Supplement to 'Common or Distinct Attention Mechanisms for Contrast and Assimilation?'

Hope K. Snyder[1], Sean M. Rafferty[1], Julia M. Haaf[1], & Jeffery N. Rouder[2,1]

[1] University of Missouri

[2] University of California, Irvine

Author Note

This document was written in R-Markdown with code for data analysis integrated into the text. The Markdown script is open and freely available at https://github.com/PerceptionAndCognitionLab/ctx-flanker/tree/public/papers/current. The data were *born open* (Rouder, 2016) and are freely available at https://github.com/PerceptionCognitionLab/data1/tree/master/ctxIndDif/flankerMorph4

Correspondence concerning this article should be addressed to Hope K. Snyder, 205 McAlester Hall, University of Missouri, Columbia, MO 65211. E-mail: hks7w2@mail.missouri.edu

Supplement to 'Common or Distinct Attention Mechanisms for Contrast and Assimilation?'

This document is the supplement to "Common or Distinct Attention Mechanisms for Contrast and Assimilation?". It provides the specification and analysis of a hierarchical Bayesian probit model for assessing the correlation across individuals' abilities to inhibit distractors in assimilation and contrast contexts. Finally, it provides a short analysis of the inhibition effects over time.

Model Specification

Let $Y = 0, 1$ denote whether a response is "A" or "H", respectively. Further, let $Y_{ijk\ell m}$ denote the response for the ith participant, $i = 1, \ldots, I$, in the jth context type ($j = 1, 2$ for word and letter contexts, respectively), for the kth context direction ($k = 1, 2$, for contexts that promote "A" and "H" responses, respectively), for the ℓth target, $\ell = 1, \ldots, L$, and for the mth replicate, $m = 1, \ldots, M_{ijk\ell}$. The context assignments are displayed in the following table:

		k=1	k=2
		A-promoting	H-promoting
j=1	word	C_T	T_E
j=2	letter	H	A

Observations $Y_{ijk\ell m}$ are dichotomous. To model the effect of covariates in dichotomous observations, we use a probit-regression specification:

$$Y_{ijk\ell m} \stackrel{ind}{\sim} \text{Bernoulli}\left[\Phi(\mu_{ijk\ell})\right].$$

Here, Φ, the cumulative distribution function of the standard normal, is the link, and $\mu_{ijk\ell} \in (\infty, \infty)$ is the combined effect of people, conditions, and the target.

To model individual inhibition effects, we additively decompose $\mu_{ijk\ell}$ into the effect of the target for a particular participant, $\gamma_{i\ell}$, the individual's *assimilation* effect when the background context is a word, α_i and the individual's *contrast* effect when the background context is a letter frame, β_i. The decomposition is:

$$\mu_{ijk\ell} = \gamma_{i\ell} + v_j x_k \alpha_i + (1 - v_j) x_k \beta_i.$$

The quantities v_j and x_k are indicators of the context type and direction, as follows:

$$v_j = \begin{cases} 0 & \text{if} \quad j = 2 \text{ (letters)} \\ 1 & \text{if} \quad j = 1 \text{ (word)} \end{cases} \qquad x_k = \begin{cases} \frac{-1}{2} & \text{if} \quad k = 1 \text{ (A-promoting)} \\ \frac{1}{2} & \text{if} \quad k = 2 \text{ (H-promoting)} \end{cases}$$

Prior distributions are needed for $\alpha_i, \beta_i,$ and $\gamma_{i\ell}$. We start with $\gamma_{i\ell}$, the effect of the target for a particular participant. These parameters are not of primary concern, and we use a broad hierarchical prior given by: $\gamma_{i\ell} \sim \mathrm{N}(\nu_\ell, \delta_\ell)$ with hyper priors of $\nu_\ell \sim \mathrm{N}(0, 1)$ and $\delta_\ell \sim$ Inverse Gamma$(.5, .01)$.

The effects of interest are an individual's assimilation effect, α_i, and an individual's contrast effect, β_i. We follow the development in Rouder et al. (2007) who describe Bayesian analyses in these settings. The joint prior over these parameters is

$$\begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} = \mathrm{N}(\boldsymbol{\lambda}, \boldsymbol{\Sigma}),$$

where $\boldsymbol{\lambda}$ is the vector (μ_α, μ_β) and $\boldsymbol{\Sigma} = \begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,2} \\ \Sigma_{2,1} & \Sigma_{2,2} \end{pmatrix}$ is a covariance matrix.

Following Rouder et al. (2007), the hyperpriors for μ_α and μ_β are given by

$$\mu_\alpha \sim N(0, 1),$$

$$\mu_\beta \sim N(0, 1).$$

The hyperprior for $\boldsymbol{\Sigma}$ is

$$\boldsymbol{\Sigma} \sim \text{Inverse Wishert}(3, \Omega),$$

where $\Omega = \begin{vmatrix} .05 & 0 \\ 0 & .05 \end{vmatrix}$. A key property of the above specification is that population covariance in $\boldsymbol{\Sigma}$ is an explicit parameter given by $\Sigma_{1,2} = \rho \sigma_a \sigma_b$. Hence, a population correlation, ρ is well defined, and subsequent inferences generalize to new data from new participants.

The choices we make above are all reasonable given the expected degree of variability in proportions. The standard normal prior in the probit space on μ_β and μ_α corresponds to flat priors in the space of proportions, and the shape value of 3 in the Inverse Wishert corresponds to a flat prior on the population-level correlation coefficient. The only substantive choice is the scalar value of .05 of the Inverse Wishert. This small value of precision corresponds to a vaguely informative prior on effects. Additional details are provided in Rouder et al. (2007).

Analyses Methods

The estimation of probit-transformed parameters follows from Albert and Chib (1995). Accordingly, a new set of latent variables is introduced for each observation, and we denote them $w_{ijk\ell m}$. These variables are defined as follows:

$$y_{ijk\ell m} = \begin{cases} 0 & \text{then } w_{ijk\ell m} < 0 \\ 1 & \text{then } w_{ijk\ell m} \geq 0 \end{cases}$$

Without loss of generality, the variable $w_{ijk\ell m}$ is assumed to be distributed as a normal with a variance of one and the mean of $\mu_{ijk\ell}$. Thus, the parameters of interest at this level are $w_{ijk\ell m}$ and $\mu_{ijk\ell}$. It is easy to compute the posterior of $\boldsymbol{w}$ given the vector of $\boldsymbol{\mu}$ and, conversely, it is easy to compute the posterior of $\boldsymbol{\mu}$ given the $\boldsymbol{w}$'s. Hence, the marginals of each may be found by Markov chain Monte Carlo (MCMC) sampling. The conditional posterior distribution of $\boldsymbol{w}$ is a truncated normal:

$$w_{ijk\ell m} | y_{ijk\ell m} \sim \begin{cases} \mathrm{N}_-(\mu_{ijk\ell}, 1) & \text{if } y_{ijk\ell m} = 0 \\ \mathrm{N}_+(\mu_{ijk\ell}, 1) & \text{if } y_{ijk\ell m} = 1, \end{cases}$$

where N_+ and N_- denote normal distributions truncated at zero from below and above, respectively.

Posterior distributions for remaining parameters may also be sampled with MCMC. In all cases, conditional posterior distributions of parameters may be derived directly from the proportional form of Bayes rule (Jackman, 2009; Rouder & Lu, 2005). Priors were chosen to be conjugate, and consequently, posterior distributions may be sampled from known distributions in Gibbs steps.

Critical parameters are individual estimates of contrast and assimilation, and the population-level correlation of these abilities. Individual estimates are parameters α_i and β_i. Posterior means and posterior standard deviations of these parameters are shown as points and ellipses, respectively, in Figure 4A in the main paper. The most critical parameter is the population correlation. On each iteration of the MCMC chain, we calculated $\rho = \Sigma_{1,2}/\sqrt{\Sigma_{1,1} \times \Sigma_{2,2}}$. The prior and posterior distribution of ρ is shown in Figure 4B. A full description of the results may be found in the main paper.

Inhibition Effects Over Time

It may be of interest to see how assimilation and contrast effects develop over time. A short examination of time effects was conducted. First, the data for the start of each experimental block was investigated. Participants appeared acclimated to the task within the first ten to twenty trials of each block. Often, these trials were removed during cleaning procedures. Second, it was asked whether the inhibition effects changed over blocks.

Figure 1 shows the average effect of the letter-context and word-context conditions across the ten blocks of the experiment. To observe the average effect in the letter-context condition, we subtracted the proportion of "H" responses for the A-letter context from that for the H-letter context, averaged over all participants. Similarly, the effect in the word-context condition, averaged over all participants, was formed by subtracting the proportion of H responses in the C_T-word context from that in the T_E-word context. In this graph, positive values indicate an assimilation effect; zero indicates no effect of context direction; negative values indicate a contrast effect. From the figure, an "acquisition period" can be observed. In the first and second block, participants became familiar with the experimental procedure and what was expected of them. As the participants seemed to learn about their task quickly, these blocks were included in the analysis presented in the main paper.

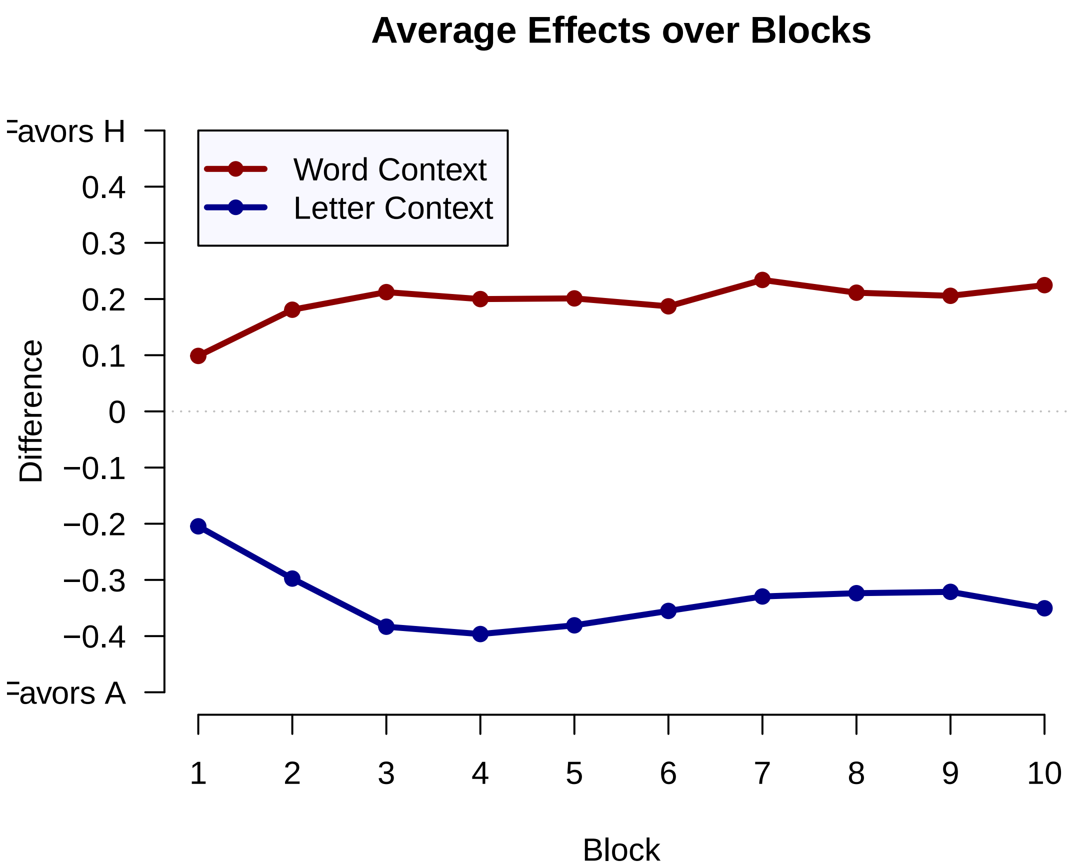

Figure 1. Average effects for assimlation and contrast for the middle morph across blocks. The negative driection denotes a robust contrast effect while the positive direction denotes a robust assimilation effect.

However, to satisfy our curiosity, it was decided to redo our analysis with the first two blocks excluded. The analysis method matched what was presented in the main paper and described here. Critical parameters remained individual estimates of contrast and assimilation, and the population-level correlation of these abilities. The parameter that holds the most interest is the population correlation. On each iteration of the MCMC chain, we again calculated $\rho = \Sigma_{1,2}/\sqrt{\Sigma_{1,1} \times \Sigma_{2,2}}$.

Figure 2A shows the assimilation and contrast effects for each participant. As before, there is a fair degree of variation across individuals as well as an overall positive relationship.

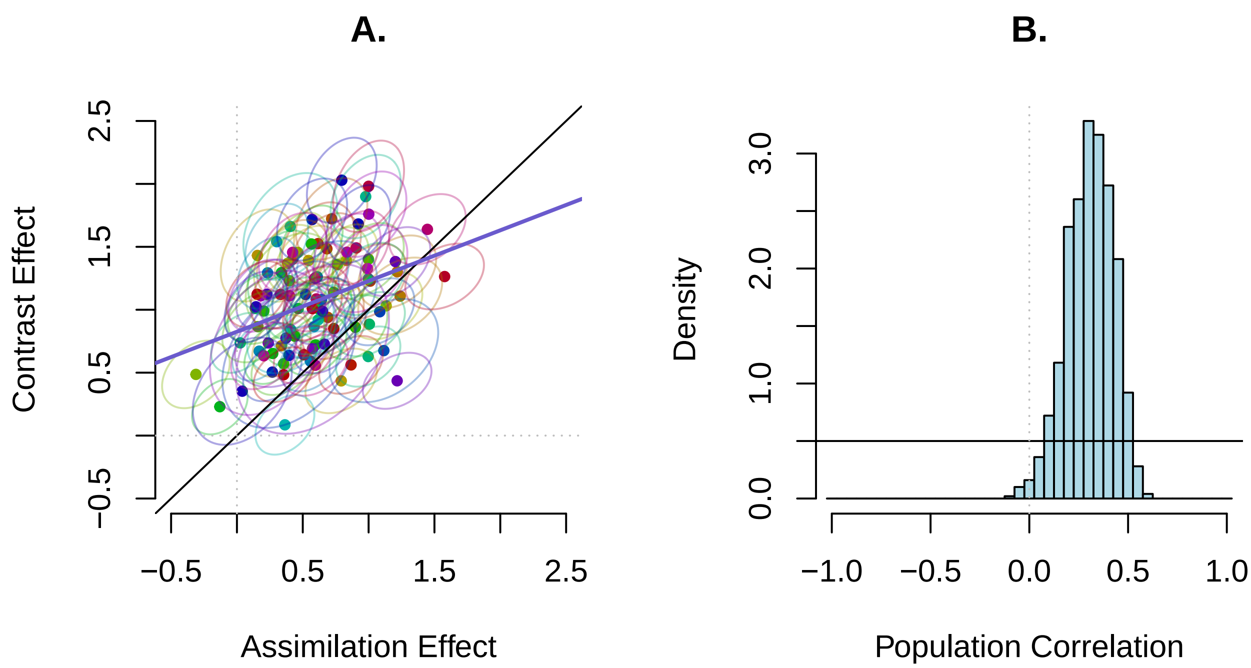

Figure 2. Model Results, excluding the first two blocks. **A.** Each participant's estimated assimilation effect against their estimated contrast effect. An ellipse denotes the standard deviations of the estimate and the blue regression line is the line of best fit. **B.** The posterior distribution of the population correlation, ρ, between assimilation and contrast. The solid line denotes the prior distribution.

To perform inference, we plot the prior and posterior distributions of the population-level correlation coefficient in Figure 2B. The prior distribution here was chosen to be flat, placing equal plausibility on all values of the correlation coefficient. The resulting posterior distribution remained localized for positive values away from zero. The mean of this posterior distribution is 0.31, which serves as a point estimate. The plausibility of a null correlation was assessed by a Savage-Dickey approach to Bayes factors (Dickey, 1971; Gelfand & Smith, 1990). The plausibility was reduced by a factor of 6.33. This indicates the data are 6.33 times more plausible under the alternative than the null.

References

Albert, J. H., & Chib, S. (1995). Bayesian residual analysis for binary response regression models. *Biometrika*, *82*, 747–759.

Dickey, J. M. (1971). The weighted likelihood ratio, linear hypotheses on normal location parameters. *The Annals of Mathematical Statistics*, *42*(1), 204–223. Retrieved from http://www.jstor.org/stable/2958475

Gelfand, A., & Smith, A. F. M. (1990). Sampling based approaches to calculating marginal densities. *Journal of the American Statistical Association*, *85*, 398–409.

Jackman, S. (2009). *Bayesian analysis for the social sciences*. Chichester, United Kingdom: John Wiley & Sons.

Rouder, J. N. (2016). The what, why, and how of born-open data. *Behavioral Research Methods*, *48*, 1062–1069. Retrieved from 10.3758/s13428-015-0630-z

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

Rouder, J. N., Lu, J., Sun, D., Speckman, P. L., Morey, R. D., & Naveh-Benjamin, M. (2007). Signal detection models with random participant and item effects. *Psychometrika*, *72*, 621–642.