Journal of Vision (2023) 23(2):12, 1–19

Hierarchical Bayesian perceptual template modeling of mechanisms of spatial attention in central and peripheral cuing

Zhong-Lin Lu

Division of Arts and Sciences, NYU Shanghai, Shanghai, China
Center for Neural Science and Department of Psychology, New York University, New York, NY, USA
NYU-ECNU Institute of Brain and Cognitive Science, Shanghai, China

Barbara Anne Dosher

Department of Cognitive Sciences, University of California, Irvine, CA, USA

The external noise paradigm and perceptual template model (PTM) have successfully been applied to characterize observer properties and mechanisms of observer state changes (e.g. attention and perceptual learning) in several research domains, focusing on individual level analysis. In this study, we developed a new hierarchical Bayesian perceptual template model (HBPTM) to model the trial-by-trial data from all individuals and conditions in a published spatial cuing study within a single structure and compared its performance to that of a Bayesian Inference Procedure (BIP), which separately infers the posterior distributions of the model parameters for each individual subject without the hierarchical structure. The HBPTM allowed us to compute the joint posterior distribution of the hyperparameters and parameters at the population, observer, and experiment levels and make statistical inferences at all these levels. In addition, we ran a large simulation study that varied the number of observers and number of trials in each condition and demonstrated the advantage of the HBPTM over the BIP across all the simulated datasets. Although it is developed in the context of spatial attention, the HBPTM and its extensions can be used to model data from the external noise paradigm in other domains and enable predictions of human performance at both the population and individual levels.

Introduction

Selective attention to a location in space or to an object often significantly improves performance accuracy or response times (Bashinski & Bacharach, 1980; Carrasco, 2011; Downing, 1988; Duncan, 1984; Han, Dosher, & Lu, 2003; Itti, Rees, & Tsotsos, 2005; Nissen, 1985; Posner, 1978; Posner, 1980; Shiffrin & Czerwinski, 1988; Smith & Ratcliff, 2009; Smith, Ratcliff, & Wolgang, 2004; Sperling & Dosher, 1986), although this is not true under some circumstances (Dosher & Lu, 2000b; Lu & Dosher, 2000; Shiu & Pashler, 1994; Solomon, 2002). Theoretically, the beneficial effects of selective attention have been attributed to facilitation of sensory analysis (Cheal, Lyon, & Gottlob, 1994; Corbetta, Miezin, Shulman, & Petersen, 1991; Mangun, Hillyard, & Luck, 1993), redistribution of limited resources or capacity (Bonnel & Miller, 1994; Henderson, 1996; Henderson & Macquistan, 1993; Palmer, Ames, & Lindsey, 1993; Shiffrin, 1988), and/or elimination of irrelevant stimuli (Enns & Di Lollo, 1997; Shiu & Pashler, 1994).

To reveal and identify mechanisms of attention in perceptual tasks, we combined the external noise paradigm (Barlow, 1956; Barlow, 1957; Carter & Henning, 1971; Greis & Röhler, 1970; Harmon & Julesz, 1973; Henning, Hertz, & Hinton, 1981; Parish & Sperling, 1991; Pavel, Sperling, Riedl, & Vanderbeek, 1987; Pelli, 1981; Pelli & Blakemore, 1990; Pollehn & Roehrig, 1970; Stromeyer & Julesz, 1972) with attention manipulations and developed the perceptual template model (PTM; Lu & Dosher, 1998). In this paradigm, effects of attention are measured when observers perform perceptual judgments on visual stimuli embedded in varying amounts of external noise in attended and unattended conditions. The threshold versus external noise contrast (TvC) functions in different attention conditions are then fit with the PTM to identify the mechanism(s) of attention.

Citation: Lu, Z.-L., & Dosher, B. A. (2023). Hierarchical Bayesian perceptual template modeling of mechanisms of spatial attention in central and peripheral cuing. *Journal of Vision*, *23*(2):12, 1–19, https://doi.org/10.1167/jov.23.2.12.

https://doi.org/10.1167/jov.23.2.12 Received July 13, 2022; published February 24, 2023 ISSN 1534-7362

Since its initial development, the external noise paradigm and the PTM have not only been applied to investigate mechanisms of attention (Baek, Dosher, & Lu, 2021; Dosher & Lu, 2000a; Dosher & Lu, 2000b; Dosher & Lu, 2013; Han et al., 2003; Hetley, Dosher, & Lu, 2014; Ling, Liu, & Carrasco, 2006; Ling, Liu, & Carrasco, 2009; Lu & Dosher, 1998; Lu & Dosher, 2000; Lu & Dosher, 2005; Lu, Lesmes, & Dosher, 2002; Lu, Tse, Dosher, Lesmes, Posner, & Chu, 2009; Luzardo & Yeshurun, 2021; Pratte, Ling, Swisher, & Tong, 2013), but also applied to characterize observer inefficiencies (Lu & Dosher, 1999; Lu & Dosher, 2008), mechanisms of perceptual learning (Chung, Levi, & Tjan, 2005; Dosher & Lu, 1998; Dosher & Lu, 1999; Dosher, Liu, & Lu, 2005; Gold, Sekuler, & Bennett, 2004; Lu, Chu, & Dosher, 2006; Lu, Chu, Dosher, & Lee, 2005a; Lu, Chu, Dosher, & Lee, 2005b; Lu & Dosher, 2004a; Lu & Dosher, 2009; Maehara & Goryo, 2007; Solomon & Tyler, 2017; Streeter, 2011; Tjan, Chung, & Levi, 2002; Xie & Yu, 2018), benefits of action video game play (Bejjanki et al., 2014), visual working memory (Chu, Dosher, Najima, & Lu, 2011; Dosher, Liu, & Lu, 2005; Najima, Dosher, Chu, & Lu, 2011; Park, Ichinose, Park, & Tadin, 2017), second-order perception (Manahilov, Simpson, & Calvert, 2005), pattern discrimination (Goris, Putzeys, Wagemans, & Wichmann, 2013), sensory adaptation (Barbot, Huxlin, Tadin, & Yoon, 2017; Dao, Lu, & Dosher, 2006), intra-saccadic suppression (Guez, Morris, & Krekelberg, 2013; Watson & Krekelberg, 2011), response times in perceptual decision making (Ludwig & Davies, 2011), mechanisms of top-down feedback in animal studies (Ding et al., 2022), perceptual effects of acute alcohol intake (Zhang et al., 2022), development and aging (Bower & Andersen, 2012; Bower, Watanabe, & Andersen, 2013; DeLoss, Watanabe, & Andersen, 2014; Jeon, Maurer, & Lewis, 2012; Jeon, Maurer, & Lewis, 2014), perceptual deficits in clinical conditions (Chan, Lee, Cheung, & Chow, 2011; Park, Schauder, Zhang, Bennetto, & Tadin, 2017; Wagner, Manahilov, Loffler, Gordon, & Dutton, 2010; Webster, Dickinson, Battista, McKendrick, & Badcock, 2012; Xu, Lu, Qiu, & Zhou, 2006), and visual rehabilitation (Cavanaugh et al., 2015; Huang, Lu, & Zhou, 2009; Opoku-Baah, Hou, & Wallace, 2020; Yan et al., 2015). The PTM has also been elaborated to model discrimination of non-orthogonal targets (Jeon, Lu, & Dosher, 2009), tuning of the perceptual template (Dosher, Liu, Blair, & Lu, 2004; He et al., 2020; Hu et al., 2021; Lu & Dosher, 2001; Lu & Dosher, 2004b; Lu, Jeon, & Dosher, 2004), binocular combination (Huang, Chen, Hou, & Lu, 2016; Zhang et al., 2021). It has also been extended from a single-channel to a multichannel model for spatial vision (Chen et al., 2014; Hou, Lu, & Huang, 2014).

The original PTM consists of four observer parameters (template gain β, exponent of the nonlinear transducer γ, standard deviation of the internal additive noise N_a, and proportional constant of multiplicative noise N_m) and additional observer state-dependent parameters associated with, for example, attention (internal additive noise reduction A_a, external noise exclusion A_f, and multiplicative internal noise reduction A_m). So far, all existing applications of the PTM have focused at the individual level. Typically, the maximum likelihood or least-squares procedure is used to fit variants of the PTM that consists of different numbers of observer state-dependent parameters (e.g. a single A_a, a single A_f, or both A_a and A_f) to each individual observer's TvC functions or psychometric functions in multiple external noise conditions, and a nested model comparison procedure is used to identify the best fitting model and therefore the associated mechanism(s) for the individual (Dosher & Lu, 2000b; Lu & Dosher, 1998; Lu & Dosher, 2013). Occasionally, bootstrap procedures are used to estimate the variabilities of the best-fitting model parameters (Lu & Dosher, 2013). Although they provide excellent point estimates of the best fitting PTM parameters at the individual level, and often the results in different individuals are similar, these modeling procedures treat data from each individual separately without explicitly capitalizing on potential regularities of model parameters across individuals. In addition, they are not designed for statistical inference at the populational level.

Building on the success of the original PTM as a generative model of human performance, the goal of the current study is to develop a new hierarchical Bayesian perceptual template model (HBPTM) to model the trial-by-trial data from all individuals and conditions in an external noise experiment within a single hierarchical structure. With hyperparameters and parameters at multiple levels and conditional dependencies that share information across levels, hierarchical Bayesian models have been developed to compute the joint posterior distribution of all the hyperparameters and parameters, capture their statistical relationships, and enable statistical inferences at multiple levels in a wide range of applications (Ahn, Krawitz, Kim, Busemeyer, & Brown, 2011; Kruschke, 2015; Lee, 2006; Lee, 2011; Lee & Mumford, 2003; Merkle, Smithson, & Verkuilen, 2011; Molloy et al., 2018; Molloy, Bahg, Lu, & Turner, 2019; Rouder & Lu, 2005; Rouder, Sun, Speckman, Lu, & Zhou, 2003), whereas the PTM provides an excellent likelihood function that relates model parameters to trial-by-trial performance. Here, we developed an HBPTM by incorporating the PTM into a hierarchical Bayesian model (HBM) framework to model the data from a published spatial cuing study of attention (Lu & Dosher, 2000), and compared the performance of the HBPTM to that of a Bayesian Inference Procedure (BIP), which separately infers the posterior distributions of the model parameters for each individual observer without the hierarchical structure of the population level hyperparameters. In addition, we ran a large simulation study that varied the number of simulated observers and number of trials in each condition and demonstrated the advantage of the HBPTM over the BIP across all the simulated datasets.

Hierarchical Bayesian perceptual template model

In this section, we first briefly describe the spatial cuing study of attention (Lu & Dosher, 2000). We then describe the PTM as the likelihood function that relates parameters of each individual observer to their trial-by-trial performance in the spatial cuing study and a BIP which separately infers the posterior distributions of the PTM parameters for each individual observer. Next, we introduce the three-level HBPTM that captures the hierarchical structure of the hierarchical experimental design, model estimation procedure, and alternative models.

The spatial cuing study

The spatial cuing study (Lu & Dosher, 2000) consisted of two experiments. In both experiments, the stimulus display consisted of four simultaneously presented "T"-like pseudo-characters at four spatial locations, each of which could occur in one of four possible orientations (Figure 1a). Observers were either precued or simultaneously cued with an arrow in the center of the display (endogenous or central cuing) or near the target location (exogenous or peripheral cuing) to identify the orientation of the "T"-like pseudo-character at one of the locations (up, down, right, or left; Figures 1b, c). The cue-target onset asynchrony (CTOA) was either 200 ms or 0 ms, respectively, for the attended and unattended conditions. The pseudo-characters were embedded in systematically varying amount of external noise. A total of eight external noise levels, with contrast standard deviations 0, 0.02, 0.04, 0.08, 0.12, 0.16, 0.24, and 0.32 were tested. The method of constant stimuli with nine appropriately placed pseudo-character contrasts was used to measure the psychometric functions for pseudo-character identification at each external noise level in both precuing and simultaneous cuing conditions in both experiments. Three undergraduate and two graduate students from the University of Southern California with normal or corrected-to-normal vision participated in the study. The undergraduate students and one graduate student were paid subjects. One graduate student volunteered for the study. All five finished the endogenous (central) cuing experiment, and three of them finished the exogenous (peripheral) cuing experiment. The three observers who finished both experiments were tested in all 288 experimental conditions, and the two observers who only finished the central cuing experiment were tested in 144 conditions, with 40 trials in each combination of central/peripheral, attended/unattended, external noise level, and signal contrast condition (see Lu & Dosher, 2000, for details about the stimuli, apparatus, and procedures).

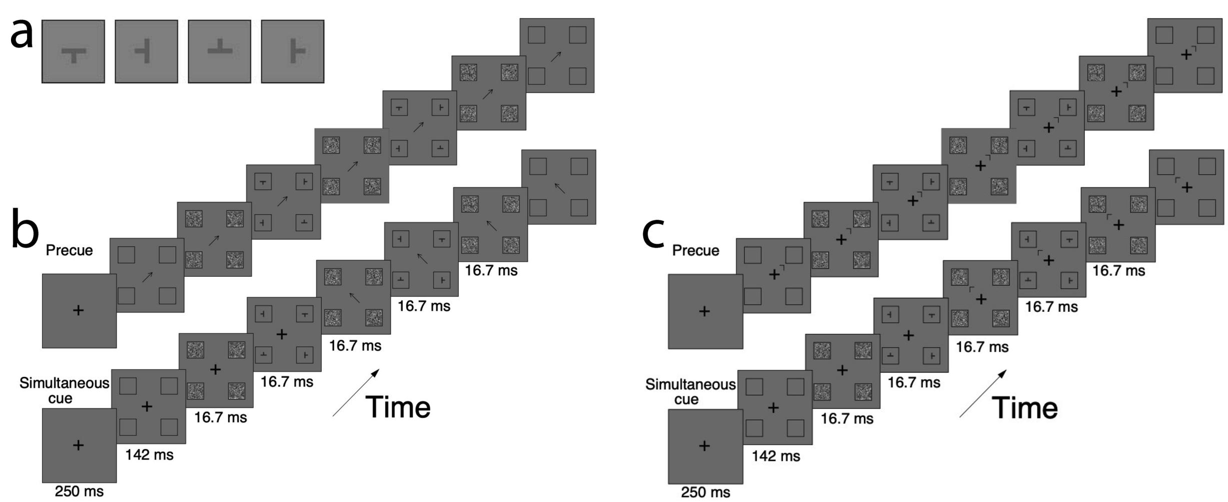

Figure 1. Sample stimuli (**a**), and stimulus sequence and layout in the (**b**) central cuing and (**c**) peripheral cuing experiment in (Lu & Dosher, 2000). Attention is manipulated by pre-cuing (CTOA = 200 ms) compared with simultaneous (CTOA = 0 ms) location cues.

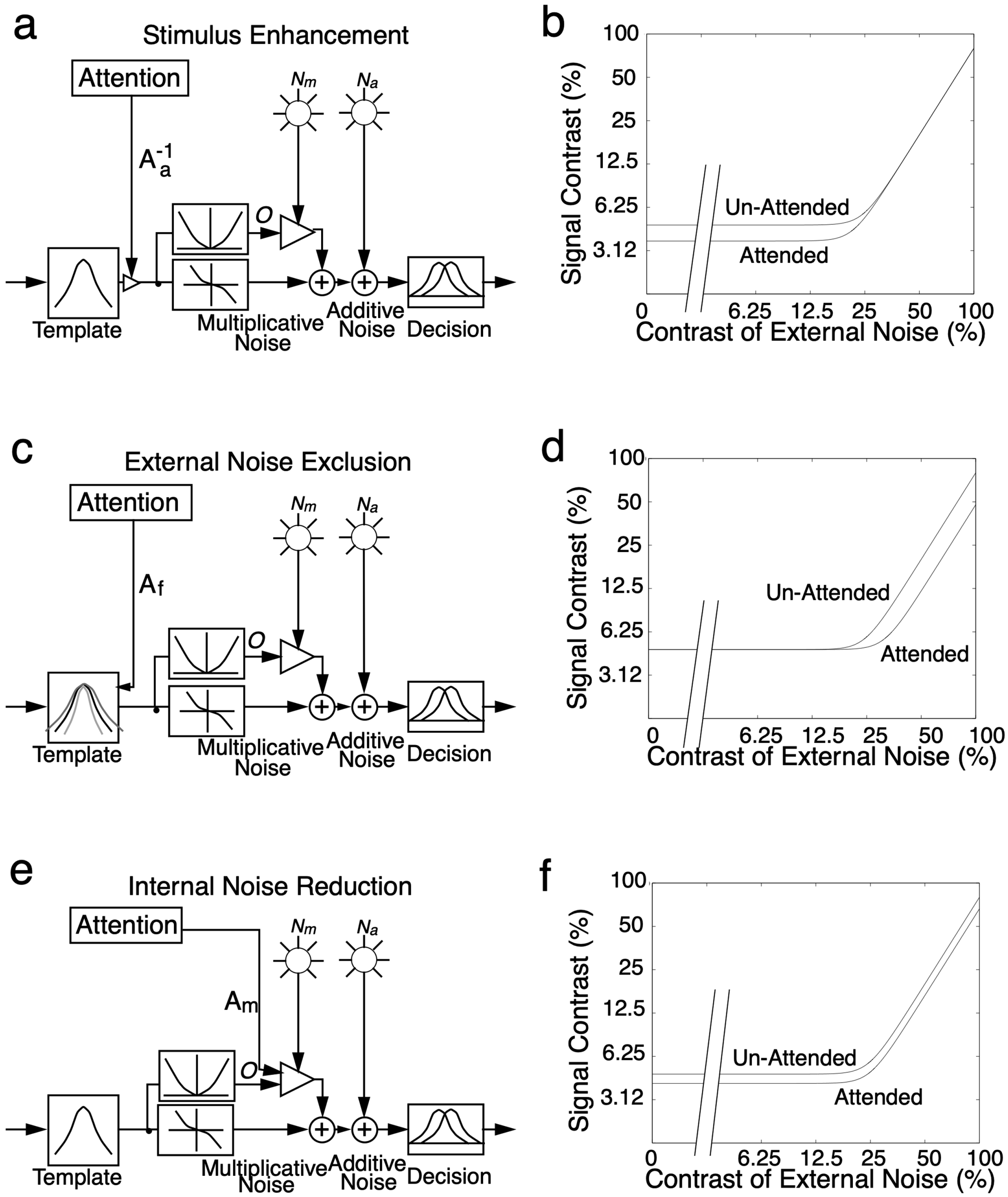

Figure 2. Three mechanisms of attention and their signature effects on threshold vs external noise contrast (TvC) functions.

Likelihood function

The likelihood function defines the probability of the observed trial-by-trial data as a function of the parameters in a model. In the PTM (Figures 2a, c, e), signal discriminability d' in condition (c, N_{ext}, cue), where c is the pseudo-character signal contrast, N_{ext} is the standard deviation of external noise contrast, and cue is either pre or simultaneous, is a function of six model parameters (N_a, N_m, β, γ, A_a, and A_f) and two stimulus parameters (c and N_{ext}; Dosher & Lu, 2000b; Lu & Dosher, 1998; Lu & Dosher, 2000):

$$d'(c, N_{ext}, cue, N_a, N_m, \beta, \gamma, A_a, A_f) = \frac{(\beta c)^{\gamma}}{\sqrt{\left(A_f N_{ext}\right)^{2\gamma} + N_m^2\left[(\beta c)^{2\gamma} + \left(A_f N_{ext}\right)^{2\gamma}\right] + (A_a N_a)^2}}, \tag{1}$$

where N_a is the standard deviation of the internal additive noise, N_m is the proportional constant of the multiplicative noise, β is the gain of the perceptual template, γ is the exponent of the transducer, A_a reflects internal additive noise reduction by attention, A_f reflects external noise exclusion by attention.[1] A_a and A_f depend on the cuing condition. The probability of obtaining a correct response in a single trial is:

$$p_{correct}(c, N_{ext}, cue, N_a, \ N_m, \beta, \gamma, A_a, A_f) = \int_{-\infty}^{+\infty} g(x - d'(c, N_{ext}, cue, \ N_a, \ N_m, \beta, \gamma, A_a, A_f)) \times G^3(x)\, dx, \quad (2)$$

where g(.) and G(.) are the probability density and cumulative probability functions of a standard Gaussian distribution. The probability of obtaining M correct responses from a total of T trials in a single condition is described by a binomial distribution B:

$$p(M|c, N_{ext}, cue, T, N_a, N_m, \beta, \gamma, A_a, A_f) = B(p_{correct}(c, N_{ext}, cue, N_m, \beta, \gamma, A_a, A_f), M, T). \quad (3a)$$

Figure 2 illustrates the signature performance effects of several attention mechanisms on the TvC functions, which graph contrast threshold as a function of the contrast of external noise. These functions are useful in illustrating the consequences of attention mechanisms. Specifically, *stimulus enhancement* reduces contrast thresholds in the region of zero or low external noise (see Figures 2a, b), accounting for effects of attention in the absence of external noise. Mathematically equivalent to internal additive noise reduction, it corresponds to claims of perceptual enhancement (Posner, Nissen, & Ogden, 1978). *External noise exclusion* reduces contrast thresholds in the region of high external noise (see Figures 2c, d), where there is external noise to exclude, by focusing perceptual analysis on the appropriate time, spatial region, and/or content characteristics of the signal stimulus (Dosher & Lu, 2000b; Shiu & Pashler, 1994). *Multiplicative internal noise reduction* reduces contrast thresholds throughout the entire range of external noise levels (see Figures 2e, f). In addition, measuring TvC functions at two or more criterion performance levels along the psychometric function resolves the individual contribution of each mechanism in when multiple mechanisms are involved (Dosher & Lu, 2000a; Lu & Dosher, 2000). In prior applications of the PTM, only stimulus enhancement and external noise exclusion have been observed, so these two mechanisms of attention are examined in our analysis.

In what follows, we use the notation $\boldsymbol{\theta}_{ij}$ to denote the PTM parameters for individual i in the jth test, which is the jth repetition of the whole experiment with all the stimulus contrast, external noise, and cuing conditions. $\mathbf{S}_{ijk}$, T_{ijk}, and M_{ijk} denote, respectively, the stimulus parameters (c, N_{ext}, and cue), the numbers of total trials and correct responses for individual i in the kth condition of the jth repetition, where condition k denotes each combination of c, N_{ext}, and cue. Table 1 shows a model structure with external noise exclusion in central cuing and a combination of external noise exclusion and internal noise reduction in peripheral cuing, which agrees with the previous analysis of the study (Lu & Dosher, 2000). Later, we will consider additional models in which central cuing causes both external noise exclusion and internal noise reduction, and a model in which attention has no effect. The system nonlinearity parameter γ of the PTM is assumed to be equal in central and peripheral cuing, consistent with many applications of the PTM (Dosher & Lu, 2000a; Dosher & Lu, 2000b; Lu & Dosher, 2000). (By convention, the A_a and A_f parameters are set to 1 in simultaneous, or unattended, conditions, and in both conditions if the respective attention mechanisms are not effective. A_a and $A_f < 1$ if attention improves performance.)

	Central cuing		Peripheral cuing	
	200 ms	0 ms	200 ms	0 ms
$\boldsymbol{N_a}$	$\boldsymbol{\theta}_{ij}(1)$	$\boldsymbol{\theta}_{ij}(1)$	$\boldsymbol{\theta}_{ij}(6)$	$\boldsymbol{\theta}_{ij}(6)$
$\boldsymbol{N_m}$	$\boldsymbol{\theta}_{ij}(2)$	$\boldsymbol{\theta}_{ij}(2)$	$\boldsymbol{\theta}_{ij}(7)$	$\boldsymbol{\theta}_{ij}(7)$
$\boldsymbol{\beta}$	$\boldsymbol{\theta}_{ij}(3)$	$\boldsymbol{\theta}_{ij}(3)$	$\boldsymbol{\theta}_{ij}(8)$	$\boldsymbol{\theta}_{ij}(8)$
$\boldsymbol{\gamma}$	$\boldsymbol{\theta}_{ij}(4)$	$\boldsymbol{\theta}_{ij}(4)$	$\boldsymbol{\theta}_{ij}(4)$	$\boldsymbol{\theta}_{ij}(4)$
$\boldsymbol{A_f}$	$\boldsymbol{\theta}_{ij}(5)$	1	$\boldsymbol{\theta}_{ij}(9)$	1
$\boldsymbol{A_a}$	1	1	$\theta_{ij}(10)$	1

Table 1. Correspondence of PTM and HBPTM parameters.

We express the probability of obtaining M_{ijk} correct responses in T_{ijk} trials as:

$$p(M_{ijk}|\mathbf{S}_{ijk}, T_{ijk}, \boldsymbol{\theta}_{ij}) = B(p_{correct}(\mathbf{S}_{ijk}, \ \boldsymbol{\theta}_{ij}), M_{ijk}, T_{ijk}). \quad (3b)$$

The likelihood of obtaining the entire dataset for a given set of parameters $\boldsymbol{\theta}_{1:I,1:J}$, is:

$$p(\boldsymbol{M}_{1:I,1:J,1:K}|\mathbf{S}_{1:I,1:J,1:K}, \boldsymbol{T}_{1:I,1:J,1:K}, \boldsymbol{\theta}_{1:I,1:J}) = \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K_i} p(M_{ijk}|\mathbf{S}_{ijk}, T_{ijk}, \boldsymbol{\theta}_{ij}). \quad (4)$$

In this study, we set J = 1 because all the individuals only repeated the experiment once and we want to estimate the PTM parameters from the whole experiment, k runs through all the (c, N_{ext}, and cuing) combinations for each individual, with $K_i = 288$ for the three observers who participated in both the central and peripheral cuing experiments, and $K_i = 144$ for the two

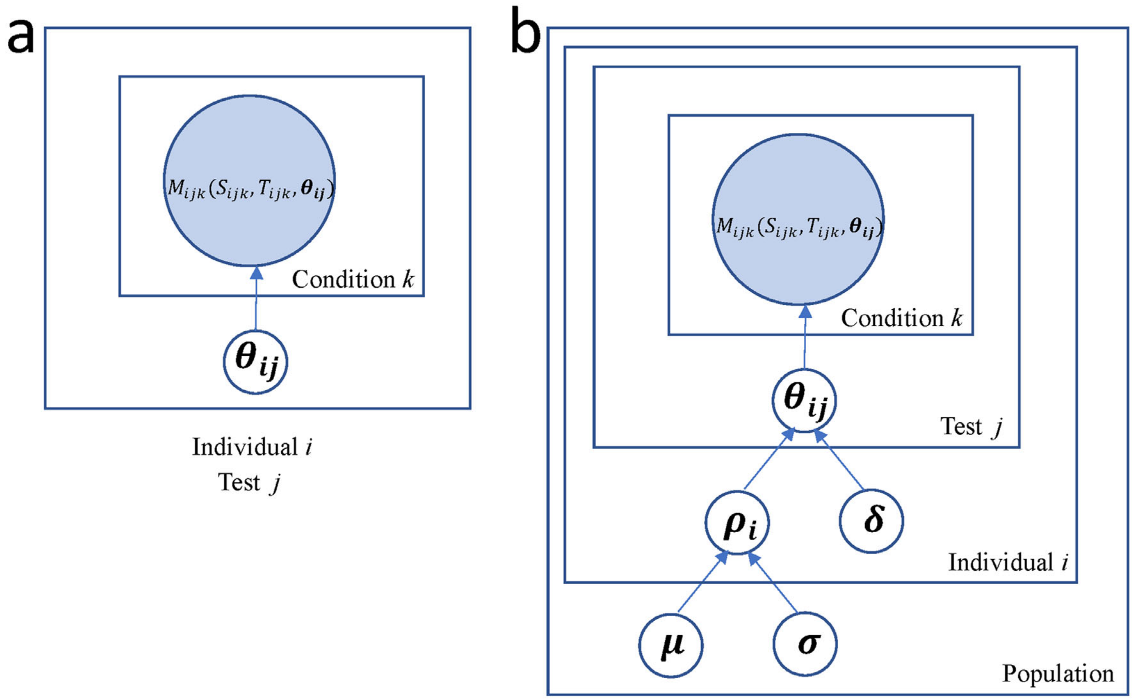

Figure 3. (**a**) The Bayesian Inference Procedure (BIP) that computes the posterior distribution of the PTM parameters for each observer independently. (**b**) A three-level hierarchical Bayesian perceptual template model (HBPTM) of spatial attention in central and peripheral cuing across multiple individuals, tests, and conditions. At the population level, $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ are the mean and standard deviation hyperparameters of the population. At the individual level, $\boldsymbol{\rho}_i$ and $\boldsymbol{\delta}$ are the mean and standard deviation hyperparameters of individual *i*, and $\boldsymbol{\delta}$ is the same across all the individuals. At the test level, $\boldsymbol{\theta}_{i,j}$ are the PTM parameters of individual *i* in test *j*. $\mathbf{S}_{ijk}$, T_{ijk}, and M_{ijk} are the stimulus condition, and numbers of total and correct responses for individual *i* in the *k*th condition of the *j*th test. In the main model, the parameters and hyperparameters are 10-dimensional.

observers who only participated in the central cuing experiment.

Bayesian inference procedure

In the BIP (Figure 3a), we treat all the observers independently in estimating PTM parameters for each individual. We can apply Bayes rule to each observer's likelihood function in Equation 4 directly to compute the posterior distributions of $\boldsymbol{\theta}_{ij}$ for each of the observers in all the tests:

$$p\left(\boldsymbol{\theta}_{i,1:J} | \boldsymbol{M}_{i,1:J,1:K} \boldsymbol{r}_{i,1,1:K}, \mathbf{S}_{i,1:J,1:K}, \boldsymbol{T}_{i,1:J,1:K}\right) = \frac{\prod_{j=1}^{J} \prod_{k=1}^{K_i} p(M_{ijk}|\mathbf{S}_{ijk}, T_{ijk}, \boldsymbol{\theta}_{ij}) p_0(\boldsymbol{\theta}_{ij})}{\int \prod_{j=1}^{J} \prod_{k=1}^{K_i} p(M_{ijk}|\mathbf{S}_{ijk}, T_{ijk}, \boldsymbol{\theta}_{ij}) p_0(\boldsymbol{\theta}_{i,j}) d\boldsymbol{\theta}_{i,1:J}}, \quad (5)$$

with uniform priors

$$p_0\left(\boldsymbol{\theta}_{ij}\right) = \mathcal{U}\left(\boldsymbol{\theta}_{ij,min}, \boldsymbol{\theta}_{ij,max}\right), \quad (6)$$

where the boundaries of the priors (Table 2) reflect prior knowledge about theoretical and empirical boundaries of the PTM parameters (Dosher & Lu, 2000a; Klein & Levi, 2009; Lu & Dosher, 2008).

	$\boldsymbol{\theta}_{ij}(1)$	$\boldsymbol{\theta}_{ij}(2)$	$\boldsymbol{\theta}_{ij}(3)$	$\boldsymbol{\theta}_{ij}(4)$	$\boldsymbol{\theta}_{ij}(5)$	$\boldsymbol{\theta}_{ij}(6)$	$\boldsymbol{\theta}_{ij}(7)$	$\boldsymbol{\theta}_{ij}(8)$	$\boldsymbol{\theta}_{ij}(9)$	$\boldsymbol{\theta}_{ij}(10)$
Min	0	0	0	0	0	0	0	0	0	0
Max	1	1	10	3	2	1	1	10	2	2

Table 2. Boundaries of the $\boldsymbol{\theta}_{ij}$ distributions.

Three-level hierarchy

Whereas the BIP can separately infer the posterior distributions of the PTM parameters for each individual observer, we developed a three-level HBPTM (Figure 3b) to capture the hierarchical structure of the experimental design and jointly infer the posterior distributions of the hyperparameters and parameters of the model across all observers. Hyperparameters $\boldsymbol{\eta}$ are used to characterize distributions of observer properties (template gain, transducer nonlinearity, internal additive and multiplicative noise, and effects of attention) at the population level, hyperparameters $\boldsymbol{\tau}_i$ are used to characterize them for individual observer i, and parameters $\boldsymbol{\theta}_{i,j}$ are used to characterize them for individual observer i in the jth test (Table 3). These

	Central cuing		Peripheral cuing	
	200 ms	0 ms	200 ms	0 ms
N_a	$\eta(1)/\tau_i(1)$	$\eta(1)/\tau_i(1)$	$\eta(6)/\tau_i(6)$	$\eta(6)/\tau_i(6)$
N_m	$\eta(2)/\tau_i(2)$	$\eta(2)/\tau_i(2)$	$\eta(7)/\tau_i(7)$	$\eta(7)/\tau_i(7)$
β	$\eta(3)/\tau_i(3)$	$\eta(3)/\tau_i(3)$	$\eta(8)/\tau_i(8)$	$\eta(8)/\tau_i(8)$
γ	$\eta(4)/\tau_i(4)$	$\eta(4)/\tau_i(4)$	$\eta(4)/\tau_i(4)$	$\eta(4)/\tau_i(4)$
A_f	$\eta(5)/\tau_i(5)$	1	$\eta(9)/\tau_i(9)$	1
A_a	1	1	$\eta(10)/\tau_i(10)$	1

Table 3. Correspondence of PTM and HBPTM hyperparameters.

hyperparameters and parameters are related through conditional probability: parameters $\theta_{i,j}$ at the test level are conditioned on the hyperparameters at the individual level, which are in turn conditioned on the hyperparameters η at the population level:

$$p\left(\theta_{ij}\right) = p\left(\theta_{ij}|\tau_i\right) p\left(\tau_i|\eta\right) p\left(\eta\right), \quad (7)$$

	$\mu_0(1)$	$\mu_0(2)$	$\mu_0(3)$	$\mu_0(4)$	$\mu_0(5)$	$\mu_0(6)$	$\mu_0(7)$	$\mu_0(8)$	$\mu_0(9)$	$\mu_0(10)$
Min	0	0	0	0	0	0	0	0	0	0
Max	1	1	10	3	2	1	1	10	2	2

Table 5. Boundaries of the μ_0 distributions.

$$p\left(M_{1:I,1:J,1:K}|\mathbf{S}_{1:I,1:J,1:K}, T_{1:I,1:J,1:K}, \theta_{1:I,1:J}, \rho_{1:I}, \mu, \sigma, \delta\right)$$
$$= \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K_i} p\left(M_{ijk}|\mathbf{S}_{ijk}, T_{ijk}, \theta_{ij}\right) p\left(\theta_{ij}|\tau_i\right)$$
$$\mathcal{N}_t\left(\tau_i, \rho_i, \delta\right) p\left(\rho_i|\eta\right) \mathcal{N}_t\left(\eta, \mu, \sigma\right) p\left(\mu\right) p\left(\sigma\right) p\left(\delta\right). \quad (10)$$

To compute the joint posterior distribution of all the parameters and hyperparameters in the HBPTM, we apply Bayes rule:

$$p\left(\theta_{1:I,1:J}, \rho_{1:I}, \mu, \sigma, \delta|M_{1:I,1:J,1:K} r_{1:I,1,1:K}, \mathbf{S}_{1:I,1:J,1:K}, T_{1:I,1:J,1:K}\right)$$
$$= \frac{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K_i} p\left(M_{ijk}|\mathbf{S}_{ijk}, T_{ijk}, \theta_{ij}\right) p\left(\theta_{ij}|\tau_i\right) \mathcal{N}_t\left(\tau_i, \rho_i, \delta\right) p\left(\rho_i|\eta\right) \mathcal{N}_t\left(\eta, \mu, \sigma\right) p_0\left(\mu\right) p_0\left(\sigma\right) p_0\left(\delta\right)}{\int \prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{K_i} p\left(M_{ijk}|\mathbf{S}_{ijk}, T_{ijk}, \theta_{ij}\right) p\left(\theta_{ij}|\tau_i\right) \mathcal{N}_t\left(\tau_i, \rho_i, \delta\right) p\left(\rho_i|\eta\right) \mathcal{N}_t\left(\eta, \mu, \sigma\right) p_0\left(\mu\right) p_0\left(\sigma\right) p_0\left(\delta\right) d\theta_{1:I,1:J} d\rho_{1:I} d\mu d\sigma d\delta}, \quad (11)$$

where $p(\eta)$ is modeled as a mixture of 10-dimensional truncated Gaussian distributions $\mathcal{N}_t$ with mean μ, standard deviation σ (Table 4), which have distributions $p(\mu)$ and $p(\sigma)$:

$$p\left(\eta\right) = \mathcal{N}_t\left(\eta, \mu, \sigma\right) p\left(\mu\right) p\left(\sigma\right), \quad (8)$$

$p(\tau_i|\eta)$ is modeled as a mixture of 10-dimensional truncated Gaussian distributions $\mathcal{N}_t$ with mean ρ_i and standard deviation δ (see Table 4), which have distributions $p(\rho_i|\eta)$ and $p(\delta)$:

$$p\left(\tau_i|\eta\right) = \mathcal{N}_t\left(\tau_i, \rho_i, \delta\right) p\left(\rho_i|\eta\right) p\left(\delta\right), \quad (9)$$

Combining Equations 4 to 7, the probability of obtaining the entire dataset is:

	$\eta(1)$ $\tau_i(1)$	$\eta(2)$ $\tau_i(2)$	$\eta(3)$ $\tau_i(3)$	$\eta(4)$ $\tau_i(4)$	$\eta(5)$ $\tau_i(5)$	$\eta(6)$ $\tau_i(6)$	$\eta(7)$ $\tau_i(7)$	$\eta(8)$ $\tau_i(8)$	$\eta(9)$ $\tau_i(9)$	$\eta(10)$ $\tau_i(10)$
Min	0	0	0	0	0	0	0	0	0	0
Max	1	1	10	3	2	1	1	10	2	2

Table 4. Boundaries of the η and τ_i distributions.

with the following prior distributions of μ, σ, and δ.

$$p_0\left(\mu\right) = \mathcal{U}\left(\mu_{0,min}, \mu_{0,max}\right), \quad (12a)$$

$$p_0\left(\frac{1}{\sigma(m)^2}\right) = \Gamma\left(5000, 3\right), \quad (12b)$$

$$p_0\left(\frac{1}{\delta(m)^2}\right) = \Gamma\left(5000, 3\right), \quad (12c)$$

where $\mathcal{U}$ is a 10-dimensional uniform distribution with $\mu_{0,min}$ and $\mu_{0,max}$ specified in Table 5, and $\Gamma(5000, 3)$ is a Gamma distribution with a shape parameter of 5000 and a rate parameter of 3. $\Gamma(5000, 3)$ is used as the prior for all the σ and δ in the model. The zero minima in Tables 4 and 5 ensure that all PTM parameters are positive. The maxima reflect prior knowledge about theoretical boundaries in the PTM (Dosher & Lu, 2000a; Klein & Levi, 2009; Lu & Dosher, 2008).

Estimating the joint distribution

All analysis was conducted on a Dell computer with Intel Xeon W-2145 @ 3.70 GHz CPU (8 cores and 16 threads) and 64 GB installed memory (RAM). The BIP and HBPTM were implemented with JAGS (Plummer, 2003) in R (R Team, 2003).

The HBPTM has a total of 110 hyperparameters and parameters: 10 $\boldsymbol{\mu}$, 10 $\boldsymbol{\sigma}$, 40 $\boldsymbol{\rho}_i$, 10 $\boldsymbol{\delta}$, and 40 $\boldsymbol{\theta}_{i,j}$ ($\boldsymbol{\rho}_i$ and $\boldsymbol{\theta}_{i,j}$ are 5 dimensional for observers 4 and 5 because they only participated in the central cuing condition). We used the Markov Chain Monte Carlo (MCMC) sampling algorithm in JAGS (Plummer, 2003) to compute the joint posterior distribution of all the hyperparameters and parameters in Equation 11. Three independent MCMC chains were simulated. Each MCMC generated 15,000 kept samples (thinning ratio = 10) via a random walk process after 30,000 burn-in and 100,000 adaptation steps. The same procedure was used to compute the posterior distributions of all the 40 $\boldsymbol{\theta}_{i,j}'s$ in Equation 5 in the BIP. Convergence of each parameter was evaluated with Gelman and Rubin's diagnostic rule (Gelman & Rubin, 1992).

Alternative models

In addition to the model with external noise exclusion in central cuing and both internal additive noise reduction and external noise exclusion in peripheral cuing, we also fit an HBPTM with no attention effect, which had a total of 91 hyperparameters and parameters, and an HBPTM with internal additive noise reduction and external noise exclusion in both central and peripheral cuing, which had a total of 123 hyperparameters and parameters.

Bayesian predictive information criterion (BPIC; Ando, 2007; Ando, 2011) was used to quantify the goodness of fit to the trial-by-trial data. The BPIC quantifies the likelihood of the data based on the joint posterior distribution of the parameters of the model and penalizes model complexity.

Results

Model convergence

The between- versus within-chain variance ratios for all hyperparameters and parameters were smaller than 1.01 for the BIP and all three HBPTM models, indicating good convergence based on Gelman and Rubin's diagnostic rule (Gelman & Rubin, 1992).

Model selection

The HBPTM with external noise exclusion in central cuing and both internal additive noise reduction and external noise exclusion in peripheral cuing was significantly better than the model assuming no attention effect ($BPIC =$ 42867.4 vs 43082.3), and provided an equivalent fit to the data compared to the model with internal additive noise reduction and external noise exclusion in both central and peripheral cuing ($BPIC =$ 42860.5). The results are consistent with the conclusions of the original study (Lu & Dosher, 2000) and a related study of central cuing only (Dosher & Lu, 2000b). We report the results from the main model in subsequent sections.

Posterior distributions from the HBPTM

η Distributions

Figure 4 shows the marginal distributions of hyperparameters $\boldsymbol{\eta}$ for the main model. The mean and half width of the 95% credible interval (95% HWCI; or half-width credible interval) of the distributions are summarized in Table 6.

At the population level, the internal additive noises, $\boldsymbol{\eta}(\mathbf{1})$ and $\boldsymbol{\eta}(\mathbf{6})$, were 0.0241 ± 0.0318 and 0.0264 ± 0.0346 in the endogenous and exogenous cuing experiments; the proportional constants of multiplicative noise, $\boldsymbol{\eta}(\mathbf{2})$ and $\boldsymbol{\eta}(\mathbf{7})$, were 0.3122 ± 0.0576 and 0.3292 ± 0.0633; the template gains, $\boldsymbol{\eta}(\mathbf{2})$ and $\boldsymbol{\eta}(\mathbf{7})$, were 1.090 ± 0.061 and 1.023 ± 0.069; and the exponent of the transducer function, $\boldsymbol{\eta}(\mathbf{4})$ was 2.580 ± 0.130. Representing endogenous cuing effects on external noise at the population level, $\boldsymbol{\eta}(\mathbf{5})$ had a mean of 0.8613 and 95% HWCI of 0.0619, indicating that

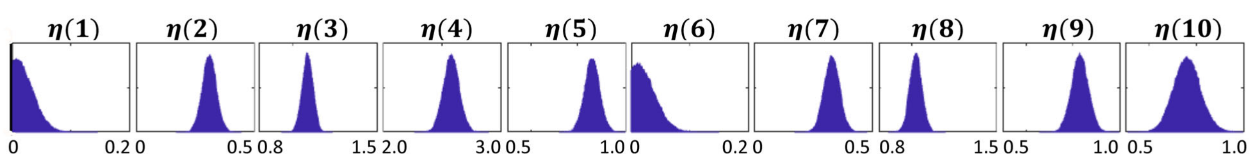

Figure 4. Marginal distributions of the population-level hyperparameters η. See Table 3 for correspondence with PTM parameters.

	$\eta(1)$	$\eta(2)$	$\eta(3)$	$\eta(4)$	$\eta(5)$	$\eta(6)$	$\eta(7)$	$\eta(8)$	$\eta(9)$	$\eta(10)$
Mean	0.0241	0.3122	1.090	2.580	0.8613	0.0264	0.3292	1.023	0.8317	0.7627
HWCI	0.0318	0.0576	0.061	0.130	0.0619	0.0346	0.0633	0.069	0.0695	0.1130

Table 6. Mean and 95% HWCI of the η distributions.

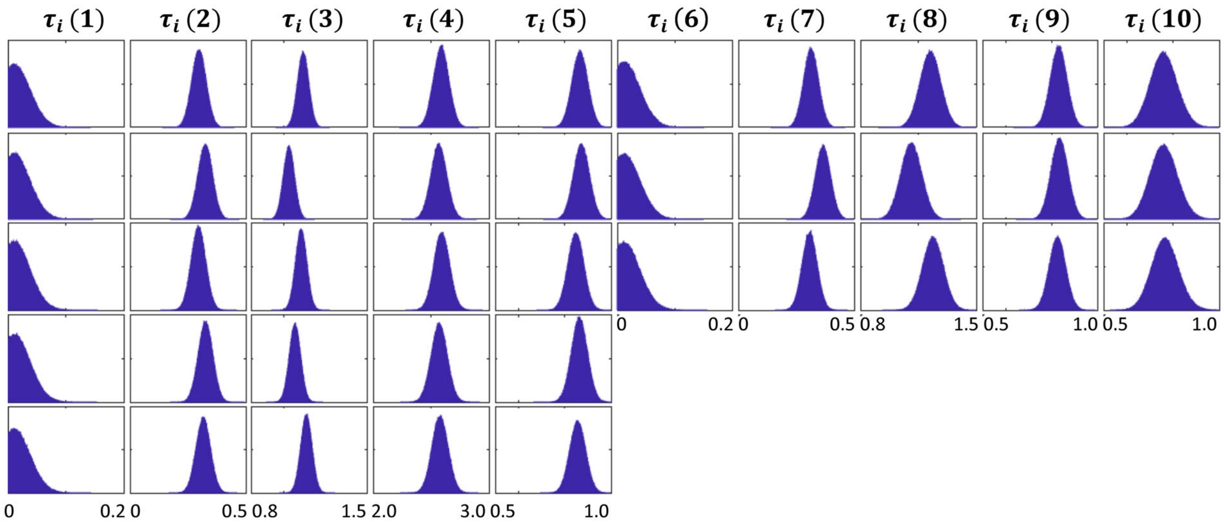

Figure 5. Marginal distributions of observer-level hyperparameters $\boldsymbol{\tau}$. Each row represents an observer. See Table 3 for correspondence with PTM parameters.

i		$\tau_i(1)$	$\tau_i(2)$	$\tau_i(3)$	$\tau_i(4)$	$\tau_i(5)$	$\tau_i(6)$	$\tau_i(7)$	$\tau_i(8)$	$\tau_i(9)$	$\tau_i(10)$
1	Mean	0.0264	0.2979	1.115	2.588	0.8661	0.0271	0.3142	1.042	0.8323	0.7606
	HWCI	0.0345	0.0617	0.068	0.133	0.0677	0.0352	0.0634	0.070	0.0712	0.1154
2	Mean	0.0263	0.3250	1.029	2.569	0.8726	0.0271	0.3655	0.9758	0.8359	0.7610
	HWCI	0.0344	0.0621	0.067	0.136	0.0683	0.0353	0.0646	0.0712	0.0735	0.1168
3	Mean	0.0263	0.2957	1.103	2.593	0.8483	0.0270	0.3082	1.051	0.8268	0.7666
	HWCI	0.0344	0.0616	0.067	0.134	0.0675	0.0351	0.0633	0.070	0.0709	0.1151
4	Mean	0.0265	0.3262	1.066	2.572	0.8638					
	HWCI	0.0345	0.0621	0.068	0.138	0.0685					
5	Mean	0.0264	0.3164	1.135	2.579	0.8564					
	HWCI	0.0345	0.0616	0.067	0.136	0.0674					

Table 7. Mean and 95% HWCI of the $\boldsymbol{\tau}$ distributions.

endogenous cuing significantly excluded external noise. Representing exogenous cuing effects on external noise at the population level, $\boldsymbol{\eta}(\mathbf{9})$ had a mean of 0.8317 and 95% HWCI of 0.0695, indicating that exogenous cuing also significantly excluded external noise. Representing exogenous cuing effects on internal additive noise at the population level, $\boldsymbol{\eta}(\mathbf{10})$ had a mean of 0.7627 and HWCI of 0.1130, indicating that exogenous cuing significantly reduced internal additive noise.

Although the pattern of results is quite consistent with Lu and Dosher (2000), the HBPTM enabled us to quantify the distributions of the observer properties and make inferences of mechanisms of attention at the population level.

τ Distributions

Figure 5 shows the marginal distributions of the observer level hyperparameters $\boldsymbol{\tau}$. The mean and 95% HWCI of the distributions are summarized in Table 7.

In general, the pattern of results at the observer level was consistent with that at the population level. Representing endogenous cuing effects on external noise at the observer level, $\boldsymbol{\tau}_i(\mathbf{5})$ ranged from 0.8483 to 0.8726, with 95% HWCI between 0.0674 and 0.0685, indicating that endogenous cuing significantly excluded external noise across all observers. The average $\boldsymbol{\tau}_i(\mathbf{5})$ across the five observers was 0.8614, with a 95% HWCI of 0.0388. In comparison, the coefficient of external noise exclusion in Lu and Dosher (2000) ranged from 0.8190 to 0.8872. Representing exogenous cuing effects on external noise at the observer level, $\boldsymbol{\tau}_i(\mathbf{9})$ ranged from 0.8268 to 0.8359, with 95% HWCI between 0.0709 and 0.0735, indicating that exogenous cuing also significantly excluded external noise across all observers. The average $\boldsymbol{\tau}_i(\mathbf{9})$ across the three observers was 0.8316, with a 95% HWCI of 0.0508. Representing exogenous cuing effects on internal additive noise at the observer level, $\boldsymbol{\tau}_i(\mathbf{10})$ ranged from 0.7606 to 0.7666, with 95% HWCI between 0.1151 and 0.1168,

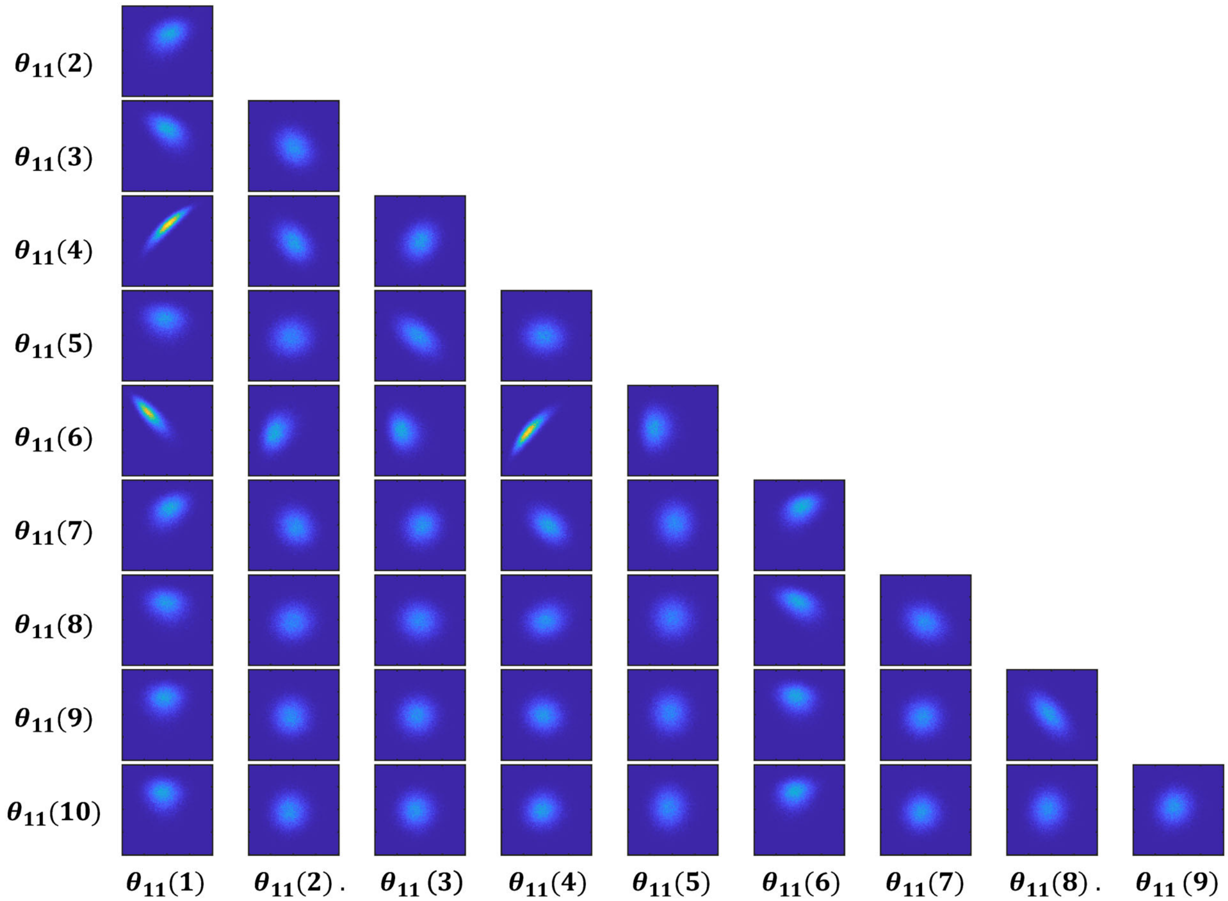

Figure 6. Distributions of pairs of $\theta_{11}(\boldsymbol{m})$ for observer 1.

	$\theta_{11}(1)$	$\theta_{11}(2)$	$\theta_{11}(3)$	$\theta_{11}(4)$	$\theta_{11}(5)$	$\theta_{11}(6)$	$\theta_{11}(7)$	$\theta_{11}(8)$	$\theta_{11}(9)$	$\theta_{11}(10)$
$\theta_{11}(1)$	1	−0.3398	0.4309	−0.9101	0.1049	0.8166	−0.3376	0.1424	−0.0456	0.0486
$\theta_{11}(2)$	−0.3398	1	0.2204	0.364	−0.0443	−0.3195	0.132	−0.0518	0.0186	−0.0253
$\theta_{11}(3)$	0.4309	0.2204	1	−0.1947	0.5027	0.176	−0.0712	0.0416	−0.0114	0.0028
$\theta_{11}(4)$	−0.9101	0.364	−0.1947	1	0.0345	−0.8834	0.369	−0.1505	0.0507	−0.0543
$\theta_{11}(5)$	0.1049	−0.0443	0.5027	0.0345	1	−0.0268	0.0106	0.0046	0.0039	−0.0046
$\theta_{11}(6)$	0.8166	−0.3195	0.176	−0.8834	−0.0268	1	−0.3441	0.4115	0.1365	−0.1559
$\theta_{11}(7)$	−0.3376	0.132	−0.0712	0.369	0.0106	−0.3441	1	0.227	−0.0199	−0.0108
$\theta_{11}(8)$	0.1424	−0.0518	0.0416	−0.1505	0.0046	0.4115	0.227	1	0.5656	−0.0543
$\theta_{11}(9)$	−0.0456	0.0186	−0.0114	0.0507	0.0039	0.1365	−0.0199	0.5656	1	−0.0909
$\theta_{11}(10)$	0.0486	−0.0253	0.0028	−0.0543	−0.0046	−0.1559	−0.0108	−0.0543	−0.0909	1

Table 8. Correlation coefficients between pairs of $\theta_{11}(\boldsymbol{m})$ for observer 1.

indicating that exogenous cuing significantly reduced internal additive noise across all observers. The average $\boldsymbol{\tau}_i(\mathbf{10})$ across the three observers was 0.7627, with a 95% HWCI of 0.1020. In comparison, the coefficient of external noise exclusion ranged from 0.8049 to 0.8348 and the coefficient of internal noise reduction ranged from 0.7628 to 0.8575 in Lu and Dosher (2000).

θ Distributions

Figure 6 shows posterior distributions of pairs of $\boldsymbol{\theta}_{11}(\mathbf{m})$ for observer 1. Table 8 lists correlation coefficients between pairs of $\boldsymbol{\theta}_{11}(\boldsymbol{m})$. Large negative correlations were found between $\boldsymbol{\theta}_{11}(\mathbf{1})$ and $\boldsymbol{\theta}_{11}(\mathbf{4})$ (−0.9101), and $\boldsymbol{\theta}_{11}(\mathbf{6})$ and $\boldsymbol{\theta}_{11}(\mathbf{4})$ (−0.8834), reflecting tradeoffs between the magnitudes of internal additive noise and the exponent of the transducer function in central and peripheral cuing, respectively. The large positive correlations between $\boldsymbol{\theta}_{11}(\mathbf{1})$ and $\boldsymbol{\theta}_{11}(\mathbf{6})$ (0.8166) reflected correlations of the magnitudes of internal additive noise in central and peripheral cueing. Large positive correlations between $\boldsymbol{\theta}_{11}(\mathbf{3})$ and $\boldsymbol{\theta}_{11}(\mathbf{5})$ (0.5027), and $\boldsymbol{\theta}_{11}(\mathbf{8})$ and $\boldsymbol{\theta}_{11}(\mathbf{9})$ (0.5656) reflected correlations between template gain and external noise exclusion. Similar results were found for the other observers (Supplementary Materials).

Comparing the performance of the HBPTM and BIP

We computed the ratios of the mean and of the 95% HWCI of the posterior distributions for each $\boldsymbol{\theta}_{ij}(\boldsymbol{m})$ from the HBPTM and BIP (i.e. the estimates of the observer level parameters). Averaged across all the observers and parameters, the ratio of the mean of the posterior distributions was 0.96 ± 0.27 (mean ± SD), indicating that the expected values of the posterior distributions of the parameters from the two methods were essentially equivalent. Averaged across all the observers and parameters, the ratio of the 95% HWCI of the posterior distributions was 0.68 ± 0.22 (mean ± SD), indicating that the 95% HWCI from the HBPTM was about 32% narrower than that from the BIP. We systematically compared the HBPTM and BIP solutions in the simulations described next.

Simulations

Methods

To evaluate the performance of the HBP relative to the BIP, we conducted a large simulation study with 15 different sample sizes: five numbers of simulated observers (3, 9, 18, 36, and 72) times three numbers of trials (10, 20, and 40 trials) in each of the 288 experimental conditions in Lu and Dosher (2000): 2 (cue condition) × 2 (cue-target SOA) × 8 (external noise levels) × 9 (stimulus contrast levels). A bootstrap procedure was used to generate the simulated datasets. In each of the 288 experimental conditions, the probability of making a correct response was randomly drawn from the probabilities of the three observers who completed both the central and peripheral cuing experiments in Lu and Dosher (2000). We then applied the HBPTM and BIP to the data.

Results

We computed the ratios of the mean and of the 95% HWCI of the posterior distributions for each of the individual observer parameters $\boldsymbol{\theta}_{ij}(\boldsymbol{m})$ from the HBPTM and BIP in each of the 15 sample sizes. Averaged across all the observers and parameters, the ratio of the mean of the posterior distributions was 1.001 ± 0.023 (mean ± SD), indicating that the expected values of the posterior distributions of the parameters from the two methods were equivalent. However, the 95% HWCI of the posterior distributions from the BIP and HBPTM were quite different.

The normalized average 95% HWCI of $\boldsymbol{\theta}_{ij}(\boldsymbol{m})$ from the HBPTM and BIP in different numbers of trials per experimental condition are shown as functions of the number of simulated observers in log10 units in Figure 7. Each 95% HWCI was normalized by the average 95% HWCI of the posterior distribution of the HBPTM solution in the 72 simulated observers and 40 trials per experimental condition dataset. As expected, the average 95% HWCI from the BIP exhibited very little variability as a function of the number of simulated observers since each observer was modeled separately, but decreased with the number of trials per experimental condition across all the parameters. On the other hand, the average 95% HWCI from the HBPTM decreased with both the number of simulated observers and the number of trials per experimental condition across all the parameters. Across different numbers of trials per experimental condition, the average 95% HWCI approached its asymptotic level between about 10 to 40 simulated observers across the different PTM parameters, suggesting that the HBPTM could benefit from increasing the number of observers up to about 40. Interestingly, the spatial attention parameters, $\boldsymbol{A}_f$ for central cuing, and $\boldsymbol{A}_f$ and $\boldsymbol{A}_a$ for peripheral cuing, reached their asymptotic levels with 20 observers.

Figure 8 shows the ratio of the average 95% HWCI of $\boldsymbol{\theta}_{ij}(\boldsymbol{m})$ from the HBPTM and BIP in different numbers of trials per experimental condition as functions of the number of simulated observers in log10 units. The early zig in the blue curves in Figures 8b, g was due to the variability of the 95% HWCI from the BIP. First, all the ratios were less than zero, indicating that the average 95% HWCI from the HBPTM was always less than that from the BIP across all the parameters and simulated sample sizes. Second, across all the parameters and number of simulated observers, the reduction of the 95% HWCI was 0.3453 log10 units (or 54.7%) when the number of trials per experimental condition per observer was 10, 0.2696 log10 units (or 46.3%) when the number of trials per experimental condition per observer was 20, and 0.2126 log10 units (or 38.7%) when the number of trials per experimental condition per observer was 40, indicating increased benefits of the HBPTM when the number of trials per experimental condition is smaller. Finally, across all the parameters and numbers of trials per experimental condition per observer, the benefit of HBPTM reached its asymptotic level when the number of observers was about 40.

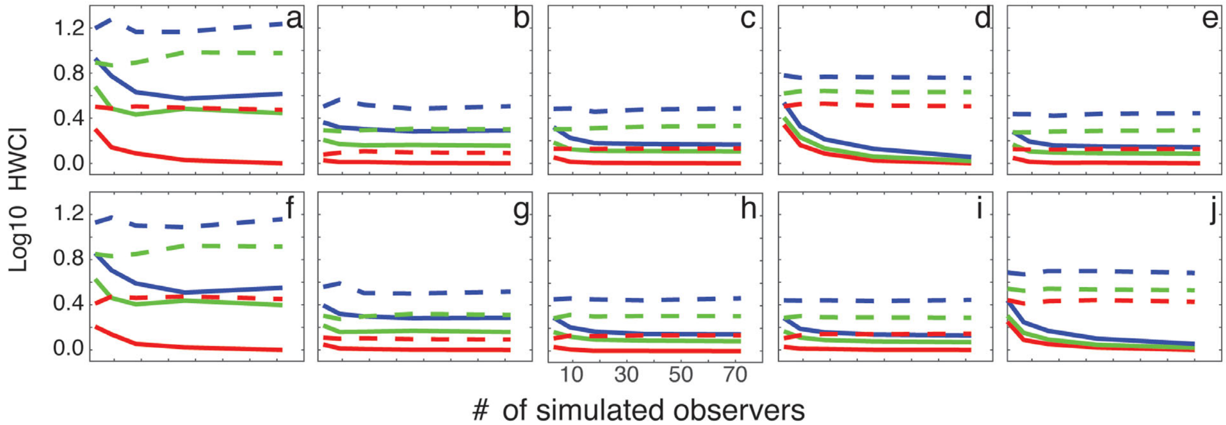

Figure 7. Normalized average 95% HWCI of $\theta_{ij}(m)$ from the HBPTM (*solid curves*) and BIP (*dotted curves*) for 10 (*blue*), 20 (*green*), and 40 (*red*) trials per experimental condition as functions of number of simulated observers, with the HWCI in log10 units. Each 95% HWCI was normalized by the average 95% HWCI of the posterior distribution of the HBPTM solution in the 72 observers and 40 trials per experimental condition dataset. HWCI's for $\theta_{ij}(m)$, $m = 1, \ldots, 10$ are shown in panels (**a**) to (**j**), corresponding with PTM parameters N_a, N_m, β, γ, and A_f for central cuing, and N_a, N_m, β, A_f, and A_a for peripheral cuing (γ is shared).

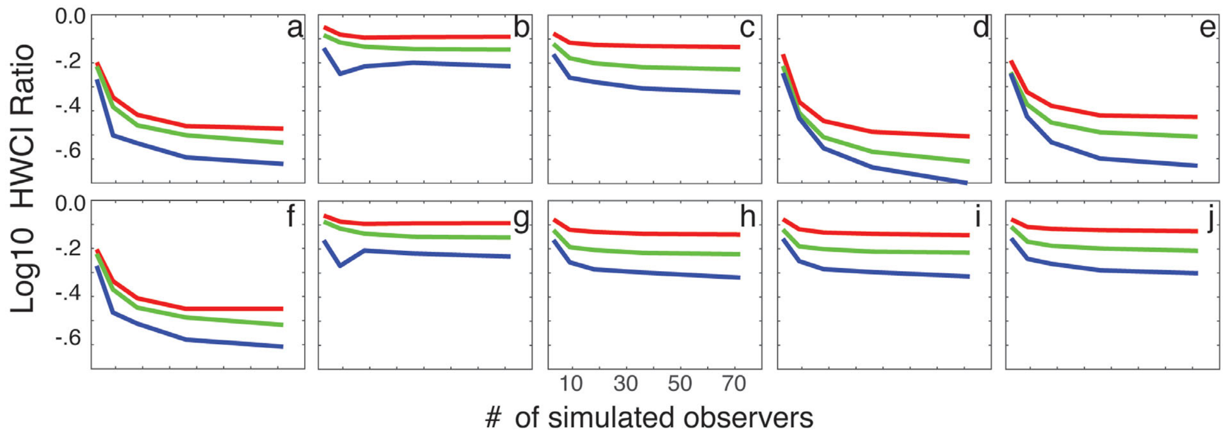

Figure 8. Ratio of the average 95% HWCI of $\theta_{ij}(m)$ from the HBPTM and BIP in 10 (*blue*), 20 (*green*), and 40 (*red*) trials per experimental condition as functions of number of simulated observers in log10 units. HWCI ratios for $\theta_{ij}(m)$, $m = 1, \ldots, 10$ are shown in panels (**a**) to (**j**), corresponding with PTM parameters N_a, N_m, β, γ, and A_f for central cuing, and N_a, N_m, β, A_f, and A_a for peripheral cuing (γ is shared).

In summary, the expected values of the posterior distributions of the parameters from the HBPTM and BIP were essentially the same. In hierarchical models, the estimated parameters at the lower levels may shrink toward the modes of the higher levels when there is not sufficient data at the lower level (Kruschke, 2015; Rouder & Lu, 2005; Rouder, Sun, Speckman, Lu, & Zhou, 2003). The observation that the HBPTM did not introduce any bias suggests that even 10 trials per experimental condition per observer was sufficient. The 95% HWCI from the HBPTM was less than that from the BIP across all the parameters and simulated sample sizes. The fewer the number of trials per experimental condition per observer, the greater the benefit of the HBPTM. The benefit of the HBPTM saturated when the number of observers was about 40.

Discussion

In this study, we developed a new HBPTM by combining the HBM and PTM to model the trial-by-trial data from all individuals and conditions in a published spatial cuing study of attention (Lu & Dosher, 2000) within a single hierarchical Bayesian structure. Using the HBMPTM, we derived the joint posterior distribution of the hyperparameters and parameters of the PTM at the populational, observer, and experiment levels, and made several statistical inferences at these levels. Specifically, the marginal distributions at the populational and observer level provided distributions of observer properties and make inferences of mechanisms of attention at both levels, and the marginal distributions at the test level revealed important covariance of the PTM parameters within each observer. Consistent with Lu and Dosher (2000), inferences at both the populational and observer levels concluded that endogenous attention significantly excluded external noise, whereas exogenous attention significantly excluded external noise and enhanced the stimulus (reduced internal additive noise). Although the pattern of results is consistent with Lu and Dosher (2000), the HBPTM enabled us to quantify the posterior distributions of the observer parameters and make inferences about the mechanisms of attention at the population level.

In addition, we compared the performance of the HBPTM to that of a BIP which separately infers the posterior distributions of the model parameters for each individual observer without the hierarchical structure and population level hyperparameters on the experimental data and in a large simulation study. We found that the expected values of the posterior distributions of the observer parameters $\boldsymbol{\theta}_{ij}(\boldsymbol{m})$ from the HBPTM and BIP were essentially the same, the 95% HWCI from the HBPTM was less than that from the BIP in the analyses of the experimental data. This pattern was also true in the simulation results, where the benefit of the HBPTM in reducing the HWCIs saturated when the number of observers was about 40.

In the current development of the HBPTM, we modeled the variance but not the covariance of the hyperparameters at the population and individual levels and revealed important covariance of the parameters at the test level. To double-check whether this is reasonable, we computed the covariance of the hyperparameters from their posterior distributions derived from the current HBPTM and did not find any significant covariances. The HBPTM can be extended by replacing the variances $\boldsymbol{\sigma}$ and $\boldsymbol{\delta}$ with covariances (Zhao, Lesmes, Dorr, & Lu, 2021; Zhao, Lesmes, Hou, & Lu, 2021).

Five observers participated in Lu and Dosher (2000). Three of them were tested in all 288 experimental conditions and two in 144 conditions, with 40 trials/condition. The dataset consisted of relatively few well practiced observers, each with many external noise and signal contrast conditions and trials. The HBPTM model converged very well, with tight distributions at the population, observer, and test levels, and no significant covariances at the observer and population levels. It is notable that the HBPTM benefitted the HWCI of the posterior distributions of PTM model parameters even in the experimental situation with a relatively smaller number of subjects and larger number of trials per condition per subject. Although the hierarchical Bayesian framework may normatively be seen as most useful for situations with more observers and smaller sample sizes per condition per observer, it nonetheless had potentially useful benefits even in the classic individual observer design such as that analyzed here (Lu & Dosher, 2000).

On the other hand, the relatively small number of well-practiced observers may or may not be representative of the general population of normal adults. A larger study with a broader sample of observers would almost surely be necessary to fully utilize the HBPTM to estimate population level distributions and covariances.

The observer models, including the PTM, specify the functional relationship between external stimuli and internal responses, as well as the decision process in human behavior. They provide the theoretical basis for generalizing the results of a particular experiment to predict the performance of the observer in other conditions and tasks based on the internal processes and intrinsic limitations of the observer. Previous observer model studies have been focused on detailed characterization of small numbers of observers (or the average over those observers), with large amount of data for each observer in many external noise levels and signal contrasts but small number of observer states (attended versus unattended) and limited range of visual stimuli (e.g. 4 pseudo characters presented at the same eccentricity). Conclusions about the generalizability of the results are based on highly consistent results across all the observers. However, the relatively small number of observers, observer states, and range of visual stimuli limit the predictive aspects of the framework. The HBPTM can be used to address this limitation. We can design studies with relatively large numbers of observers, observer states, and range of visual stimuli, collect fewer trials from each observer in a subset of the conditions (covering all the conditions across observers and with overlapping conditions between observers), and use the HBPTM to compute the joint distribution at the population, observer, and test levels. The marginal posterior hyperparameter distributions in the not-observed conditions for an observer can serve as a prior for the observer if new data are to be collected in those conditions, greatly improving the efficiency of the experiment. In addition, this important development would enable us to make predictions of human performance at the populational level and performance predictions concerning new observers or existing observers in the not-observed conditions (Lu, Zhao, Lesmes, & Dorr, 2022).

Although it is developed in the context of an existing spatial attention study, the HBPTM and its extensions can be used as a framework to model data from the external noise paradigm in other domains, including perceptual learning, visual working memory, aging, visual deficits, and visual rehabilitation. The hierarchical Bayesian (HB) framework can also be combined with various elaborated PTMs and multichannel PTM to model discrimination of non-orthogonal targets, tuning of the perceptual template, and binocular combination.

Previously, we developed a Bayesian adaptive procedure, qTvC, to measure TvC (threshold versus [external noise] contrast) functions with high efficiency (Lesmes, Jeon, Lu, & Dosher, 2006). Combining HBPTM with the adaptive testing procedure in qTvC can lead to hierarchical adaptive design optimization (HADO) of TvC measurements (Kim, Pitt, Lu, Steyvers, & Myung, 2014), making full use of informative priors across observers and conditions from the hierarchical model.

Keywords: perceptual template model (PTM), hierarchical Bayesian perceptual template model (HBPTM), spatial attention, stimulus enhancement, external noise exclusion

Acknowledgments

Supported by the National Eye Institute (R01EY17491).

Conflicts of interest/competing interests: Z.-L.L. holds intellectual property interests in visual function measurement and rehabilitation technologies, and equity interests in Adaptive Sensory Technology, Inc. (San Diego, CA, USA) and Jiangsu Juehua Medical Technology, Ltd (Jiangsu, China). B.A.D. declares no competing interests.

Statement about availability of data and code: The data and R-code are available by request.

Commercial relationships: none.
Corresponding author: Zhong-Lin Lu.
Email: zhonglin@nyu.edu.
Address: Division of Arts and Sciences, NYU Shanghai, 567 West Yangsi Road, Pudong New Area, Shanghai 200126, China; Center for Neural Science and Department of Psychology, New York University, 4 Washington Place, New York, NY 10003, USA.

Footnote

[1]Although the PTM could be used to model three mechanisms of attention, stimulus enhancement, external noise exclusion, and multiplicative noise reduction, only stimulus enhancement and external noise exclusion have been observed in spatial attention studies. We only included these two mechanisms in the current development.

References

Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J. R., & Brown, J. W. (2011). A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Journal of Neuroscience, Psychology, and Economics, 4*(2), 95–110.

Ando, T. (2007). Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models. *Biometrika, 94*(2), 443–458.

Ando, T. (2011). Predictive Bayesian model selection. *American Journal of Mathematical and Management Sciences, 31*(1-2), 13–38.

Baek, J., Dosher, B. A., & Lu, Z.-L. (2021). Visual attention in spatial cueing and visual search. *Journal of Vision, 21*(3), 1.

Barbot, A., Huxlin, K. R., Tadin, D., & Yoon, G. (2017). Neural compensation mechanisms following long-term adaptation to severe optical defects. *Journal of Vision, 17*(7), 37.

Barlow, H. B. (1956). Retinal noise and absolute threshold. *Josa, 46*(8), 634–639.

Barlow, H. B. (1957). Increment thresholds at low intensities considered as signal/noise discriminations. *The Journal of Physiology, 136*(3), 469.

Bashinski, H. S., & Bacharach, V. R. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. *Perception & Psychophysics, 28*(3), 241–248.

Bejjanki, V. R., Zhang, R., Li, R., Pouget, A., Green, C. S., & Lu, Z.-L. et al. (2014). Action video game play facilitates the development of better perceptual templates. *Proceedings of the National Academy of Sciences, 111*(47), 16961–16966.

Bonnel, A., & Miller, J. (1994). Attentional effects on concurrent psychophysical discriminations: Investigations of a sample-size model. *Perception and Psychophysics, 55*(2), 162–179.

Bower, J. D., & Andersen, G. J. (2012). Aging, perceptual learning, and changes in efficiency of motion processing. *Vision Research, 61*, 144–156.

Bower, J. D., Watanabe, T., & Andersen, G. J. (2013). Perceptual learning and aging: Improved performance for low-contrast motion discrimination. *Frontiers in Psychology, 4*, 66.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51*(13), 1484–1525.

Carter, B. E., & Henning, G. B. (1971). The detection of gratings in narrow-band visual noise. *The Journal of Physiology, 219*(2), 355.

Cavanaugh, M. R., Zhang, R., Melnick, M. D., Das, A., Roberts, M., Tadin, D., . . . Huxlin, K. R. (2015). Visual recovery in cortical blindness is limited by high internal noise. *Journal of Vision, 15*(10), 9.

Chan, M., Lee, T., Cheung, C., & Chow, Y. (2011). *Dual-channel portable amblyopia treatment system with perceptual template model. Paper presented at the 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI)*. Shanghai, China, *2011*, pp. 1305–1307.

Cheal, M., Lyon, D. R., & Gottlob, L. R. (1994). A Framework for Understanding the Allocation of Attention in Location-precued Discrimination. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *47*(3), 699–739.

Chen, G., Hou, F., Yan, F.-F., Zhang, P., Xi, J., Zhou, Y., . . . Huang, C.-B. (2014). Noise provides new insights on contrast sensitivity function. *PLoS One, 9*(3), e90579.

Chu, W., Dosher, B. A., Najima, R., & Lu, Z.-L. (2011). Resolution of Representations in Spatially Cued Attention and Visual Working Memory: A Multi-Alternative Perceptual Template Decision Model. *Journal of Vision, 11*(11), 1271.

Chung, S. T., Levi, D. M., & Tjan, B. S. (2005). Learning letter identification in peripheral vision. *Vision Research, 45*(11), 1399–1412.

Corbetta, M., Miezin, F. M., Shulman, G. L., & Petersen, S. E. (1991). Selective attention modulates extrastriate visual regions in humans during visual feature discrimination and recognition. *Ciba Found Symp, 163*, 165–175; discussion 175-180.

Dao, D. Y., Lu, Z.-L., & Dosher, B. A. (2006). Adaptation to sine-wave gratings selectively reduces the contrast gain of the adapted stimuli. *Journal of Vision, 6*(7), 6.

DeLoss, D. J., Watanabe, T., & Andersen, G. J. (2014). Optimization of perceptual learning: Effects of task difficulty and external noise in older adults. *Vision Research, 99*, 37–45.

Ding, J., Ye, Z., Xu, F., Hu, X., Yu, H., Zhang, S., . . . Lu, Z.-L. (2022). Effects of top-down influence suppression on behavioral and V1 neuronal contrast sensitivity functions in cats. *Iscience, 25*(1), 103683.

Dosher, B. A., Liu, S.-H., Blair, N., & Lu, Z.-L. (2004). The spatial window of the perceptual template and endogenous attention. *Vision Research, 44*(12), 1257–1271.

Dosher, B. A., Liu, S.-H., & Lu, Z.-L. (2005). The decay of perceptual representations in iconic memory. *Journal of Vision, 5*(8), 912.

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, 95*, 13988–13993.

Dosher, B. A., & Lu, Z.-L. (1999). Mechanisms of perceptual learning. *Vision Research, 39*, 3197–3221.

Dosher, B. A., & Lu, Z.-L. (2000a). Mechanisms of perceptual attention in precuing of location. *Vision Research, 40*, 1269–1292.

Dosher, B. A., & Lu, Z.-L. (2000b). Noise exclusion in spatial attention. *Psychological Science, 11*, 139–146.

Dosher, B. A., & Lu, Z.-L. (2005). Perceptual learning in clear displays optimizes perceptual expertise: Learning the limiting process. *Proceedings of the National Academy of Sciences, 102*(14), 5286–5290.

Dosher, B. A., & Lu, Z.-L. (2013). Mechanisms of visual attention. In C.F. Chubb, B. Dosher, Z.-L. Lu, & R. Shiffrin (Ed.), *Human Information Processing: Vision, Memory, Attention* (pp. 140–165). Washington, DC.: American Psychological Association.

Downing, C. J. (1988). Expectancy and Visual-Spatial Attention : Effects on Perceptual Quality. *Journal of Experimental Psychology. Human Perception and Performance, 14*, 188–202.

Duncan, J. (1984). Selective Attention and the Organization of Visual Information. *Journal of Experimental Psychology: General, 113*, 501–517.

Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science, 8*(2), 135–139.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472.

Gold, J. M., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science, 28*(2), 167–207.

Goris, R. L., Putzeys, T., Wagemans, J., & Wichmann, F. A. (2013). A neural population model for visual pattern detection. *Psychological Review, 120*(3), 472.

Greis, U., & Röhler, R. (1970). Untersuchung der subjektiven Detailerkennbarkeit mit Hilfe der Ortsfrequenzfilterung. *Optica Acta: International Journal of Optics, 17*(7), 515–526.

Guez, J., Morris, A. P., & Krekelberg, B. (2013). Intrasaccadic suppression is dominated by reduced detector gain. *Journal of Vision, 13*(8), 4.

Han, S., Dosher, B. A., & Lu, Z.-L. (2003). Object attention revisited: Identifying mechanisms and boundary conditions. *Psychological Science, 14*(6), 598–604.

Harmon, L. D., & Julesz, B. (1973). Masking in visual recognition: Effects of two-dimensional filtered noise. *Science, 180*(4091), 1194–1197.

He, X., Shen, M., Cui, R., Zheng, H., Ruan, X., & Lu, Z.-L. et al. (2020). The temporal window of visual processing in Aging. *Investigative Ophthalmology & Visual Science, 61*(5), 60.

Henderson, J. M. (1996). Spatial Precues Affect Target Discrimination in the Absence of Visual Noise. *Journal of Experimental Psychology. Human Perception and Performance, 22*, 780–787.

Henderson, J. M., & Macquistan, A. D. (1993). The spatial distribution of attention following an exogenous cue. *Perception & Psychophysics, 53*, 221–230.

Henning, G. B., Hertz, B. G., & Hinton, J. (1981). Effects of different hypothetical detection mechanisms on the shape of spatial-frequency filters inferred from masking experiments: I. Noise masks. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 71*(5), 574–581.

Hetley, R., Dosher, B. A., & Lu, Z.-L. (2014). Generating a taxonomy of spatially cued attention for visual discrimination: effects of judgment precision and set size on attention. *Attention, Perception, & Psychophysics, 76*(8), 2286–2304.

Hou, F., Lu, Z.-L., & Huang, C.-B. (2014). The external noise normalized gain profile of spatial vision. *Journal of Vision, 14*(13), 9.

Hu, X., Qin, Y., Ying, X., Yuan, J., Cui, R., Ruan, X., . . . Hou, F. (2021). Temporal Characteristics of Visual Processing in Amblyopia. *Frontiers in Neuroscience, 15*, 657.

Huang, C.-B., Chen, G., Hou, F., & Lu, Z.-L. (2016). Binocular Combination: Data and Binocular Perceptual Template Model. *Investigative Ophthalmology & Visual Science, 57*(12), 1512.

Huang, C.-B., Lu, Z.-L., & Zhou, Y. (2009). Mechanisms underlying perceptual learning of contrast detection in adults with anisometropic amblyopia. *Journal of Vision, 9*(11), 24.

Itti, L., Rees, G., & Tsotsos, J. K. (2005). *Neurobiology of Attention*. New York, NY: Elsevier.

Jeon, S.-T., Lu, Z.-L., & Dosher, B. A. (2009). Characterizing perceptual performance at multiple discrimination precisions in external noise. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 26*(11), B43–B58.

Jeon, S.-T., Maurer, D., & Lewis, T. L. (2012). Characterizing the mechanisms behind improvements in visual sensitivity during childhood. *Journal of Vision, 12*(9), 1366.

Jeon, S.-T., Maurer, D., & Lewis, T. L. (2014). Developmental mechanisms underlying improved contrast thresholds for discriminations of orientation signals embedded in noise. *Frontiers in Psychology, 5*, 977.

Kim, W., Pitt, M. A., Lu, Z.-L., Steyvers, M., & Myung, J. I. (2014). A hierarchical adaptive approach to optimal experimental design. *Neural Computation, 26*(11), 2465–2492.

Klein, S. A., & Levi, D. M. (2009). Stochastic model for detection of signals in noise. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 26*(11), B110–B126.

Kruschke, J. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan*. San Diego, CA: Academic Press.

Lee, M. D. (2006). A hierarchical Bayesian model of human decision-making on an optimal stopping problem. *Cognitive Science, 30*(3), 1–26.

Lee, M. D. (2011). How cognitive modeling can benefit from hierarchical Bayesian models. *Journal of Mathematical Psychology, 55*(1), 1–7.

Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 20*(7), 1434–1448.

Lesmes, L. A., Jeon, S.-T., Lu, Z.-L., & Dosher, B. A. (2006). Bayesian adaptive estimation of threshold versus contrast external noise functions: The quick TvC method. *Vision Research, 46*, 3160–3176.

Ling, S., Liu, T., & Carrasco, M. (2006). The influence of attention on motion selective channels: An equivalent noise approach. *Journal of Vision, 6*(6), 519.

Ling, S., Liu, T., & Carrasco, M. (2009). How spatial and feature-based attention affect the gain and tuning of population responses. *Vision Research, 49*(10), 1194–1204.

Lu, Z.-L., Chu, W., & Dosher, B. A. (2006). Perceptual learning of motion direction discrimination in fovea: Separable mechanisms. *Vision Research, 46*(15), 2315–2327.

Lu, Z.-L., Chu, W., Dosher, B. A., & Lee, S. (2005a). Independent perceptual learning in monocular and binocular motion systems. *Proceedings of the National Academy of Sciences, 102*(15), 5624–5629.

Lu, Z.-L., Chu, W., Dosher, B. A., & Lee, S. (2005b). Perceptual learning of Gabor orientation identification in visual periphery: Complete inter-ocular transfer of learning mechanisms. *Vision Research, 45*(19), 2500–2510.

Lu, Z.-L., & Dosher, B. A. (1998). External Noise Distinguishes Attention Mechanisms. *Vision Research, 38*, 1183–1198.

Lu, Z.-L., & Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 16*, 764–778.

Lu, Z.-L., & Dosher, B. A. (2000). Spatial attention: Different mechanisms for central and peripheral temporal precues? *Journal of Experimental Psychology: Human Perception and Performance, 26*(5), 1534.

Lu, Z.-L., & Dosher, B. A. (2001). Characterizing the spatial-frequency sensitivity of perceptual templates. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 18*(9), 2041–2053.

Lu, Z.-L., & Dosher, B. A. (2004a). Perceptual learning retunes the perceptual template in foveal orientation identification. *Journal of Vision, 4*(1), 5.

Lu, Z.-L., & Dosher, B. A. (2004b). Spatial attention excludes external noise without changing the spatial frequency tuning of the perceptual template. *Journal of Vision, 4*(10), 10.

Lu, Z.-L., & Dosher, B. A. (2005). External noise distinguishes mechanisms of attention. In *Neurobiology of Attention* (pp. 448–453): New York, NY: Elsevier.

Lu, Z.-L., & Dosher, B. A. (2008). Characterizing observer states using external noise and observer models: Assessing internal representations with external noise. *Psychological Review, 115*(1), 44–82.

Lu, Z.-L., & Dosher, B. A. (2009). Mechanisms of perceptual learning. *Learning & Perception, 1*(1), 19–36.

Lu, Z.-L., & Dosher, B. A. (2013). *Visual psychophysics: From laboratory to theory*. Cambridge, MA: MIT Press.

Lu, Z.-L., Jeon, S.-T., & Dosher, B. A. (2004). Temporal tuning characteristics of the perceptual template and endogenous cuing of spatial attention. *Vision Research, 44*(12), 1333–1350.

Lu, Z.-L., Lesmes, L. A., & Dosher, B. A. (2002). Spatial attention excludes external noise at the target location. *Journal of Vision, 2*(4), 4.

Lu, Z.-L., Tse, H. C.-H., Dosher, B. A., Lesmes, L. A., Posner, C., & Chu, W. (2009). Intra-and cross-modal cuing of spatial attention: Time courses and mechanisms. *Vision Research, 49*(10), 1081–1096.

Lu, Z.-L., Zhao, Y., Lesmes, L. A., & Dorr, M. (2022). Predicting contrast sensitivity functions with digital twins generated by Hierarchical Bayesian modelling. *Investigative Ophthalmology & Visual Science, 63*(7), 2328–2328.

Ludwig, C. J., & Davies, J. R. (2011). Estimating the growth of internal evidence guiding perceptual decisions. *Cognitive Psychology, 63*(2), 61–92.

Luzardo, F., & Yeshurun, Y. (2021). Inter-individual variations in internal noise predict the effects of spatial attention. *Cognition, 217*, 104888.

Maehara, G., & Goryo, K. (2007). Perceptual learning in monocular pattern masking: experiments and explanations by the twin summation gain control model of contrast processing. *Perception & Psychophysics, 69*(6), 1009–1021.

Manahilov, V., Simpson, W. A., & Calvert, J. (2005). Why is second-order vision less efficient than first-order vision? *Vision Research, 45*(21), 2759–2772.

Mangun, G. R., Hillyard, S. A., & Luck, S. J. (1993). Electrocortical substrates of visual selective attention. In D.E. Meyer, & S. Kornblum (Eds.), *Attention and performance 14: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 219–243). Cambridge, MA: The MIT Press.

Merkle, E. C., Smithson, M., & Verkuilen, J. (2011). Hierarchical models of simple mechanisms underlying confidence in decision making. *Journal of Mathematical Psychology, 55*(1), 57–67.

Molloy, M. F., Bahg, G., Li, X., Steyvers, M., Lu, Z.-L., & Turner, B. M. (2018). Hierarchical Bayesian analyses for modeling BOLD time series data. *Computational Brain & Behavior, 1*(2), 184–213.

Molloy, M. F., Bahg, G., Lu, Z.-L., & Turner, B. M. (2019). Individual differences in the neural dynamics of response inhibition. *Journal of Cognitive Neuroscience, 31*(12), 1976–1996.

Najima, R., Dosher, B. A., Chu, W., & Lu, Z.-L. (2011). Visual working memory performance with contrast and external noise: A load-dependent perceptual template model account. *Journal of Vision, 11*(11), 1254.

Nissen, M. J. (1985). Accessing features and objects: Is location special? In M. I. Posner, & O. S. Martin (Eds.), *Attention and Performance XI* (pp. 205–220). Hillsdale, NJ: Erlbaum.

Opoku-Baah, C., Hou, B., & Wallace, M. (2020). A multisensory perspective on low vision rehabilitation: audition facilitates visual learning of contrast detection. *Investigative Ophthalmology & Visual Science, 61*(7), 4269.

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 108–130.

Parish, D. H., & Sperling, G. (1991). Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination. *Vision Research, 31*(7-8), 1399–1415.

Park, W. J., Ichinose, M., Park, S., & Tadin, D. (2017). Perceptual inefficiencies predict individual differences in working memory both in typical adults and in schizophrenia. *Journal of Vision, 17*(10), 1110.

Park, W. J., Schauder, K. B., Zhang, R., Bennetto, L., & Tadin, D. (2017). High internal noise and poor external noise filtering characterize perception in autism spectrum disorder. *Scientific Reports, 7*(1), 1–12.

Pavel, M., Sperling, G., Riedl, T., & Vanderbeek, A. (1987). Limits of visual communication: the effect of signal-to-noise ratio on the intelligibility of American Sign Language. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 4*(12), 2355–2365.

Pelli, D. (1981). *Effects of visual noise*. (PhD dissertation). Cambridge, UK: University of Cambridge,

Pelli, D., & Blakemore, C. (1990). The quantum efficiency of vision. *Vision: Coding and Efficiency*, pp. 3–24. Cambridge, MA: Cambridge University Press.

Plummer, M. (2003). *JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Paper presented at the Proceedings of the 3rd international workshop on distributed statistical computing*. Retrieved from https://www.R-project.org/conferences/DSC-2003/.

Pollehn, H., & Roehrig, H. (1970). Effect of noise on the modulation transfer function of the visual channel. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 60*(6), 842–848.

Posner, M. I. (1978). *Chronometric explorations of mind*. Oxford, England: Lawrence Erlbaum.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Posner, M. I., Nissen, M. J., & Ogden, W. C. (1978). Attended and unattended processing modes: The role for spatial location. In N.H.L. Pick, & I.J. Saltzman (Eds.), *Modes of perceiving and processing information* (pp. 137–157). Hillsdale, NJ: Erlbaum.

Pratte, M. S., Ling, S., Swisher, J. D., & Tong, F. (2013). How attention extracts objects from noise. *Journal of Neurophysiology, 110*(6), 1346–1356.

Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review, 12*(4), 573–604.

Rouder, J. N., Sun, D., Speckman, P. L., Lu, J., & Zhou, D. (2003). A hierarchical Bayesian statistical framework for response time distributions. *Psychometrika, 68*(4), 589–606.

R Team, . (2003). *R: A language and environment for statistical computing, Version 2.9. 2*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://cran.-archive.r-project.org/bin/windows/base/old/2.9.2/.

Shiffrin, R. M. (1988). Attention. In R.C. Atkinson, & R.J. Herrnstein (Eds.), *Stevens' Handbook of Experimental Psychology, Vol. 1: Perception and Motivation; Vol. 2: Learning and Cognition* (2nd ed.) (pp. 739–811). Oxford, England: John Wiley & Sons.

Shiffrin, R. M., & Czerwinski, M. P. (1988). A model of automatic attention attraction when mapping is partially consistent. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 14*(3), 562–569.

Shiu, L.-P., & Pashler, H. (1994). Negligible Effect of Spatial Precuing on Identification of Single Digits. *Journal of Experimental Psychology. Human Perception and Performance, 20*, 1037–1054.

Smith, P. L., & Ratcliff, R. (2009). An integrated theory of attention and decision making in visual signal detection. *Psychol Rev, 116*(2), 283–317.

Smith, P. L., Ratcliff, R., & Wolgang, B. J. (2004). Attention orienting and the time course of perceptual decisions: response time distributions with masked and unmasked displays. *Vision Research, 44*, 1297–1320.

Solomon, J. A. (2002). *Covert attention does NOT affect contrast sensitivity. Paper presented at the Vision Sciences Society Annual Meeting Abstract (2nd annual meeting)*, Sarasota, Florida. Journal of Vision, *2*, 436.

Solomon, J. A., & Tyler, C. W. (2017). Improvement of contrast sensitivity with practice is not compatible with a sensory threshold account. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 34*(6), 870–880.

Sperling, G., & Dosher, B. A. (1986). Strategy and optimization in human information processing. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and performance* (Vol. 1, pp. 1–85). Hoboken, NJ: Wiley & Sons.

Streeter, K. D. (2011). A Study in Texture Segmentation: Investigating the Role of Template Retuning in Perceptual Learning.

Stromeyer, C. F., & Julesz, B. (1972). Spatial-frequency masking in vision: Critical bands and spread of masking. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 62*(10), 1221–1232.

Tjan, B., Chung, S. L., & Levi, D. (2002). Limitation of ideal-observer analysis in understanding perceptual learning. *Investigative Ophthalmology & Visual Science, 43*(13), 2916.

Wagner, D., Manahilov, V., Loffler, G., Gordon, G. E., & Dutton, G. N. (2010). Visual noise selectively degrades vision in migraine. *Investigative Ophthalmology & Visual Science, 51*(4), 2294–2299.

Watson, T., & Krekelberg, B. (2011). An equivalent noise investigation of saccadic suppression. *Journal of Neuroscience, 31*(17), 6535–6541.

Webster, K. E., Dickinson, J. E., Battista, J., McKendrick, A. M., & Badcock, D. R. (2012). Evidence for increased internal noise in migraineurs for contrast and shape processing. *Cephalalgia, 32*(2), 125–139.

Xie, X.-Y., & Yu, C. (2018). Double training downshifts the threshold vs. noise contrast (TvC) functions with perceptual learning and transfer. *Vision Research, 152*, 3–9.

Xu, P., Lu, Z.-L., Qiu, Z., & Zhou, Y. (2006). Identify mechanisms of amblyopia in Gabor orientation identification with external noise. *Vision Research, 46*, 3748–3760.

Yan, F.-F., Zhou, J., Zhao, W., Li, M., Xi, J., & Lu, Z.-L. et al. (2015). Perceptual learning improves neural processing in myopic vision. *Journal of Vision, 15*(10), 12.

Zhang, P., Guo, Y., Qiao, Y., Yan, N., Zhang, Y., Ren, W., . . . Wu, D. (2022). Acute Alcohol Intake Affects Internal Additive Noise and the Perceptual Template in Visual Perception. *Frontiers in Neuroscience, 16*, 873671.

Zhang, P., Wang, H., Ren, W., Lu, Q., Li, C., Chen, G., . . . Wang, Z. (2021). Lower Internal Additive Noise and Better Perceptual Template Characterize Binocular Contrast Sensitivity Summation. *Frontiers in Psychology, 12*, 740759.

Zhao, Y., Lesmes, L. A., Dorr, M., & Lu, Z.-L. (2021). Quantifying uncertainty of the estimated visual acuity behavioral function with hierarchical Bayesian modeling. *Translational Vision Science & Technology, 10*(12), 18.

Zhao, Y., Lesmes, L. A., Hou, F., & Lu, Z.-L. (2021). Hierarchical Bayesian modeling of contrast sensitivity functions in a within-subject design. *Journal of Vision, 21*(12), 9.