoding model may deterministically specify b as a function of $\mathbf{r}$, ω , as in the example shown in Figure 1b, which compares the decision variable $u = \sum_{i=1}^N \omega_i r_i$ to a fixed decision threshold. Alternatively, the behavioral decoding model may also specify b probabilistically in order to model stimulus-independent “decision noise” (Shadlen, Britten, Newsome, & Movshon, 1996). The joint probability of observing a behavior b and neural response $\mathbf{r}$ as a function of a stimulus $\mathbf{s}$ may be written as the product of a neural encoding model and behavioral decoding model using the basic probability law $P(A, B) = P(A|B)P(B)$ (Bishop, 2006), yielding the expression

$$P(b, \mathbf{r}|\mathbf{s}, \omega, \theta) = P(b|\mathbf{r}, \omega)P(\mathbf{r}|\mathbf{s}, \theta). \quad (1)$$

By marginalizing the joint probability $P(b, \mathbf{r}|\mathbf{s}, \omega, \theta)$ over $\mathbf{r}$, we can express the probability of a behavior entirely as a function of the stimulus parameters $\mathbf{s}$ and model parameters θ , ω without any dependence on unobserved neural responses. This follows from the basic probability law $\int P(A, B) dB = P(A)$. Marginalizing Equation 1 over $\mathbf{r}$ yields the equation

$$P(b|\mathbf{s}, \theta, \omega) = \int P(b|\mathbf{r}, \omega)P(\mathbf{r}|\mathbf{s}, \theta) d\mathbf{r}. \quad (2)$$

Note that the integrand in Equation 2 is the product of the behavioral decoding model and the neural encoding model integrated over all possible neural responses $\mathbf{r}$ conditioned on the stimulus $\mathbf{s}$. In the case of fixed decoding model parameters $\hat{\omega}$, so that $P(\omega) = \delta(\omega - \hat{\omega})$ (where δ denotes the Dirac delta function), we can use Equation 2 to derive an expression for the posterior probability of the neural encoding model parameters θ given only psychophysical trial data $\mathcal{D}_\psi = \{\mathbf{s}_i, b_i\}_{i=1}^n$ (Appendix A):

$$P(\theta|\mathcal{D}_\psi) = \frac{1}{Z} \left[\prod_{i=1}^n \int P(b_i|\mathbf{r}, \hat{\omega})P(\mathbf{r}|\mathbf{s}_i, \theta) d\mathbf{r} \right] P(\theta). \quad (3)$$

In the case in which one does not make informative prior assumptions $P(\theta)$ about the neural encoding model parameters, Equation 3 becomes the likelihood. In the application presented in this study, we do not incorporate informative priors on θ and simply attain maximum likelihood point estimates.

Although Equations 2 and 3 make explicit the dependence of the psychometric function on the neural encoding model and show that one can in principle estimate neural parameters from behavior, these

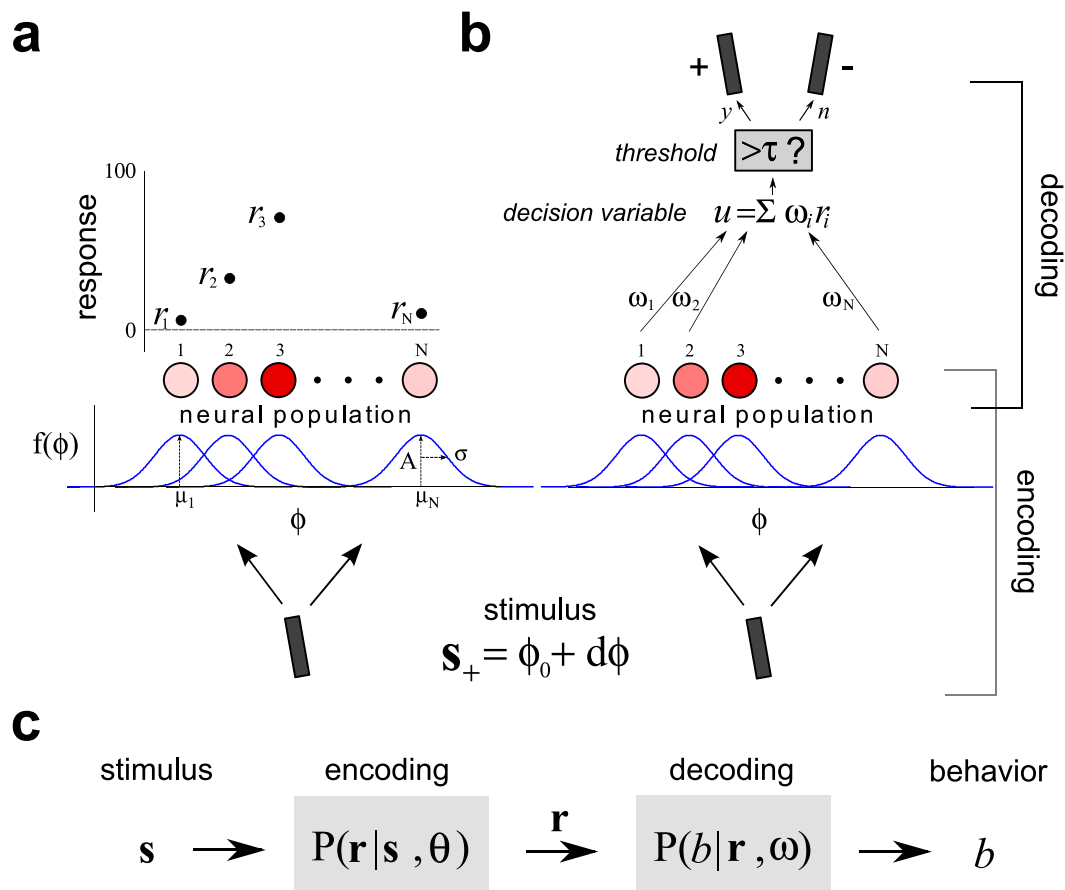


Figure 1. Schematic illustration of neural encoding and behavioral decoding models. (a) A neural encoding model $P(\mathbf{r}|\mathbf{s}, \theta)$ specifies the probability of observing stimulus-dependent neural population responses $\mathbf{r}$. Bottom: An oriented bar stimulus elicits noisy responses from orientation-tuned neurons whose tuning curves are specified by parameters $\theta = (A, \sigma, \mu_1, \dots, \mu_N)^T$. Top: Observed noisy single-trial responses $\mathbf{r} = (r_1, r_2, \dots, r_N)^T$ of each neuron. (b) A behavioral decoding model takes as input the stimulus-evoked neural responses $\mathbf{r} = (r_1, r_2, \dots, r_N)^T$ and uses them to determine the probability of a behavior b . In the deterministic model shown here, neural responses $\mathbf{r}$ are multiplied by weights $\omega = (\omega_1, \dots, \omega_N)^T$ and summed to form a decision variable ($u = \sum \omega_i r_i$), which is compared to a threshold (τ) to predict a binary perceptual decision. (c) One can define a biologically interpretable psychometric function by using the output $\mathbf{r}$ of a neural encoding model as the input to a behavioral decoding model.

equations are of little practical use without specific assumptions about the neural encoding, i.e., $P(\mathbf{r}|\mathbf{s}, \theta)$, or behavioral decoding, i.e., $P(b|\mathbf{r}, \omega)$ models. Even with such assumptions, one must be aware that there are practical limitations on the number of neuronal parameters θ that can be accurately estimated during the course of a psychophysical experiment. As studies with classification images show (Ahumada, 1996; Eckstein & Ahumada, 2002; Mineault, Barthelmé, & Pack, 2009; Murray, 2011), binomial (e.g., yes/no) responses provide relatively little information per trial, necessitating a large number of trials to attain accurate estimates of the perceptual filter. However, we demonstrate here that it is very realistic to use psychophysical data to estimate and compare low-dimensional analytical models (e.g., May and Solomon, 2015a,

2015b; Pestilli et al., 2011; Pestilli et al., 2009) in a process of focused hypothesis testing.

Orientation discrimination model

We now consider the application of our modeling framework to a simple orientation discrimination task, in which a subject has to determine in which of two directions (clockwise: –, counterclockwise: +) a sinusoidal grating stimulus having contrast c ($0 \leq c \leq 100\%$) has been tilted (by $d\phi^\circ$) with respect to vertical. In order to do this, we must specify concretely the hypothesized neural code $\mathbf{r}$, the observable behaviors b , the hypothesized neural encoding model $P(\mathbf{r}|\mathbf{s}, \theta)$, and the hypothesized behavioral decoding model $P(b|\mathbf{r}, \omega)$.

A fairly straightforward derivation of a psychometric function defined using the neural encoding model shown in Figure 1 and with linear decoding (Fisher linear discriminant) is given in Appendix B. This analysis is similar to those presented in several previous studies (e.g., Ma, 2010; Pestilli et al., 2009). Our derivation yields the final model

$$P(b = 1 | \mathbf{s}, \boldsymbol{\theta}, \boldsymbol{\omega}) = \Phi\left(K\sqrt{\psi(c)}d\phi\right), \quad (4)$$

where $\psi(c)$ denotes the contrast tuning (also called *contrast gain*) of neurons in the population, and $K \propto \sqrt{I_F(\phi_0)}$ is a parameter describing population sensitivity to changes in orientation around the vertical reference ($\phi_0 = \pi/2$) at 100% contrast.

In this paper, we consider three different functional forms for the contrast gain function $\psi(c)$. One form suggested from neurophysiological findings (Albrecht & Hamilton, 1982) is the Naka-Rushton function

$$\psi_1(c, \boldsymbol{\eta}^{(1)}) = \frac{c^n}{c^n + c_{50}^n}, \quad (5)$$

having parameters $\boldsymbol{\eta}^{(1)} = (n, c_{50})^T$. This functional form (Equation 5) is also sometimes referred to as the hyperbolic ratio function (Albrecht & Hamilton, 1982). Another form is the hyperbolic tangent (tanh) function

$$\psi_2(c, \boldsymbol{\eta}^{(2)}) = \tanh(bc) = \frac{e^{bc} - e^{-bc}}{e^{bc} + e^{-bc}}, \quad (6)$$

commonly used in machine learning (Bishop, 2006), having parameter $\boldsymbol{\eta}^{(2)} = (b)^T$. Both of these functional forms (Naka-Rushton, Tanh) are shown in Figure 2. Finally, we consider a Gaussian form that allows for the possibility of a nonmonotonic relationship between contrast and firing rate, given by

$$\psi_3(c, \boldsymbol{\eta}^{(3)}) = \exp\left(-\frac{1}{2\sigma^2}(c - \mu)^2\right), \quad (7)$$

with parameters $\boldsymbol{\eta}^{(3)} = (\mu, \sigma)^T$.

Our interest in fitting multiple models to the same data set is to test the efficacy of psychophysical data for distinguishing between competing hypotheses of neural encoding. This approach follows previous work using fits of multiple models to behavioral data to gain insight into sensory or cognitive mechanisms (Qamar et al., 2013; van den Berg, Awh, & Ma, 2014). The comparison between the Naka-Rushton model and the Gaussian model is a coarse-grained *qualitative* comparison because the two models are qualitatively very different (monotonic vs. nonmonotonic) whereas the comparison between Naka-Rushton and Tanh is a fine-grained *quantitative* comparison because the two models are both monotonic and qualitatively very similar (Figure 2).

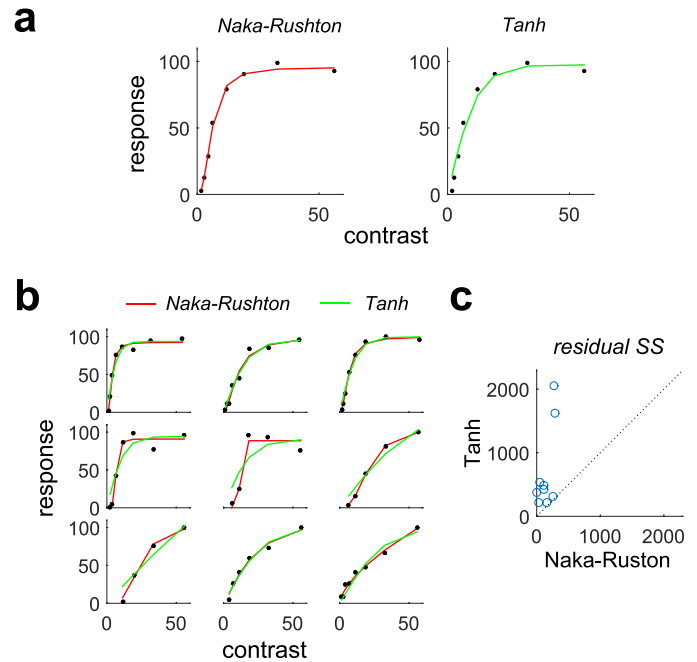


Figure 2. Two competing hypotheses for the functional form of contrast gain tuning. Despite the qualitative similarity of the Naka-Rushton (Equation 5) and Tanh (Equation 6) models, we observe a better quantitative fit to neurophysiological data by the Naka-Rushton function, particularly at lower contrasts. (a) Fits of both models (Equations 5 and 6) to contrast gain responses of a representative V1 neuron. Data points graphically adapted from figure 3 of Albrecht and Hamilton (1982). (b) Fits of both models (Equations 5 and 6) to contrast gain responses of several V1 neurons. Data points graphically adapted from figure 1 of Albrecht and Hamilton (1982). (c) Residual sum-of-squares error for the fits of both models in (b). We see a better fit for the Naka-Rushton model (sign-rank test, $n = 9$, $p = 0.0039 < 0.01$).

Recovering neural encoding model parameters

Fitting thresholds

Because we can write our psychometric function (Equation 4) in terms of d' (25), we can use thresholds taken at multiple contrasts to estimate the psychometric function parameters using least-squares curve fitting. Figure 3 shows the best fit of the model (Equation 4) with Naka-Rushton contrast gain (Equation 5) to the data from Skottun, Bradley, Sclar, Ohzawa, and Freeman (1987; their figure 1). We see in Figure 3 that this model provides an excellent fit to their data (Supplementary Figure S1). We find that the values recovered for the Naka-Rushton contrast function parameters n , c_{50} from their threshold data lie within the range measured in previous neurophysiological work (Albrecht & Hamilton, 1982) as shown in Figure 4 (red circles).

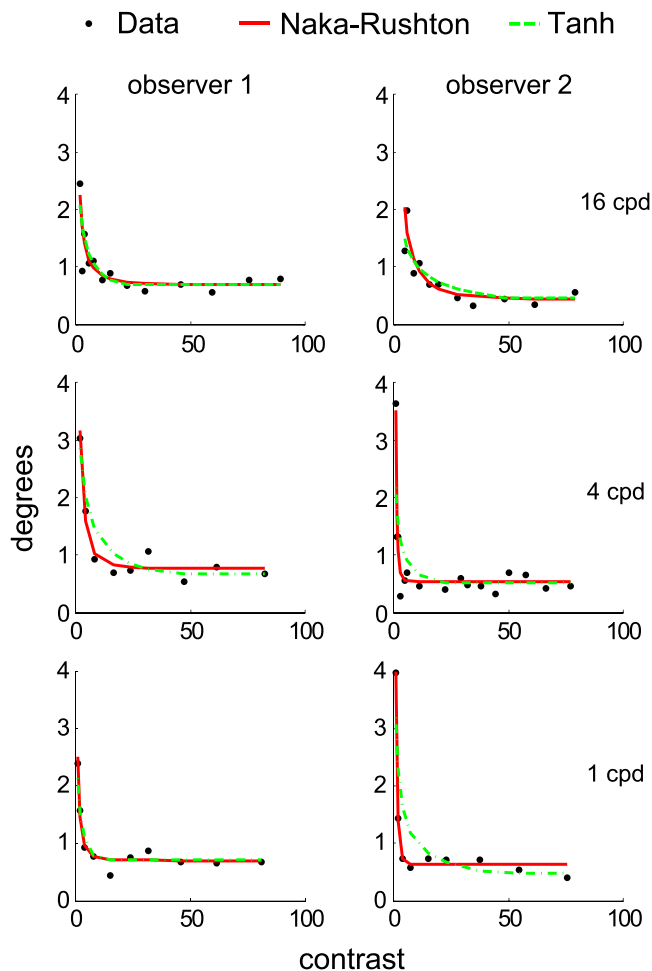


Figure 3. Fits of the behavioral decoding model (Equation 4) with Naka-Rushton contrast gain (Equation 5) to threshold data (79% performance) graphically adapted from figure 1 of Skottun et al. (1987). Plot of residual sum-of-squares error for models with Naka-Rushton (red) and Tanh (green) contrast gain (Equation 6) are given in Supplementary Figure S1.

Experiment 1: Direct estimation from psychophysical trials

The data in Skottun et al. (1987) only provides thresholds, and therefore our estimates of $\eta^{(1)} = (n, c_{50})^T$ were not obtained as in most psychophysical experiments, in which one finds the maximum likelihood estimate of model parameters using stimulus-response data $\mathcal{D}_\psi = \{s_i, b_i\}_{i=1}^n$ (Kingdom & Prins, 2010). In order to directly test the use of psychophysical data to recover the parameters of neural tuning curves, we ran an orientation discrimination experiment (Experiment 1) on nine subjects (seven naive) in which we covaried orientation and contrast. Additional details of Experiment 1 are described in the Supplementary Methods. Contour plots of subject performance $P(b = 1 | s = (c, d\phi)^T, K, \eta^{(1)})$ are shown in Figure 5 (and Supplementary Figure S3) with fits of the model (Equation 4) with Naka-Rushton gain (Equa-

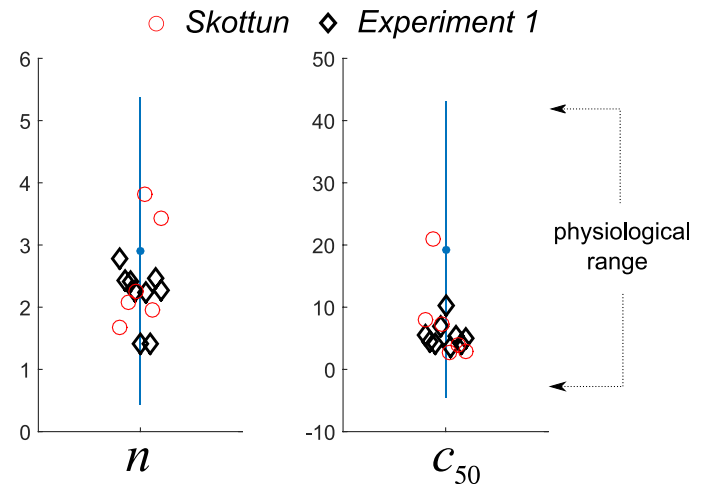


Figure 4. Estimates of neural contrast gain function parameters n and c_{50} (Naka-Rushton) from psychophysical data. Red dots denote estimates from threshold data (Skottun et al., 1987), black diamonds are estimates from fitting the model directly to psychophysical trial data (Experiment 1). We see that all of the estimates lie within the physiological range (blue lines = $\mu \pm 1.96 \cdot \sigma$) (Albrecht & Hamilton, 1982).

tion 5) to subject data in the middle column. We found in a subsequent experiment (Supplementary Material) that this model could also generalize reasonably well for most (but not all) subjects to predict responses to a small validation set of novel stimuli (Supplementary Figure S4).

We see in Figure 4 that the values of the Naka-Rushton parameters n , c_{50} estimated from our Experiment 1 data (black diamonds) lie within the neurophysiologically observed range. Numerical values of these parameters are given in Supplementary Tables S1 and S2. Interestingly, we find that all of our estimates of the half-saturation parameter c_{50} obtained in these experiments (along with five of six estimates of c_{50} from Skottun et al., 1987) lie toward the lower end of the physiologically observed range (i.e., around 5% contrast; see Albrecht & Hamilton, 1982). This suggests the subjects may be using the neurons that are most sensitive to contrast when they perform the task, consistent with the “lower envelope” principle of sensory coding (Egger & Britten, 2013; Mountcastle, LaMotte, & Carli, 1972; L. Wang et al., 2007).

Comparing competing models

Exploring model space

In Experiment 1, whose goal was to show that one can estimate neural model parameters from psychophysical data, we assumed a known form (Equation 5) of the contrast gain function based on previous

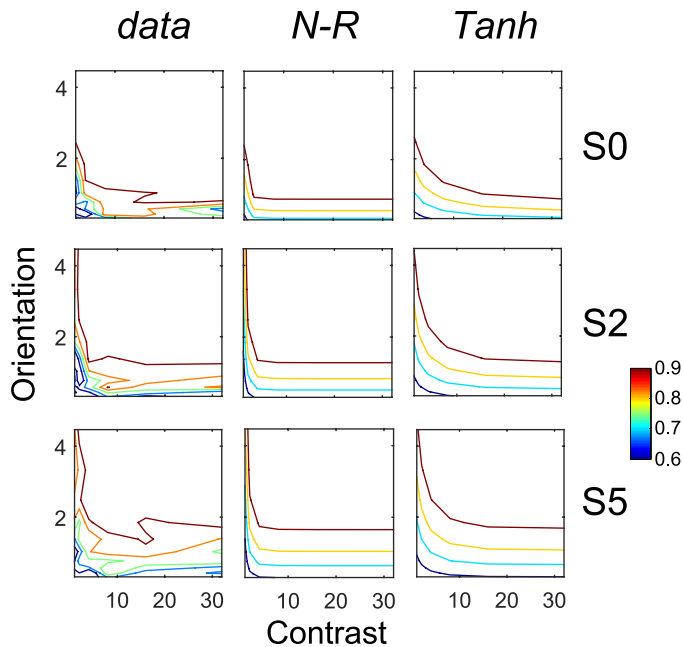


Figure 5. Contour plots of the psychometric function $P(b=1|\mathbf{s}=(c, d\phi)^T)$ as a function of orientation ($d\phi$) and contrast (c) for three subjects in Experiment 1. Other subjects shown in Supplementary Figure S3. Left: Raw data. Middle: Fits of model (Equation 4) with Naka-Rushton (Equation 5) contrast gain (model 1) to data. Right: Fits of (Equation 4) with Tanh (Equation 6) contrast gain (model 2) to data.

neurophysiological investigations (Albrecht & Hamilton, 1982). Supposing that the correct functional form of the contrast gain function $\psi(c)$ was not known beforehand from physiological recordings, we may be interested in evaluating various possibilities by fitting the model (Equation 4) to psychophysical data with different choices for $\psi(c)$ and seeing which best accounts for the observed results. Such information derived from relatively fast and inexpensive psychophysical experiments could provide important clues to guide subsequent neurophysiology research.

In order to test the ability of psychophysical experiments to compare competing models of neural contrast gain, we will also consider two other possibilities for the contrast gain, given by the hyperbolic tangent (Tanh) function (Equation 6) and the familiar Gaussian tuning curve (Equation 7). These three possible choices (Equations 5, 6, and 7) of contrast gain function $\psi(c)$ define a discrete space of three competing neural encoding models, which we index by $i = 1, 2, 3$. By fitting each model to psychophysical data, we may evaluate their relative likelihoods using the Akaike Information Criterion (AIC), which measures goodness-of-fit while penalizing model complexity (Akaike, 1974; Burnham & Anderson, 2003). Previous work has shown that it is important that any model comparison method takes

complexity into account because an overly complex model often fits training data well but fails to generalize to novel observations (Bishop, 2006; Pitt & Myung, 2002).

We denote the value of the AIC for the i -th model by AIC_i , with model i being preferred to model j if $AIC_i > AIC_j$. We define a *model preference index*

$$P_{i-j} = AIC_i - AIC_j, \quad (8)$$

where a positive value of P_{i-j} indicates model i is preferred to model j , and a negative value indicating j is preferred to i . The model preference index is defined implicitly with respect to a fixed number of observations, i.e., $P_{i-j} = P_{i-j}(n)$, where n is the number of trials used to compute the AIC. We define a *change in model preference* after k additional trials as

$$\Delta P_{i-j} = P_{i-j}(n+k) - P_{i-j}(n). \quad (9)$$

In our analysis, model 1 assumes Naka-Rushton contrast tuning (Equation 5), model 2 assumes Tanh tuning (Equation 6), and model 3 assumes Gaussian tuning (Equation 7).

Computing the AIC for fits of all three models to the data collected in Experiment 1 allows us to determine the model preferences P_{1-2} (Naka-Rushton–Tanh) and P_{1-3} (Naka-Rushton–Gaussian). We see in Figure 6a that the Naka-Rushton model is preferred over the Gaussian model for all nine subjects and over the Tanh model for seven of nine subjects, with the preference being quite strong for many subjects. Statistical tests show that over these nine subjects, both model preferences are significantly different from zero (sign-rank test, $n=9$; $P_{1-2} > 0$: $p = 0.02$, $P_{1-3} > 0$: $p = 0.004$). Figure 6b shows how this model preference P_{1-2} evolves with the number of experimental trials. We see that, as more trials are collected, the model preference (for most subjects) seems to change in favor of the Naka-Rushton model, whose better ability to fit the data overcomes the complexity penalty imposed by the AIC. We also see from Figure 6b that the final model preferences are established after about 1,000–1,200 trials. Similar results were obtained using the Bayes Information Criterion, which more severely penalizes model complexity (Bishop, 2006), changing the final model preference for only one subject (Supplementary Figures S5 and S6).

Experiment 2: Optimizing stimuli for model comparison

In Experiment 1, data was collected using the method of constant stimuli, which previous work has suggested may be suboptimal for purposes of model estimation and comparison (Watson & Fitzhugh, 1990). Therefore, we conducted a second experiment (Experiment 2) in order to determine if stimuli

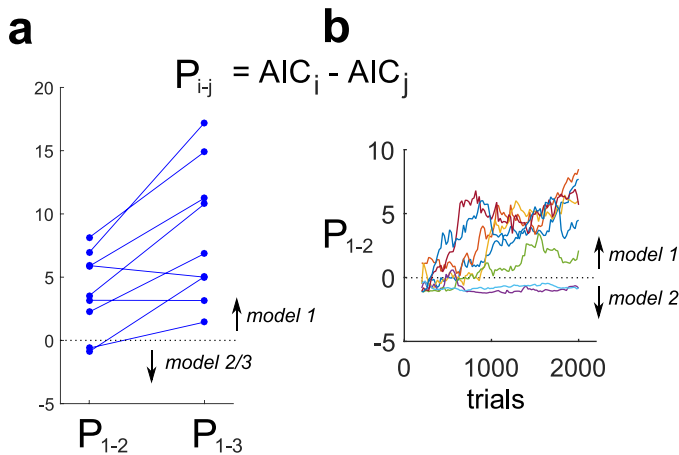


Figure 6. Model preferences P_{i-j} based on fits of three competing neural encoding models to data from Experiment 1. Model 1 assumes Naka-Rushton (Equation 5) contrast gain, model 2 assumes Tanh (Equation 6) contrast gain, and model 3 assumes Gaussian (Equation 7) contrast gain. (a) Final model preferences P_{1-2} and P_{1-3} based on fits to all Experiment 1 trials. For most subjects, we see a final preference ($P_{1-2} > 0$) for model 1 (Naka-Rushton) over model 2, and for all subjects, we see a preference ($P_{1-3} > 0$) for model 1 over model 3. (b) Dynamics of model preference P_{1-2} for the two qualitatively similar models (Naka-Rushton–Tanh) for the $n = 8$ subjects completing 2,000+ trials. Final model preferences are established by $\sim 1,000$ trials.

explicitly optimized for purposes of model comparison were more effective for this goal than the stimuli used in Experiment 1 (Supplementary Figure S2).

There are several ways to define the optimal comparison stimulus (OCS) in neurophysiology and psychophysics experiments (Cavagnaro et al., 2010; DiMattina & Zhang, 2011; Z. Wang & Simoncelli, 2008), and in the current study, we used an information-theoretic criterion that finds the stimulus that minimizes the expected entropy of the posterior density over model space (Cavagnaro et al., 2010). This stimulus $\mathbf{s} = (c, d\phi)^T$ may be found by maximizing the expression

$$U^{(C)}(\mathbf{s}) = \sum_{i=1}^m P_0(i) D_{\text{KL}}[p(b|\mathbf{s}, i), p(b|\mathbf{s})], \quad (10)$$

where $P_0(i)$ is the prior probability of each model, D_{KL} the Kullback-Liebert divergence (Cover & Thomas, 2006), $p(b|\mathbf{s}, i)$ is the response probability conditioned on the stimulus and model, and $p(b|\mathbf{s})$ is the overall response probability averaged across models. Intuitively, this method minimizes uncertainty about which model is true by presenting stimuli that are expected to yield a posterior density with most of the probability mass on one or a few models, i.e., a density with minimum entropy (Cover & Thomas, 2006). This

information-theoretic criterion has been used in cognitive science to choose stimuli optimized for testing competing hypotheses of memory decay and decision making under risk (Cavagnaro, Gonzalez, Myung, & Pitt, 2013; Cavagnaro, Pitt, & Myung, 2011).

Data was obtained during a two-phase experiment conducted on a single testing day: an estimation phase (E-phase, Experiment 1) in which data is collected for model-fitting purposes, followed by a comparison phase (C-phase, Experiment 2) in which stimuli optimized for model discrimination were presented (DiMattina & Zhang, 2011). Immediately after the conclusion of Experiment 1 (E-phase, $N_E = 1,200$ trials) a single OCS was found by optimizing (Equation 10) based on fits of model 1 (Naka-Rushton) and model 2 (Tanh) to Experiment 1 data. Search for the OCS was restricted to contrasts greater than 1% and orientations from 0° to 20° , based on observation of at what point the two models seemed to differ the most as well as the fact that stimuli presented at values less than 1% contrast are often barely visible (Campbell & Robson, 1968). The OCS for each subject are illustrated in Figure 7 (left panels). Note that many of these stimuli have contrast $c \approx 1$ and orientation $d\phi > 5^\circ$ and hence lie outside the range of stimuli (contrasts and orientations) used to estimate the models (Supplementary Figures S2 and S4).

In Experiment 2, the OCS was repeatedly presented to the subject for $N_C = 200$ trials during the Experiment 2 C-phase, interleaved with 200 stimuli chosen at random with uniform probability from the stimulus grid used during the Experiment 1 (Supplementary Figure S2) for 400 trials total. We will heretofore refer to these randomly chosen Experiment 1 (E-phase) stimuli as IID stimuli. We see from Figure 7 (right panels) that for many (but not all) subjects the OCS (blue curves) does a much better job than the IID stimuli (green curves) of shifting the model preference P_{1-2} in the direction of the Naka-Rushton model, $\Delta P_{1-2} = P_{1-2}(N_E + N_C) - P_{1-2}(N_E) > 0$. Statistical analysis demonstrates that over all subjects, the median value of ΔP_{1-2} is significantly larger for the OCS (median $\Delta P_{1-2} = 5.41$) than IID (median $\Delta P_{1-2} = -0.04$) trials (sign-rank test, $n = 9$, $p = 0.0117$).

Our goal in Experiment 2 was not to do an in-depth investigation of adaptive stimulus optimization methods for model comparison (a very important problem needing more research) but rather to demonstrate the potential utility of such an approach. Our results suggest that utilizing stimuli optimized for neural encoding model comparison is certainly no worse, and in many cases much better, than continued presentation of the (IID) stimuli used in Experiment 1.

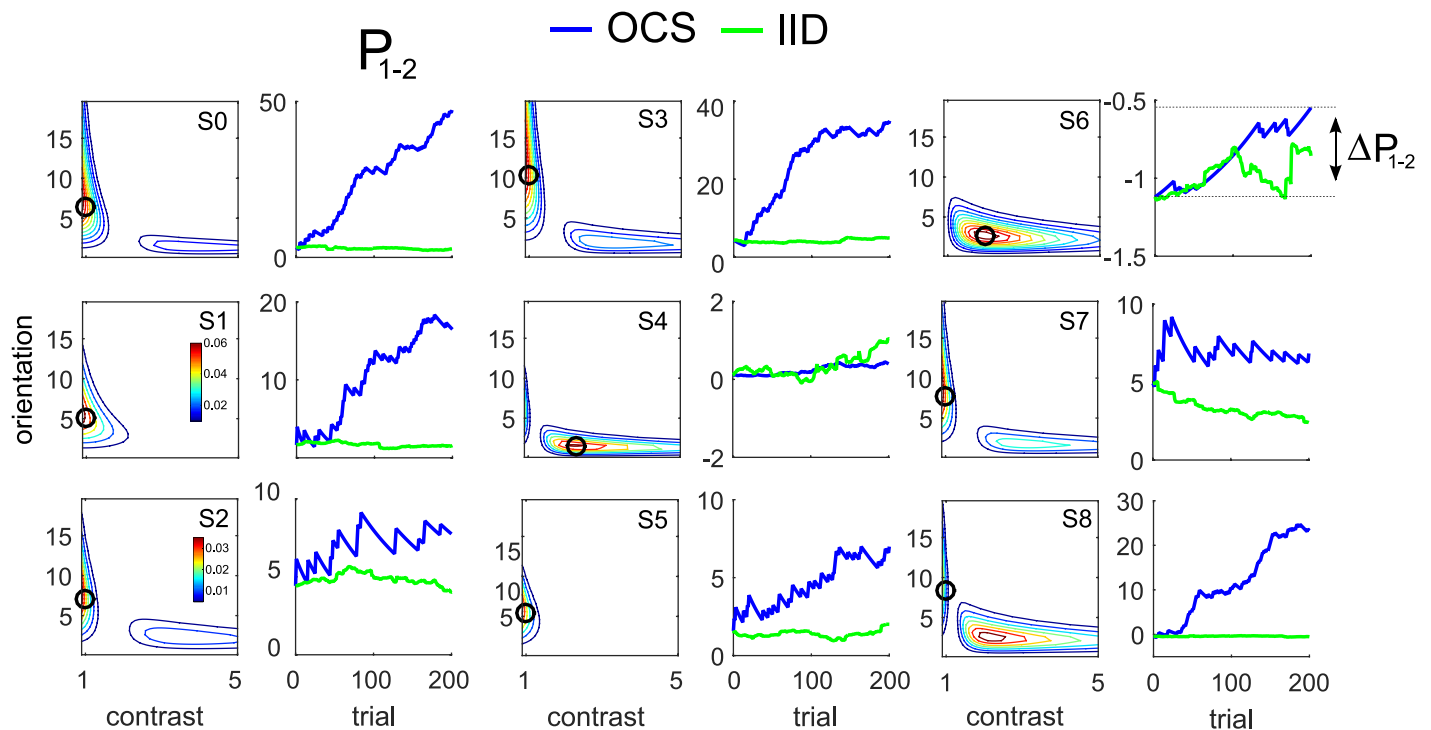


Figure 7. Left panels: OCS $\mathbf{s} = (c, d\phi)^T$ for discriminating models 1 and 2 (black circles), superimposed on a contour plot of the model comparison utility function (Equation 10). Color bars shown for only two subjects to minimize clutter. Right panels: Evolution of the model preference P_{1-2} during Experiment 2 for both OCS (blue curves) and stimuli chosen at random from the grid used in Experiment 1 (IID: green curves). Top right panel graphically illustrates the change in model preference (ΔP_{1-2}) defined in the text.

Numerical simulations of model comparison experiments

In order to more rigorously examine the potential utility of adaptive OCS in the ideal case in which one of the candidate models is actually the true process generating the data, we performed a simulation of Experiment 2 (C-phase) for all subjects. In these simulations, we took as the ground truth the Naka-Rushton model (model 1) and used the fit of this model to actual E-phase (Experiment 1) data to generate synthetic C-phase (Experiment 2) data. We quantified the C-phase change in model preference index ΔP_{1-2} for both IID and OCS data collection strategies in which the Naka-Rushton model was assumed true. In the actual experiments, at the end of the E-phase, there was already a model preference ($P_{1-2} \neq 0$, see Figure 6a), so in order to determine how often the two data collection strategies (OCS, IID) would result in a correct choice given no initial preference, we set the initial model preference to zero so that $\Delta P_{1-2} = P_{1-2}$.

Results of $N_{mc} = 100$ Monte Carlo simulations of Experiment 2 are shown in Figure 8. In each panel, we plot the median value of P_{1-2} (thick lines: blue = OCS, green = IID), the range containing 95% of simulations (thin lines), and the trajectory of P_{1-2}

observed experimentally (red lines). For many (but not all) subjects, we see a reasonably good agreement between the simulation predictions and the observed change in model preferences during the C-phase. We find that over the group of subjects, there is a correlation (Pearson, $n = 9$, $r = 0.71$, $p = 0.03$) between the predictions of ΔP_{1-2} predicted by the simulations and those observed experimentally (Supplementary Figure S7). The simulations tend to predict a larger value of ΔP_{1-2} than observed experimentally (median: experiments = 5.41, simulations = 13.59) although, just like the experiments, the median ΔP_{1-2} obtained is larger for simulations using OCS than IID (median = 0.67) data collection strategies. We also find that one is more likely to make a correct model choice using the OCS data collection method (Supplementary Table S3) with IID yielding a correct choice after $N_C = 200$ trials (given no initial preference) in 80% of simulations but OCS in about 99%. Additional simulations also reveal that OCS stimuli can also be more effective for model comparison in cases in which model 2 is the ground truth (Supplementary Figure S8). These simulations suggest the potential usefulness of this adaptive stimulus optimization method for comparing competing models of neural encoding.

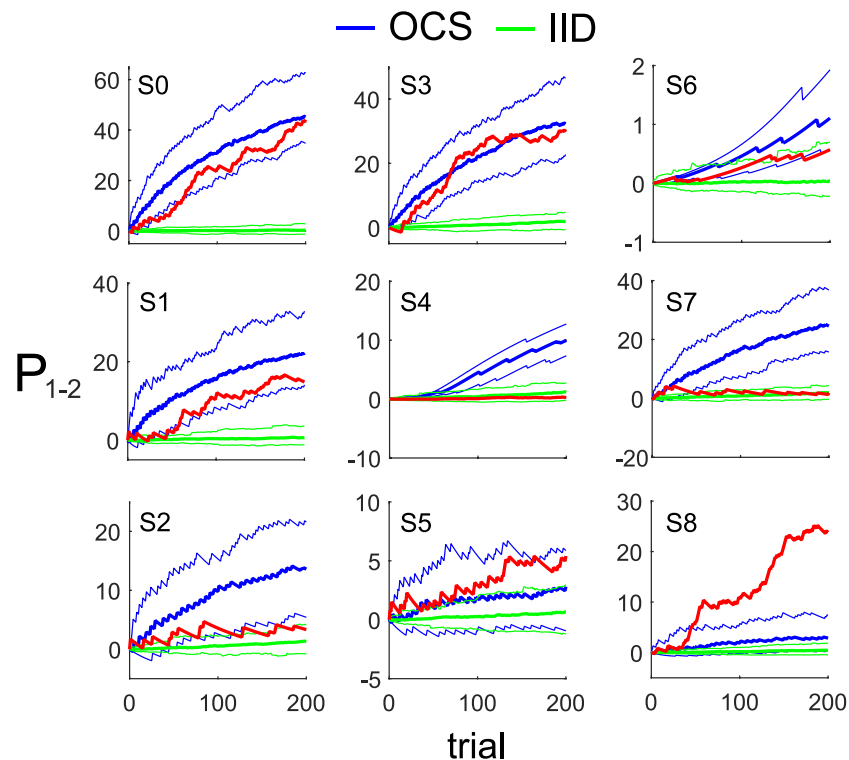


Figure 8. Results from $N_{mc} = 100$ Monte Carlo simulations of Experiment 2 in which synthetic data is generated by fits of the Naka-Rushton model to Experiment 1 data. Simulation results are shown for OCS (blue curves) and IID (green curves) stimuli with thick lines denoting median values of P_{1-2} and thin lines denoting the middle 95% of values. Superimposed on these plots are the dynamic model preferences (red curves) actually observed during the real Experiment 2 performed on subjects (Figure 7).

Discussion

Neural codes from behavior

For more than 40 years there has been a rich, two-way traffic of ideas between sensory neurophysiology and psychophysics with computational modeling often forming the bridge between these levels of analysis (Gold & Shadlen, 2007; Nienborg et al., 2012; Parker & Newsome, 1998; Romo & de Lafuente, 2013). Most often, computational modeling has been applied to neural data in order to predict or explain behavior (e.g., Kiani et al., 2014; Purushothaman & Bradley, 2005; Shadlen & Newsome, 2001) rather than being applied to behavioral data to gain insight about neural mechanisms. However, in a number of recent studies, a growing number of investigators have taken the complementary approach of using behavioral experiments or neural modeling of optimal behavior to inform and connect with neural encoding models. Here we briefly review some of this work before relating it to the present study.

One example of deriving neural codes from behavioral considerations is *accuracy maximization analysis*, which finds optimal neural encoding models for specific natural perception tasks (Burge, Fowlkes, & Banks,

2010; Burge & Geisler, 2011, 2014, 2015; W. Geisler, Perry, Super, & Gallogly, 2001; W. S. Geisler, 2008; W. S. Geisler et al., 2009). This methodology has been applied to determine the neural receptive fields that would be optimal for performing natural vision tasks, such as separating figure from ground (Burge et al., 2010; W. S. Geisler et al., 2009), estimating retinal disparity (Burge & Geisler, 2014), and estimating the speed of visual motion (Burge & Geisler, 2015). The neural encoding models derived account for experimentally observed neural tuning properties, and although these models were not estimated by fitting psychophysical data (as done here), a Bayesian ideal observer reading out these optimal neural codes manages to accurately account for human psychophysical performance (e.g., Burge & Geisler, 2015).

Another line of research which fits theoretically optimal performance has employed neural implementations of Bayesian ideal observers to understand how optimal or near-optimal behavioral performance can be explained in terms of probabilistic population coding (Beck et al., 2008; Ma, 2010; Ma, Beck, Latham, & Pouget, 2006; Ma et al., 2011; Qamar et al., 2013). One recent study of this kind has demonstrated that one can account for near-optimal visual search behavior seen in human observers using a neural model implementing probabilistic population codes that represent stimulus

reliability (Ma et al., 2011). Another recent study (Qamar et al., 2013) demonstrated that a neural model that accounts for the ability of subjects to make trial-by-trial adjustments of decision boundaries in a categorization task automatically learns to perform divisive computations like those seen in visual neurons (Carandini & Heeger, 2012).

Another recent example of predicting receptive field properties by modeling behavioral performance comes from a recent study (Yamins et al., 2014) that explored a large number of computational models of the ventral visual stream using a high-throughput modeling technique (Pinto, Doukhan, DiCarlo, & Cox, 2009). This work revealed that models that could account for human behavioral performance on a challenging object recognition task (but not fit to neural data) had intermediate and output-layer units whose responses closely matched neural tuning observed in visual areas V4 and IT (Yamins et al., 2014). Another neural modeling study (Salinas, 2006) showed that one can explain the shape of tuning curves used by different sensory systems by taking into account the downstream motor behavior that decodes these sensory representations, using examples as diverse as binocular disparity in vision and echo delay in bats. These studies suggest that behavior can provide strong constraints on the nature of neural computation in the sensory systems.

Other efforts to use behavior to inform theories of neural mechanism come from the perceptual learning literature, in which investigators have proposed neural models that account for improvements in performance with experience despite relatively stable early-stage sensory encoding (Doshier, Jeter, Liu, & Lu, 2013; Doshier & Lu, 1998, 1999; Petrov et al., 2005). One recent model demonstrates that Hebbian modifications to the task-specific readout of a stable neural population is sufficient to explain perceptual learning and explains the empirically observed “switch cost” when the background noise context changes (Petrov et al., 2005). Other work (Bejjanki, Beck, Lu, & Pouget, 2011) has suggested that perceptual learning can be construed as improved probabilistic inference, in which altering only feed-forward weights input weights to a recurrent neural network can yield a modest sharpening of tuning curves as observed experimentally (Yang & Maunsell, 2004).

A number of studies have used classification images (Ahumada, 1996; Murray, 2011) to make direct comparisons between the properties of perceptual filters and neural response properties (see review by Neri & Levi, 2006). For instance, one such study demonstrated that performance on a bar detection task could be explained using a combination of linear matched filtering and contrast energy detection, similar to mechanisms known to exist in V1 simple and complex cells (Neri & Heeger, 2002). Other studies have revealed striking relationships between the optimal

perceptual filter for orientation discrimination and the receptive fields of V1 neurons (Ringach, 1998; Solomon, 2002) or have demonstrated multiplicative perceptual combination of visual cues similar to that observed physiologically (Neri, 2004). Classification image studies (Eckstein, Shimozaki, & Abbey, 2002; Murray et al., 2003; Neri, 2004) have demonstrated that, consistent with physiological studies of attentional effects on neurons, the shape of perceptive fields do not change with attention. Taken as a whole, this body of work suggests that the classification image technique can potentially shed light on neural mechanisms.

Several recent studies have considered the optimal distribution of neuronal tuning curves for efficiently encoding sensory variables and the implications of anisotropic neural populations for perceptual behavior (Ganguli & Simoncelli, 2010, 2014; Girshick, Landy, & Simoncelli, 2011; Wei & Stocker, 2015). In addition to comparing the predictions of theoretical models to the distribution of neural tuning curves observed experimentally, models of population decoding with such anisotropic populations have also been shown to explain psychophysical data, such as orientation and spatial frequency discrimination thresholds (Ganguli & Simoncelli, 2010) and perceptual biases (Girshick et al., 2011; Wei & Stocker, 2015). Although these studies do not directly infer physiological properties from fits of psychometric models to behavioral data, they do demonstrate that behaviorally relevant considerations (i.e., optimal representation of the world and perceptual decisions) can explain some features of neural encoding.

In the pattern vision literature, a number of investigators have utilized numerical simulations of early visual processing aimed at explaining psychophysical performance on contrast detection tasks (Chirimuuta & Tolhurst, 2005; Clatworthy, Chirimuuta, Lauritzen, & Tolhurst, 2003; Goris et al., 2013; Goris, Wichmann, & Henning, 2009). One study of this kind demonstrated that a large number of well-known results in the contrast detection literature could be accounted for by a neural population model of the early visual system that takes into account known biological nonlinearities (Goris et al., 2013). Another notable study modeling spatial pooling in human contrast detection (Morgenstern & Elder, 2012) was able to define analytical models specified in terms of local Gabor receptive field parameters. The authors found that the best model to account for their psychophysical data had local receptive fields approximately the size of those seen in V1 whose outputs were passed through an energy filter and summed, similar to known mechanisms in visual cortex. As in the present study (Figure 4), these authors presented a direct comparison with their estimated parameter values and previously published physiological data (Morgenstern & Elder, 2012, figure 14).

One of the limitations of numerical models is that it can often be difficult to directly relate the neural model parameters to behavioral performance. Therefore, another recent study (May & Solomon, 2015a, 2015b) took the approach of deriving analytical models of psychophysical performance on contrast detection and discrimination tasks using the natural link between psychophysical performance and the Fisher information of a neural population (Dayan & Abbott, 2001). The authors managed to demonstrate a surprisingly simple and intuitive relationship between the parameters of the neural code and perceptual performance and were able to account for the results of previous numerical simulation studies (Chirimuuta & Tolhurst, 2005; Clatworthy et al., 2003). Along these same lines, two recent studies (perhaps most closely related to the present work) used fits of low-dimensional analytically defined neural models to psychophysical data in order to predict how contrast gain encoding in orientation-tuned visual neurons may be modulated by attention and understand mechanisms of attentional pooling (Pestilli et al., 2011; Pestilli et al., 2009). This work demonstrates very elegantly how one can use data obtained from behavioral experiments to make precise, quantitative predictions about neural encoding.

Relationship to previous work

As detailed above, a number of previous studies have used fits of neural models to behavioral data in order to gain insight about neural encoding and decoding mechanisms. The current work is complementary to these studies and makes a number of novel contributions to extend this general approach further.

First, we present for didactic purposes a simple mathematical derivation that frames the psychometric function explicitly in terms of the neural encoding model, showing that one can in principle use psychophysical data to estimate neural encoding model parameters (Appendix A). Although the main result (Equation 3) is not directly useful without specific assumptions about the neural encoding and behavioral decoding models, it serves to make explicit the general approach taken here and in related work.

Second, unlike many previous studies that either fit or compare neural models to previously collected psychophysical data (e.g., Goris et al., 2013; May & Solomon, 2015a, 2015b) or to theoretically optimal ideal observer performance (e.g., Burge & Geisler, 2014, 2015; Geisler et al., 2009), we performed our own psychophysical experiments on human subjects and estimated model parameters directly by fitting to psychophysical trial data.

Third, because our model was a simple low-dimensional analytical model (e.g., May & Solomon,

2015a, 2015b; Morgenstern & Elder, 2012; Pestilli et al., 2009), as opposed to being defined by a complex numerical simulation (Chirimuuta & Tolhurst, 2005; Clatworthy et al., 2003; Goris et al., 2013; Goris et al., 2009) or a high-dimensional perceptual filter (Ahumada, 1996; Murray, 2011; Neri & Levi, 2006), it was possible to do fast model fitting online during the course of the experimental session rather than doing so post hoc as in previous work. As illustrated previously (DiMattina & Zhang, 2011, 2013; Tam, 2012), the ability to estimate models in real time during the course of the experiment is essential if one wishes to generate novel stimuli to compare competing models.

Fourth, in contrast to many previous studies, we present direct comparisons (Figure 4) between the values of model parameters estimated from fitting psychophysical trial data and those independently measured (Albrecht & Hamilton, 1982) in physiological studies (but see Morgenstern & Elder, 2012; Neri & Levi, 2006, for exceptions). This direct comparison made by ourselves and others (e.g., Morgenstern & Elder, 2012) with previously published physiology data makes a strong case that one can get accurate estimates of neural system parameters by fitting psychophysical trial data.

Finally and perhaps most importantly, although other studies have demonstrated how one can use psychophysical data for post hoc model comparison (Morgenstern & Elder, 2012; Pestilli et al., 2009; Qamar et al., 2013; van den Berg et al., 2014), we extend this idea further by considering how one can adaptively optimize stimuli explicitly during the experimental session for purposes of model comparison. We show that adaptively optimized stimuli are far more effective for model comparison than post hoc analyses using both experiments (Figure 7) and numerical simulations (Figure 8). Although qualitatively very different models (Naka-Rushton and Gaussian contrast gain) can be well discriminated without this technique (Figure 6), it can be very helpful to distinguish between qualitatively similar models (Naka-Rushton and Tanh: see Figure 2). We feel that this general approach of adaptive stimulus generation offers great promise for psychophysical and physiological experiments (Cavagnaro et al., 2013; Cavagnaro et al., 2011; DiMattina & Zhang, 2013; Myung et al., 2013; Z. Wang & Simoncelli, 2008) and is of great interest for future work.

Limitations

In the example analyzed in this study, we only estimated a relatively modest number of biologically interpretable parameters from psychophysical data. However, although the example we use is fairly modest, the theoretical results we present here are fully general and can be applied in a variety of contexts, subject only

to the practical limitations imposed by the amount of psychophysical data that is needed to reliably estimate high-dimensional models (Mineault et al., 2009; Murray, 2011). It is of great interest for us to extend this methodology to higher dimensional examples, for instance, estimating a parametric model of orientation-tuning anisotropy in area V1 (see below).

In both the present example as well as many previous efforts to fit neural models to behavioral data, the neural encoding models were fairly simple early-stage neural encoders, for instance, a population of V1 cells tuned to orientation and/or contrast (Goris et al., 2013; May & Solomon, 2015a, 2015b; Pestilli et al., 2011; Pestilli et al., 2009). In our opinion, this method is likely to be most useful for recovering physiological properties of low-level neural encoders that can be specified by a few parameters. Although previous studies have fit neural models to behavioral data arising from higher level cognitive tasks, such as visual search or working memory (Bays, 2014; Ma et al., 2011), it is less likely that the method presented here will be able to provide much direct physiological insight in these cases.

Another limitation of the present study is that we relied on fairly simple assumptions about the neuronal noise (independent responses) and a linear decoding strategy. However, even with these simplifications, we attained excellent fits of our derived model to the psychophysical data (Figure 5, Supplementary Figure S3), and the model we derived had reasonably good predictive validity for novel stimuli (Supplementary Figure S4). Previous work has demonstrated that simple linear decoders are adequate for estimating stimulus parameters from neural data (Berens et al., 2012) and has suggested that naive decoding strategies that do not take noise correlations into account can be nearly as effective as decoding that assumes such knowledge (May & Solomon, 2015a).

Finally, a further limitation is that by design of the experiment (which focused on low-contrast gratings), we were only concerned with estimating the parameters of the contrast gain functions for a subpopulation of the most sensitive neurons, which we assumed to be identically tuned. In reality, there is diversity in the contrast thresholds (c_{50}) and shapes (n) of contrast gain functions (Albrecht & Hamilton, 1982). Therefore, our results only demonstrate that our subpopulation of interest is sufficient to explain the observed psychophysical behavior and does not rule out the possibility that other neurons not considered by our model may contribute as well.

Future directions

We feel that there is a lot of potential for this general methodological approach to be applied to test hy-

potheses of the large-scale organization of heterogeneous neural population codes using psychophysical experiments. One well-studied example is the population of orientation-tuned neurons in V1, which are more densely located and more narrowly tuned near the cardinal (horizontal, vertical) than near oblique orientations (Li, Peterson, & Freeman, 2003). This orientation-tuning anisotropy matches the statistics of natural images (Girshick et al., 2011), and when such an anisotropic neural population is combined with a Bayesian decoder, it can explain a number of biases observed in orientation discrimination tasks (Wei & Stocker, 2015). One can in principle apply our methodology to estimate a parametric model describing heterogeneity in V1 tuning curve parameters (e.g., variations in density and tuning width as a function of preferred grating orientation). This could be accomplished by defining a parametric model of the stimulus-dependent Fisher information $I_F(\phi) = F(\phi, \mathbf{v})$ as a function of reference orientation ϕ , which would serve as the link between the neural population code and psychophysical performance (May & Solomon, 2015a; Wei & Stocker, 2015). After estimating the parameters $\hat{\mathbf{v}}$ from psychophysical experiments, one can then optimize neural population code parameters θ to minimize $\int_{[0, \pi)} (F(\phi, \hat{\mathbf{v}}) - I_F(\phi, \theta))^2 d\phi$, where $I_F(\phi, \theta)$ denotes the Fisher information predicted by a neural encoding model having parameters θ . Because many different neural population codes are capable of giving rise to very similar Fisher information profiles (Wei & Stocker, 2015), additional constraints, such as coding efficiency (Ganguli & Simoncelli, 2014), may be necessary in order to get a unique solution for neural population code parameters. Conducting such technically challenging psychophysics experiments aimed at understanding the large-scale organization of neural population codes is an interesting direction of future research.

Conclusions

Although psychophysics can certainly never supplant physiological studies, several recent modeling studies suggest that modeling behavioral data can provide insights into neural encoding mechanisms. Perhaps this ability of behavior to provide guidance to neurophysiology is not too surprising given the long history of psychophysical observations accurately predicting physiological mechanisms many years before their discovery, for instance, the neural encoding of color (Helmholtz, 1925; Read, 2015; Wald, 1964; Young, 1802). We believe that behavioral studies will continue to play an important role in guiding neurophysiological research, making the development of better computational methodology for integrating

behavioral and neurophysiological studies an important and worthwhile goal.

Keywords: *computational modeling, neural encoding, psychometric functions, psychophysics*

Acknowledgments

The author would like to thank FGCU student Steven Davis for help with the experiments and thanks Curtis L. Baker Jr., Nick Prins, and Wei-Ji Ma for comments on the manuscript. The author declares no competing financial interests.

Commercial relationships: none.

Corresponding author: Christopher DiMattina.

Email: cdimattina@fgcu.edu.

Address: Computational Perception Laboratory, Department of Psychology, Florida Gulf Coast University, Fort Myers, FL, USA.

References

- Ahumada, A. J. (1996). Perceptual classification images from vernier acuity masked by noise. *Perception*, 25, ECVF abstract supplement.
- Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6), 716–723.
- Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology*, 48(1), 217–237.
- Barbour, D. L., & Wang, X. (Feb 14, 2003). Contrast tuning in auditory cortex. *Science*, 299(5609), 1073–1075.
- Bays, P. M. (2014). Noise in neural populations accounts for errors in working memory. *The Journal of Neuroscience*, 34(10), 3632–3645.
- Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., . . . Pouget, A. (2008). Probabilistic population codes for Bayesian decision making. *Neuron*, 60(6), 1142–1152.
- Bejjanki, V. R., Beck, J. M., Lu, Z.-L., & Pouget, A. (2011). Perceptual learning as improved probabilistic inference in early sensory areas. *Nature Neuroscience*, 14(5), 642–648.
- Bensmaia, S. J., Denchev, P. V., Dammann, J. F., Craig, J. C., & Hsiao, S. S. (2008). The representation of stimulus orientation in the early stages of somatosensory processing. *The Journal of Neuroscience*, 28(3), 776–786.
- Berens, P., Ecker, A. S., Cotton, R. J., Ma, W. J., Bethge, M., & Tolias, A. S. (2012). A fast and simple population code for orientation in primate V1. *The Journal of Neuroscience*, 32(31), 10618–10626.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. New York: Springer.
- Bollimunta, A., Totten, D., & Ditterich, J. (2012). Neural dynamics of choice: Single-trial analysis of decision-related activity in parietal cortex. *The Journal of Neuroscience*, 32(37), 12684–12701.
- Borst, A., & Theunissen, F. E. (1999). Information theory and neural coding. *Nature Neuroscience*, 2(11), 947–957.
- Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *The Journal of Neuroscience*, 12(12), 4745–4765.
- Burge, J., Fowlkes, C. C., & Banks, M. S. (2010). Natural-scene statistics predict how the figure-ground cue of convexity affects human depth perception. *The Journal of Neuroscience*, 30(21), 7269–7280.
- Burge, J., & Geisler, W. S. (2011). Optimal defocus estimation in individual natural images. *Proceedings of the National Academy of Sciences, USA*, 108(40), 16849–16854.
- Burge, J., & Geisler, W. S. (2014). Optimal disparity estimation in natural stereo images. *Journal of Vision*, 14(2):1, 1–18, doi:10.1167/14.2.1. [PubMed] [Article]
- Burge, J., & Geisler, W. S. (2015). Optimal speed estimation in natural image movies predicts human performance. *Nature Communications*, 6, 7900.
- Burnham, K. P., & Anderson, D. R. (2003). *Model selection and multimodel inference: A practical information-theoretic approach*. New York: Springer Science & Business Media.
- Campbell, F. W., & Robson, J. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology*, 197(3), 551–566.
- Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1), 51–62.
- Cavagnaro, D. R., Gonzalez, R., Myung, J. I., & Pitt, M. A. (2013). Optimal decision stimuli for risky choice experiments: An adaptive approach. *Management Science*, 59(2), 358–375.
- Cavagnaro, D. R., Myung, J. I., Pitt, M. A., & Kujala,

- J. V. (2010). Adaptive design optimization: A mutual information-based approach to model discrimination in cognitive science. *Neural Computation*, 22(4), 887–905.
- Cavagnaro, D. R., Pitt, M. A., & Myung, J. I. (2011). Model discrimination through adaptive experimentation. *Psychonomic Bulletin & Review*, 18(1), 204–210.
- Chirimuuta, M., & Tolhurst, D. J. (2005). Does a Bayesian model of V1 contrast coding offer a neurophysiological account of human contrast discrimination? *Vision Research*, 45(23), 2943–2959.
- Clatworthy, P., Chirimuuta, M., Lauritzen, J., & Tolhurst, D. (2003). Coding of the contrasts in natural images by populations of neurons in primary visual cortex (V1). *Vision Research*, 43(18), 1983–2001.
- Cohen, M. R., & Newsome, W. T. (2009). Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *The Journal of Neuroscience*, 29(20), 6635–6648.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*. Hoboken, NJ: John Wiley & Sons.
- Dayan, P., & Abbott, L. F. (2001). *Theoretical neuroscience, volume 806*. Cambridge, MA: MIT Press.
- DiMattina, C. (2015). Fast adaptive estimation of multidimensional psychometric functions. *Journal of Vision*, 15(9):5, 1–20, doi:10.1167/15.9.5. [PubMed] [Article]
- DiMattina, C., & Zhang, K. (2011). Active data collection for efficient estimation and comparison of nonlinear neural models. *Neural Computation*, 23(9), 2242–2288.
- DiMattina, C., & Zhang, K. (2013). Adaptive stimulus optimization for sensory systems neuroscience. *Frontiers in Neural Circuits*, 7, 101.
- Dosher, B. A., Jeter, P., Liu, J., & Lu, Z.-L. (2013). An integrated reweighting theory of perceptual learning. *Proceedings of the National Academy of Sciences, USA*, 110(33), 13678–13683.
- Dosher, B. A., & Lu, Z.-L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences, USA*, 95(23), 13988–13993.
- Dosher, B. A., & Lu, Z.-L. (1999). Mechanisms of perceptual learning. *Vision Research*, 39(19), 3197–3221.
- Eckstein, M. P., & Ahumada, A. J. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision*, 2(1):i, doi:10.1167/2.1.i. [PubMed] [Article]
- Eckstein, M. P., Shimozaki, S. S., & Abbey, C. K. (2002). The footprints of visual attention in the Posner cueing paradigm revealed by classification images. *Journal of Vision*, 2(1):3, 25–45, doi:10.1167/2.1.3. [PubMed] [Article]
- Egger, S. W., & Britten, K. H. (2013). Linking sensory neurons to visually guided behavior: Relating MST activity to steering in a virtual environment. *Visual Neuroscience*, 30(5–6), 315–330.
- Ganguli, D., & Simoncelli, E. P. (2010). Implicit encoding of prior probabilities in optimal neural populations. In *Advances in neural information processing systems* (pp. 658–666). Cambridge, MA: MIT Press.
- Ganguli, D., & Simoncelli, E. P. (2014). Efficient sensory encoding and Bayesian inference with heterogeneous neural populations. *Neural Computation*, 26, 2103–2134.
- Geisler, W., Perry, J., Super, B., & Gallogly, D. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, 41(6), 711–724.
- Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, 59, 167–192.
- Geisler, W. S., Najemnik, J., & Ing, A. D. (2009). Optimal stimulus encoders for natural tasks. *Journal of Vision*, 9(13):17, 1–16, doi:10.1167/9.13.17. [PubMed] [Article]
- Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: Visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, 14(7), 926–932.
- Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30, 535–574.
- Goris, R. L., Putzeys, T., Wagemans, J., & Wichmann, F. A. (2013). A neural population model for visual pattern detection. *Psychological Review*, 120(3), 472–496.
- Goris, R. L., Wichmann, F. A., & Henning, G. B. (2009). A neurophysiologically plausible population code model for human contrast discrimination. *Journal of Vision*, 9(7):15, 1–22, doi:10.1167/9.7.15. [PubMed] [Article]
- Graf, A. B., Kohn, A., Jazayeri, M., & Movshon, J. A. (2011). Decoding the activity of neuronal populations in macaque primary visual cortex. *Nature Neuroscience*, 14(2), 239–245.

- Helmholtz, H. V. (1925). *Handbook of physiological optics*. Mineola, NY: Dover.
- Kiang, N. Y. (1965). *Discharge patterns of single fibers in the cat's auditory nerve*. Cambridge, MA: MIT.
- Kiani, R., Cueva, C. J., Reppas, J. B., & Newsome, W. T. (2014). Dynamics of neural population responses in prefrontal cortex indicate changes of mind on single trials. *Current Biology*, 24(13), 1542–1547.
- Kingdom, F., & Prins, N. (2010). *Psychophysics: A practical introduction*. London: Academic Press.
- Lewi, J., Butera, R., & Paninski, L. (2009). Sequential optimal design of neurophysiology experiments. *Neural Computation*, 21(3), 619–687.
- Li, B., Peterson, M. R., & Freeman, R. D. (2003). Oblique effect: A neural basis in the visual cortex. *Journal of Neurophysiology*, 90(1), 204–217.
- Ma, W. J. (2010). Signal detection theory, uncertainty, and Poisson-like population codes. *Vision Research*, 50(22), 2308–2319.
- Ma, W. J., Beck, J. M., Latham, P. E., & Pouget, A. (2006). Bayesian inference with probabilistic population codes. *Nature Neuroscience*, 9(11), 1432–1438.
- Ma, W. J., Navalpakkam, V., Beck, J. M., van den Berg, R., & Pouget, A. (2011). Behavior and neural basis of near-optimal visual search. *Nature Neuroscience*, 14(6), 783–790.
- May, K. A., & Solomon, J. A. (2015a). Connecting psychophysical performance to neuronal response properties I: Discrimination of suprathreshold stimuli. *Journal of Vision*, 15(6):8, 1–26, doi:10.1167/15.6.8. [PubMed] [Article]
- May, K. A., & Solomon, J. A. (2015b). Connecting psychophysical performance to neuronal response properties II: Contrast decoding and detection. *Journal of Vision*, 15(6):9, 1–21, doi:10.1167/15.6.9. [PubMed] [Article]
- Mineault, P. J., Barthelmé, S., & Pack, C. C. (2009). Improved classification images with sparse priors in a smooth basis. *Journal of Vision*, 9(10):17, 1–24, doi:10.1167/9.10.17. [PubMed] [Article]
- Morgenstern, Y., & Elder, J. H. (2012). Local visual energy mechanisms revealed by detection of global patterns. *The Journal of Neuroscience*, 32(11), 3679–3696.
- Mountcastle, V. B., LaMotte, R. H., & Carli, G. (1972). Detection thresholds for stimuli in humans and monkeys: Comparison with threshold events in mechanoreceptive afferent nerve fibers innervating the monkey hand. *Journal of Neurophysiology*, 35(1), 122–136.
- Muniak, M. A., Ray, S., Hsiao, S. S., Dammann, J. F., & Bensmaia, S. J. (2007). The neural coding of stimulus intensity: Linking the population response of mechanoreceptive afferents with psychophysical behavior. *The Journal of Neuroscience*, 27(43), 11687–11699.
- Murray, R. F. (2011). Classification images: A review. *Journal of Vision*, 11(5):2, 1–25, doi:10.1167/11.5.2. [PubMed] [Article]
- Murray, R. F., Sekuler, A. B., & Bennett, P. J. (2003). A linear cue combination framework for understanding selective attention. *Journal of Vision*, 3(2):2, 116–145, doi:10.1167/3.2.2. [PubMed] [Article]
- Myung, J. I., Cavagnaro, D. R., & Pitt, M. A. (2013). A tutorial on adaptive design optimization. *Journal of Mathematical Psychology*, 57(3), 53–67.
- Neri, P. (2004). Attentional effects on sensory tuning for single-feature detection and double-feature conjunction. *Vision Research*, 44(26), 3053–3064.
- Neri, P., & Heeger, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience*, 5(8), 812–816.
- Neri, P., & Levi, D. M. (2006). Receptive versus perceptive fields from the reverse-correlation viewpoint. *Vision Research*, 46(16), 2465–2474.
- Nienborg, H., Cohen, R., & Cumming, B. G. (2012). Decision-related activity in sensory neurons: Correlations among neurons and with behavior. *Annual Review of Neuroscience*, 35, 463–483.
- Paninski, L., Pillow, J., & Lewi, J. (2007). Statistical models for neural encoding, decoding, and optimal stimulus design. *Progress in Brain Research*, 165, 493–507.
- Parker, A. J., & Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience*, 21(1), 227–277.
- Pestilli, F., Carrasco, M., Heeger, D. J., & Gardner, J. L. (2011). Attentional enhancement via selection and pooling of early sensory responses in human visual cortex. *Neuron*, 72(5), 832–846.
- Pestilli, F., Ling, S., & Carrasco, M. (2009). A population-coding model of attention's influence on contrast response: Estimating neural effects from psychophysical data. *Vision Research*, 49(10), 1144–1153.
- Petrov, A. A., Doshier, B. A., & Lu, Z.-L. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review*, 112(4), 715–743.
- Pinto, N., Doukhan, D., DiCarlo, J. J., & Cox, D. D. (2009). A high-throughput screening approach to

- discovering good forms of biologically inspired visual representation. *PLoS Computational Biology*, 5(11), e1000579.
- Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, 6(10), 421–425.
- Purushothaman, G., & Bradley, D. C. (2005). Neural population code for fine perceptual decisions in area MT. *Nature Neuroscience*, 8(1), 99–106.
- Qamar, A. T., Cotton, R. J., George, R. G., Beck, J. M., Prezhdo, E., Laudano, A., . . . Ma, W. J. (2013). Trial-to-trial, uncertainty-based adjustment of decision boundaries in visual categorization. *Proceedings of the National Academy of Sciences, USA*, 110(50), 20332–20337.
- Read, J. (2015). The place of human psychophysics in modern neuroscience. *Neuroscience*, 296, 116–129.
- Ringach, D. L. (1998). Tuning of orientation detectors in human vision. *Vision Research*, 38(7), 963–972.
- Romo, R., & de Lafuente, V. (2013). Conversion of sensory signals into perceptual decisions. *Progress in Neurobiology*, 103, 41–75.
- Sachs, M. B., & Abbas, P. J. (1974). Rate versus level functions for auditory-nerve fibers in cats: Tone-burst stimuli. *The Journal of the Acoustical Society of America*, 56(6), 1835–1847.
- Sadagopan, S., & Wang, X. (2008). Level invariant representation of sounds by populations of neurons in primary auditory cortex. *The Journal of Neuroscience*, 28(13), 3415–3426.
- Salinas, E. (2006). How behavioral constraints may determine optimal sensory representations. *PLoS Biology*, 4(12), e387.
- Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *The Journal of Neuroscience*, 16(4), 1486–1510.
- Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area lip) of the rhesus monkey. *Journal of Neurophysiology*, 86(4), 1916–1936.
- Skottun, B. C., Bradley, A., Sclar, G., Ohzawa, I., & Freeman, R. D. (1987). The effects of contrast on visual orientation and spatial frequency discrimination: A comparison of single cells and behavior. *Journal of Neurophysiology*, 57(3), 773–786.
- Solomon, J. A. (2002). Noise reveals visual mechanisms of detection and discrimination. *Journal of Vision*, 2(1):7, 105–120, doi:10.1167/2.1.7. [PubMed] [Article]
- Tam, W. (2012). *Adaptive modeling of marmoset inferior colliculus neurons in vivo*. PhD thesis, Johns Hopkins University, Baltimore, MD.
- van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review*, 121(1), 124–149.
- Vogels, R., & Orban, G. (1990). How well do response changes of striate neurons signal differences in orientation: A study in the discriminating monkey. *The Journal of Neuroscience*, 10(11), 3543–3558.
- Wald, G. (Sept 4, 1964). The receptors of human color vision. *Science*, 145, 1007–1016.
- Wang, L., Narayan, R., Graña, G., Shamir, M., & Sen, K. (2007). Cortical discrimination of complex natural stimuli: Can single neurons match behavior? *The Journal of Neuroscience*, 27(3), 582–589.
- Wang, Z., & Simoncelli, E. P. (2008). Maximum differentiation (MAD) competition: A methodology for comparing computational models of perceptual quantities. *Journal of Vision*, 8(12):8, 1–13, doi:10.1167/8.12.8. [PubMed] [Article]
- Watson, A. B., & Fitzhugh, A. (1990). The method of constant stimuli is inefficient. *Perception & Psychophysics*, 47(1), 87–91.
- Wei, X.-X., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain “anti-Bayesian” percepts. *Nature Neuroscience*, 18(10), 1509–1517.
- Yamins, D. L., Hong, H., Cadieu, C. F., Solomon, E. A., Seibert, D., & DiCarlo, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences, USA*, 111(23), 8619–8624.
- Yang, T., & Maunsell, J. H. (2004). The effect of perceptual learning on neuronal responses in monkey visual area V4. *The Journal of Neuroscience*, 24(7), 1617–1626.
- Young, T. (1802). The Bakerian lecture: On the theory of light and colours. *Philosophical Transactions of the Royal Society of London*, 92, 12–48.

Appendix A

A psychophysical experiment will yield stimulus–response data $\mathcal{D}_\psi = \{\mathbf{s}_i, b_i\}_{i=1}^n$. Assuming that subject responses are independent across trials, we can write the data likelihood

$$P(\mathcal{D}_\psi | \boldsymbol{\theta}, \boldsymbol{\omega}) = \prod_{i=1}^n P(b_i | \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\omega}). \quad (11)$$

Bayes' rule expresses the posterior probability of θ, ω in terms of the data likelihood (Equation 11), yielding

$$P(\theta, \omega | \mathcal{D}_\psi) = \frac{1}{Z} P(\mathcal{D}_\psi | \theta, \omega) P(\theta, \omega), \quad (12)$$

where Z is a normalizing constant, and $P(\theta, \omega)$ reflects any prior beliefs about the neural code and the decoding parameters. Using Equations 11 and 2, we may rewrite Equation 12 as

$$P(\theta, \omega | \mathcal{D}_\psi) = \frac{1}{Z} \left[\prod_{i=1}^n \int P(b_i | \mathbf{r}, \omega) P(\mathbf{r} | \mathbf{s}_i, \theta) d\mathbf{r} \right] P(\theta, \omega). \quad (13)$$

Assuming that θ and ω are independent, so $P(\theta, \omega) = P(\theta)P(\omega)$, marginalizing Equation 13 over ω yields

$$P(\theta | \mathcal{D}_\psi) = \frac{1}{Z} \left[\prod_{i=1}^n \int \int P(b_i | \mathbf{r}, \omega) P(\mathbf{r} | \mathbf{s}_i, \theta) P(\omega) d\mathbf{r} d\omega \right] P(\theta). \quad (14)$$

In the case of fixed decoding model parameters $\hat{\omega}$, so that $P(\omega) = \delta(\omega - \hat{\omega})$ (where δ denotes the Dirac delta function), we obtain

$$P(\theta | \mathcal{D}_\psi) = \frac{1}{Z} \left[\prod_{i=1}^n \int P(b_i | \mathbf{r}, \hat{\omega}) P(\mathbf{r} | \mathbf{s}_i, \theta) d\mathbf{r} \right] P(\theta). \quad (15)$$

from Equation 14. A symmetrical argument of the same form as that presented above can be used to show that we can use psychophysical data to estimate the parameters of a behavioral decoding model given a neural encoding model with known parameters $\hat{\theta}$ using the equation

$$P(\omega | \mathcal{D}_\psi) = \frac{1}{Z} \left[\prod_{i=1}^n \int P(b_i | \mathbf{r}, \hat{\theta}, \omega) P(\mathbf{r} | \mathbf{s}_i, \hat{\theta}) d\mathbf{r} \right] P(\omega). \quad (16)$$

This allows a similar process of model fitting and comparison to be used in order to test competing hypotheses of neural decoding, for instance, whether simple linear decoding (Berens et al., 2012) or more complicated decoding mechanisms (Graf, Kohn, Jazayeri, & Movshon, 2011) are needed to accurately recover stimulus parameters or explain behavior.

Appendix B

Let stimulus $\mathbf{s}_- = (\phi_0 - d\phi, c)^T$ denote the clockwise stimulus and $\mathbf{s}_+$ denote the counterclockwise stimulus with parameters $\mathbf{s}_+ = (\phi_0 + d\phi, c)^T$. We will assume that orientation and contrast are coded by a population of N independent neurons, whose expected noisy (Poisson)

response r_i to a stimulus $\mathbf{s} = (\phi, c)^T$ is given by the 2-D contrast-modulated tuning curve $f_i(\phi, c) = \psi(c)f_i(\phi)$, where $f_i(\phi)$ describes the orientation tuning of the i th unit and $\psi(c)$ describes the contrast gain with $0 \leq \psi(c) \leq 1$ for contrast (in percentage) $0 \leq c \leq 100$. We will also assume that all of the units decoded for the behavioral decision have the same contrast gain function $\psi(c)$. For tractability, we approximate the Poisson noise response by Gaussian noise with mean and variance $\mu = \sigma^2 = f_i$. This fully specifies our neural encoding model $P(\mathbf{r} | \mathbf{s}, \theta)$ as a factorial Gaussian distribution.

To specify the behavioral decoding model $P(b | \mathbf{r}, \omega)$, where $b = 1$ indicates a correct response ($b = 0$ incorrect), we assume that the responses of all units are pooled linearly to form a new decision variable

$$u = \sum_{i=1}^N \omega_i r_i = \omega^T \mathbf{r}, \quad (17)$$

where the ω_i are dependent on the perceptual task. Because the weighted sum of Gaussian variables is also Gaussian, this new decision variable (Equation 17) is Gaussian, and the expected value for stimulus $\mathbf{s}_0 = (\phi_0, c)^T$ is given by

$$\mu_0 = \psi(c) \sum_{i=1}^N \omega_i f_i(\phi_0), \quad (18)$$

with variance

$$\sigma^2 = \sum_{i=1}^N \omega_i^2 \text{Var}[r_i] = \psi(c) \sum_{i=1}^N \omega_i^2 f_i(\phi_0). \quad (19)$$

Because our perturbed stimuli $\mathbf{s}_-$ and $\mathbf{s}_+$ are assumed to be very close to the reference $\mathbf{s}_0$, we will assume the variance of the response to these stimuli is also equal to the same σ^2 in Equation 19. The expected value of responses to stimulus $\mathbf{s}_\pm = (\phi_0 \pm d\phi, c)^T$ is given by

$$\mu_\pm = \psi(c) \sum_{i=1}^N \omega_i f_i(\phi_0 \pm d\phi), \quad (20)$$

and from Equations 18 through 20, we obtain an expression for the well-known psychophysical quantity

$$d'(\phi_0, d\phi, c) = \frac{\mu_+ - \mu_-}{\sigma} \quad (21)$$

for a perturbation of size $d\phi$

$$\begin{aligned} d'(\phi_0, d\phi, c) &= \sqrt{\psi(c)} \frac{\sum_{i=1}^N \omega_i [f_i(\phi_0 + d\phi) - f_i(\phi_0 - d\phi)]}{\sqrt{\sum_{i=1}^N \omega_i^2 f_i(\phi_0)}}, \end{aligned} \quad (22)$$

where Equation 22 is obtained readily by plugging Equations 19 and 20 into Equation 21.

Introducing the notation $\mathbf{f}_\phi = (f_1(\phi), \dots, f_N(\phi))^T$ and $\Sigma_\phi = \text{diag}[\mathbf{f}_\phi]$ and suppressing arguments, we can rewrite Equation 22 as

$$(d')^2 = 2\psi(c) \left[\frac{[\boldsymbol{\omega}^T(\mathbf{f}_{\phi_+} - \mathbf{f}_{\phi_-})]^2}{\boldsymbol{\omega}^T 2\Sigma_{\phi_0} \boldsymbol{\omega}} \right], \quad (23)$$

and we recognize the term in brackets as the ratio of variability between groups to that within groups when observations are projected onto the vector $\boldsymbol{\omega}$. The vector maximizing this ratio is known as the Fisher linear discriminant and is given by

$$\boldsymbol{\omega}_F \propto \Sigma_{\phi_0}^{-1}(\mathbf{f}_{\phi_+} - \mathbf{f}_{\phi_-}). \quad (24)$$

For small perturbations $d\phi$, the direction of the vector $\boldsymbol{\omega}_F$ does not depend on $d\phi$ because we may approximate $(\mathbf{f}_{\phi_+} - \mathbf{f}_{\phi_-}) \approx \mathbf{f}'_{\phi_0} 2d\phi$, because $\phi_+ - \phi_- = 2d\phi$. Substituting Equation 24 into Equation 23 and using this approximation yields

$$d'(\phi_0, d\phi, c) = 2\sqrt{\psi(c)} d\phi \sqrt{I_F(\phi_0)}, \quad (25)$$

where

$$I_F(\phi_0) = \sum_{i=1}^N \frac{(f'_i(\phi_0))^2}{f_i(\phi_0)} \quad (26)$$

is the population Fisher information about the stimulus orientation ϕ around the reference stimulus ϕ_0 , a well-known result from population coding theory (Dayan & Abbott, 2001). Given our expression (Equation 25) for d' and using the fact that the probability of correct response ($b = 1$) in the two-alternative forced choice task is $\Phi(d'/2)$ (single interval) or $\Phi(d'/\sqrt{2})$ (two-interval), we obtain the final model

$$P(b = 1 | \mathbf{s}, \boldsymbol{\theta}, \boldsymbol{\omega}) = \Phi\left(K\sqrt{\psi(c)}d\phi\right), \quad (27)$$

where $K \propto \sqrt{I_F(\phi_0)}$ is a parameter describing population sensitivity to changes in orientation around ϕ_0 at 100% contrast.