Supplemental Material

Information on the Diffusion Model for Conflict Tasks and Fitting Procedure

Of the models developed to fit tasks with response competition, such as the Stroop task or flanker task, perhaps none is more theoretically driven than the Diffusion Model for Conflict Tasks (DMC) (Ulrich et al., 2015), which is a model that is impressive in both its breadth and its fit to empirical data. The DMC is a task-general model designed to probe data more deeply from any task with response conflict (i.e., tasks with both incongruent and congruent trials). In line with many models of executive attention and control (Heatherton and Wagner, 2011; Nigg, 2000; Ridderinkhof et al., 2004; Shields et al., 2017), the DMC posits that response selection is driven by superimposed automatic (i.e., bottom-up) and controlled (i.e., top-down) attentional processes. Salient, goal-irrelevant stimuli (e.g., flanking arrows) are presumed to automatically capture or draw attention, which may or may not facilitate task performance. Goal-relevant stimuli (e.g., the centered, focal arrow) are presumed to be attended to by a controlled (i.e., goal-directed) attentional process that necessarily facilitates task performance. In addition to fitting parameters for attentional processes, the model also fits parameters for the time that it takes to encode and execute a chosen motor action (i.e., nondecision time), variability in the time it takes to encode and execute a chosen motor action (i.e., nondecision time variability), the amount of information needed before making a decision (i.e., decision boundary), and variability in the starting point of information accumulation (i.e., starting-point variability). Owing to its strong theoretical basis, this model is the only extant model that can fit the observed shapes of delta plots of both the Simon and flanker tasks (Ulrich et al., 2015). This model shows good to excellent parameter recovery (White et al., 2018) and its model parameters correspond with EMG and EEG data (Servant et al., 2016). Thus, the DMC permits deeper analyses—as well as precise estimates—of the component processes that underlie task performance.

To fit the model to trial data, the model was fit to cumulative density functions (CDFs), capturing response time distributions (.1, .3, .5, .7, .9 quantiles), and conditional accuracy functions (CAFs), describing error proportions in five equally spaced response time bins, which constitutes the error data considered in the fitting procedure (Servant et al., 2016; White et al., 2018). Parameters were constrained

to be positive numbers (and in the case of *a* and *b*, greater than or equal to 1). Following work examining the parameter recovery of this model (White et al., 2018), the observed and predicted CDFs and CAFs were then fit by minimizing the following χ^2 statistic:

$$\chi^2 = \sum_{i=1}^{2} N_i \sum_{j=1}^{X} \frac{(p_{ij} - \pi_{ij})^2}{\pi_{ij}}$$

where p_{ij} represents the observed and π_{ij} represents the predicted proportion of trials in bin *j* of trial type (i.e., congruent, incongruent) *i*, and N_i represents the number of trials per trial type *i*. We fit the model using the subplex optimization method (Rowan, 1990), which divides the parameter space into subspaces and then optimizes these subspaces by the Nelder-Mead simplex method. The subplex optimization method was chosen because subplex was developed and is much better suited for optimization problems involving numerous parameters, dismissal of unacceptable parameters (e.g., parameters with negative values), and noisy data (i.e., models with random sampling, such as a Wiener diffusion process) making this optimization method more appropriate for this model than the simplex method (Farrell and Lewandowsky, 2018).

Parameter estimates were obtained for each participant according to the following procedure. First, to determine plausible starting values, we selected the best fitting parameters from previously published work fitting this model to flanker data (Ulrich et al., 2015). We then compared the observed and predicted means and standard deviations from 10,000 simulated trials to gross behavioral outcomes on the flanker task (e.g., mean congruent and incongruent response times and error rates) averaged across all participants and manually modified the parameters to better match these gross behavioral metrics. Next, we used those parameters to randomly generate 14 sets of additional parameters (15 total sets) that were constrained to be greater than zero and normally distributed around each respective initial parameter value with a variance of 50% of the respective initial parameter value. These 15 parameter sets were then fit to the data from the entire sample by simulating 10,000 trials of each type (i.e., congruent and incongruent; 20,000 total trials) and comparing the observed and predicted CDFs and CAFs using the procedure described in the preceding paragraph. The final parameter values from the best-fitting model

from the 15 parameter sets were then selected as the best-fitting parameters for the entire sample. Once the best-fitting parameters for the entire sample were found, we used those parameters as the starting values to randomly generate nine sets of additional parameters (ten total sets) for each participant that were constrained to be greater than zero and normally distributed around each respective initial parameter value with a variance of 50% of the respective initial parameter value. These nine randomly generated sets of parameter values differed for each participant, but the best-fitting parameters for the whole sample were fit to each participant's data. These parameter sets for each participant were then fit sequentially to the participant's data by simulating 10,000 trials of each type (i.e., congruent and incongruent; 20,000 total trials) and comparing the observed and predicted CDFs and CAFs using the procedure described in the preceding paragraph. The final parameter values from the best-fitting model from the 10 parameter sets were selected as the best-fitting parameters for that person.

Cortisol Values by Condition and Sex

For the interested reader, we list the means and standard deviations of cortisol values by condition and sex in Supplemental Table 1 below.

Supplemental Table 1

Cortisol (nmol/L) by Condition and Sex

	Pre-Manipulation	Post-Manipulation
Condition	Mean (SD)	Mean (SD)
Stress		
Females (*n* = 37)	6.83 (4.74)	7.74 (4.33)
Males (*n* = 12)	8.53 (5.58)	12.41 (6.22)
Control		
Females (*n* = 40)	7.97 (4.36)	7.36 (3.60)
Males (*n* = 15)	8.79 (7.63)	8.64 (5.57)

Note: *n*s reflect participants with nonmissing cortisol at both timepoints.

Computational Modeling on the First and Second Half of the Classic Flanker

The DMC is a data-hungry computational model, with 500 trials only achieving an average correlation between true and recovered parameters of $r = .86$ (White et al., 2018). Because of this, to

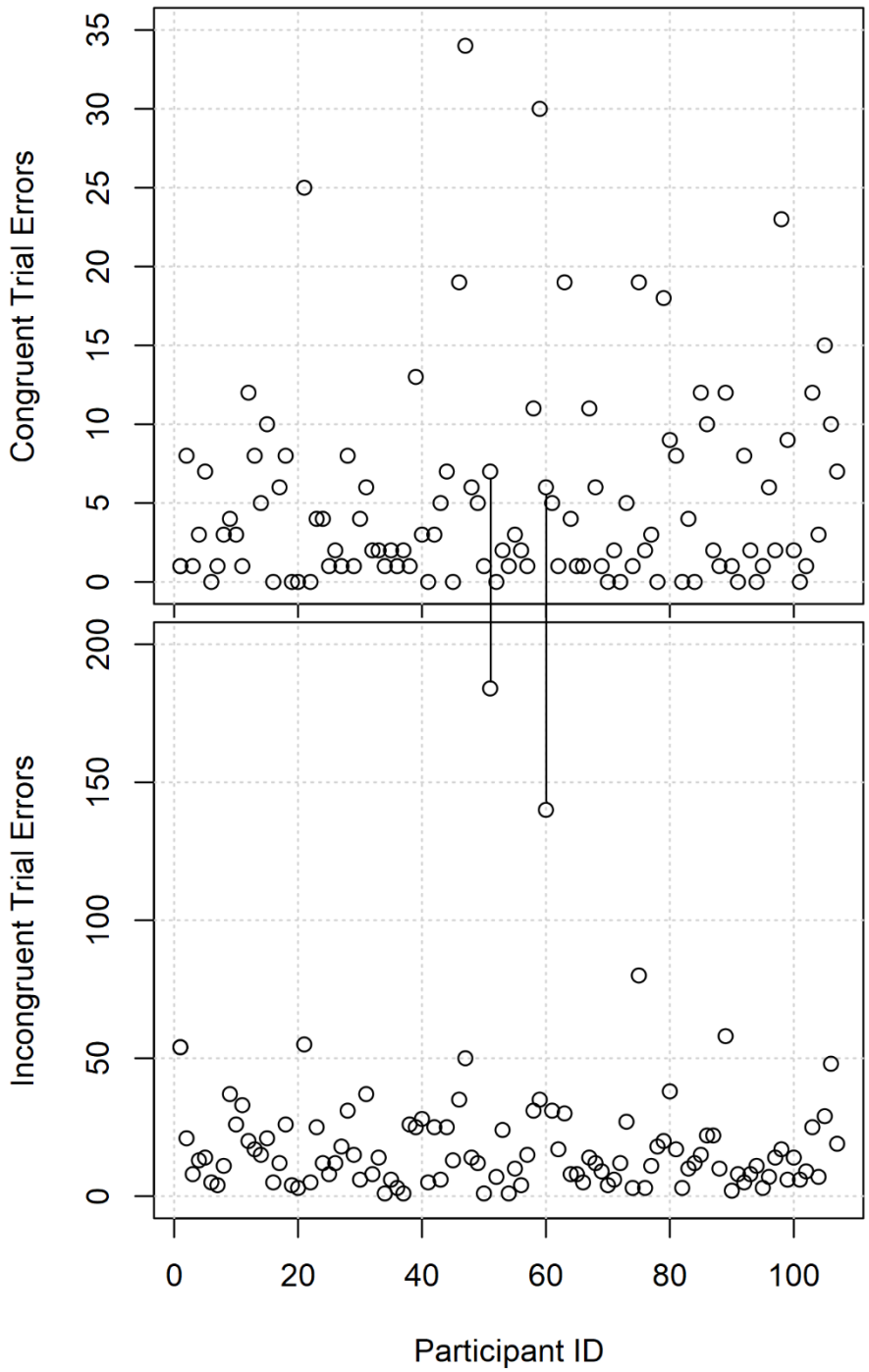
preserve the fidelity of the estimates, we fit the DMC to the data from the entire task in the results presented in the main text. We acknowledge that doing so obscures time-dependent effects emerging over the course of the task, and this is a limitation of the model. However, and importantly, the only effect of the stress manipulation on any model parameter was on nondecision time, which is a parameter that shows excellent recovery even with only 200 trials (r = .98 between true and recovered parameter values; White et al., 2018). Because of this, we present the results of model fits for nondecision time from the first half of the task and the second half of the task separately here.

Results from fitting the DMC to the first and second halves of the task show what may have been expected from the raw behavioral data. In the first half of the task, participants in the stress condition (M = 380.53ms, SE = 3.17) evidenced significantly less nondecision time than participants in the control condition (M = 399.03, SE = 5.57), $t(105) = -2.78$, $p = .006$, $d = -0.54$. In the second half of the task, participants in the stress condition (M = 383.02ms, SE = 4.17) did not significantly differ in nondecision time from participants in the control condition (M = 392.49, SE = 6.13), $t(105) = -1.24$, $p = .217$, $d = -0.24$. This change was because the control condition decreased in nondecision time from the first half of the task (M = 399.03ms) to the second half of the task (M = 392.49ms), $p = .019$, whereas the stress condition remained constant in nondecision time from the first half of the task (M = 380.53ms) to the second half (M = 383.02ms), p =.481. This is consistent with our interpretation that mild stress resulted in a bypass of the typical shift from monitored execution of chosen motor actions to an unmonitored, automatic execution of chosen actions (see manuscript discussion).

Assessing Potential Effects on Mind Wandering

Mind wandering can reduce response times, but this reduction in response time as a result of mind wandering is not beneficial for task performance. Because of this, we assessed whether stress influenced a common index of mind wandering, the response time coefficient of variation (i.e., mean RT divided by SD RT for each participant; Hu et al., 2012; Mrazek et al., 2011). A larger response time coefficient of variation indicates greater response time variability across the task, and greater mind wandering (Hu et al., 2012; Mrazek et al., 2011). We found that participants in the stress condition did not

differ in response time coefficients of variation (M = 5.98, SE = 0.17) from participants in the control condition (M = 6.10, SE = 0.17), $t(105) = 0.51$, $p = .613$—with participants in the stress condition showing smaller response time coefficients of variation, if anything. These results were similar analyzing response time standard deviations; participants in the stress condition (M = 80.06, SE = 2.47) did not differ in response time standard deviations from participants in the control condition (M = 80.95, SE = 2.74), $t(105) = 0.24$, $p = .812$. Thus, participants in the stress condition did not show greater response time variability than participants in the control condition, indicating that they did not evidence greater mind wandering.



Supplementary Figure 1. Number of errors committed in the classic flanker task by participant ID. Notably, both control participants who committed an error in more than 50% of incongruent trials were not outliers in errors on congruent trials, as shown by the line connecting these participants across the trial types.

References

Farrell, S., Lewandowsky, S., 2018. Computational modeling of cognition and behavior. Cambridge University Press, New York, NY. https://doi.org/10.1017/9781316272503

Heatherton, T.F., Wagner, D.D., 2011. Cognitive neuroscience of self-regulation failure. Trends Cogn. Sci. 15, 132–139. https://doi.org/10.1016/j.tics.2010.12.005

Hu, N., He, S., & Xu, B. (2012). Different efficiencies of attentional orienting in different wandering minds. *Consciousness and Cognition*, *21*, 139-148. https://doi.org/10.1016/j.concog.2011.12.007

Mrazek, M. D., Chin, J. M., Schmader, T., Hartson, K. A., Smallwood, J., & Schooler, J. W. (2011). Threatened to distraction: Mind-wandering as a consequence of stereotype threat. *Journal of Experimental Social Psychology*, *47*, 1243-1248. https://doi.org/10.1016/j.jesp.2011.05.011

Nigg, J.T., 2000. On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. Psychol. Bull. 126, 220–246. https://doi.org/10.1037/0033-2909.126.2.220

Ridderinkhof, K.R., van den Wildenberg, W.P.M., Wijnen, J., Burle, B., 2004. Response inhibition in conflict tasks is revealed in delta plots, in: Posner, M.I. (Ed.), Cognitive Neuroscience of Attention. Guilford Press, New York, NY, pp. 369–377.

Rowan, T.H., 1990. Functional stability analysis of numerical algorithms. Ph.D. thesis. University of Texas at Austin.

Servant, M., White, C., Montagnini, A., Burle, B., 2016. Linking theoretical decision-making mechanisms in the Simon task with electrophysiological data: A model-based neuroscience study in humans. J. Cogn. Neurosci. 28, 1501–1521. https://doi.org/10.1162/jocn_a_00989

Shields, G.S., Moons, W.G., Slavich, G.M., 2017. Inflammation, self-regulation, and health: An immunologic model of self-regulatory failure. Perspect. Psychol. Sci. 12, 588–612. https://doi.org/10.1177/1745691616689091

Ulrich, R., Schröter, H., Leuthold, H., Birngruber, T., 2015. Automatic and controlled stimulus processing in conflict tasks: Superimposed diffusion processes and delta functions. Cogn. Psychol.

78, 148–174. https://doi.org/10.1016/J.COGPSYCH.2015.02.005

White, C.N., Servant, M., Logan, G.D., 2018. Testing the validity of conflict drift-diffusion models for use in estimating cognitive processes: A parameter-recovery study. Psychon. Bull. Rev. 25, 286–301. https://doi.org/10.3758/s13423-017-1271-2