**Hierarchical Drift-Diffusion Modeling of Flanker Task Behavior as Related to Remitted Depression and Recent Suicidal Ideation**

Jiani Li, B.S.

Department of Psychology

University of Southern California

PSYC 573: Bayesian Data Analysis

Instructor: Mark Lai, Ph.D.

**Introduction**

Cognitive control refers to the ability to carry out goal-directed actions in face of distraction and flexibly switch between actions depending on contexts (Badre, 2024). A canonical task to measure this ability is the Eriksen flanker task, which involves quickly determining the direction of a central stimulus flanked by stimuli that either point in the same direction as the central stimulus (the Congruent condition) or in the opposite direction (the Incongruent condition) (Eriksen & Eriksen, 1974). Compared to a congruent trial, an incongruent trial typically results in a slowdown in reaction time and a decline in accuracy, which is termed the interference effect. The Eriksen flanker task has been widely applied to measure cognitive control in relation to psychopathology, such as major depressive disorder (MDD), obsessive-compulsive disorder, attention-deficit/hyperactivity disorder, and externalizing behavior (Dillon et al., 2015; Ekman & Friesen, 1971; Hall et al., 2021; Haller et al., 2021; Pe et al., 2013; Riesel et al., 2019). Such studies typically reported a discrepancy between individuals with and without a psychological disorder in accuracy and reaction time. While one would expect lower accuracies and slower reaction times to be associated with greater symptom severity, the literature has sometimes suggested otherwise. For instance, two studies showed that participants with severe anhedonia were slower but more accurate at identifying the direction of the central stimulus than those without (Dubal et al., 2000; Dubal & Jouvent, 2004). These counterintuitive findings gave rise to the hypothesis that MDD is associated with a deliberative cognitive style that prioritizes accuracy over speed, which may be motivated by greater aversion to loss and the resulting negative mood or by a need to focus processing resources on identifying the causes of low mood (Andrews & Thomson, 2009; Robinson et al., 2007). However, this hypothesis cannot be directly tested just by analyzing accuracies and reaction times themselves, as other mechanisms, such as sluggish motor response, can also give rise to a slower and more accurate response style. Drift diffusion model (DDM), which decomposes choices into several processes, is well suited to testing multiple alternative explanations of an observed response style (Ratcliff & McKoon, 2008).

In brief, a DDM postulates that in a two-choice task, the two choices are defined by an upper and a lower decision boundary respectively, with α being the *distance* between the two boundaries. A choice is made through a drift process that starts from a starting point with *bias* z, proceeds according to accumulated evidence, and finishes when one of the decision boundaries is surpassed that triggers the corresponding response. This drift process occurs at a *drift rate* of v and the time taken to execute non-decision behaviors (e.g., perception, motor preparation, execution) is termed *non-decision time* t.

Applying a DDM to flanker task data collected from individuals with and without MDD, Dillon and colleagues (2015) found that individuals with MDD showed a slower drift rate specific to executive control that tracked the severity of anhedonia. This confirms the utility of disentangling interrelated processes of decision-making through DDM in understanding possible mechanisms of psychopathology. The present study seeks to extend prior literature on cognitive control among individuals with current MDD to those with remitted MDD, who are not currently influenced by MDD yet may still preserve vulnerability factors that explain their past MDD. Thus, any differences we observe in decision-making behavior between healthy controls and those with remitted MDD, while potentially reflecting residual effects of past depressive episodes, may also represent trait-like vulnerability factors of MDD and therefore advance us towards a better dissociation between trait- versus state-like mechanisms of MDD.

**Methods**

**Participants**

139 young adults participated in the study. Using the exclusion criteria adopted by Dillon and colleagues (2015), we excluded eight participants who had more than five outlier trials (i.e., >10% of total trials), which were defined as having a reaction time that is greater or smaller than three standard deviations from the person mean. Of the remaining 131 participants (age = 18-40 years, 45.8% non-white), there were 51 healthy controls (HCs), 51 individuals with rMDD, and 29 individuals with current MDD and active suicidal ideation in the past month. The Diagnostic Interview for Genetic Studies (Nurnberger et al., 1994) was used to evaluate participants' diagnostic status, and the Columbia Suicide Severity Rating Scale was used to assess participants’ suicidal ideation (Posner et al., 2011). Those with a history of depression met DSM-5 criteria for lifetime MDD but had been free from major depressive episodes or significant depressive symptoms for at least two months (American Psychiatric Association, 2013). HCs did not have any current or past psychiatric disorders or first-degree family members with known psychiatric disorders.

**The Flanker Task**

We used the flanker task from the NIH EXAMINER toolbox (Kramer et al., 2014) to index cognitive control. On each trial, participants are shown a row of five arrows, with the central arrow either pointing in the same direction as (i.e., the congruent condition) or in the opposite direction from (i.e., the incongruent condition) the four flanking arrows. Participants need to indicate the direction of the central arrow as quickly and accurately as possible by pressing the left or right arrow key. There are 24 trials in each condition and 48 trials in total. Participants’ responses and reaction times are recorded.

**Hierarchical Bayesian DDM**

We implemented hierarchical Bayesian DDM analyses using the HSSM Python package version 0.2.4 (Fengler et al., 2024). As stated above, a DDM assumes that when making a two-stimulus choice, an individual makes a speed-accuracy tradeoff by continuously sampling evidence from the external environment until the decision threshold for a stimulus is reached and the corresponding action is executed (Ratcliff & McKoon, 2008). We utilized the likelihood function in Navarro & Fuss (2009) supported by the HSSM package, which assumes that the observed choices and reaction times follow a *Wiener first-passage time* distribution (c, td) ~ WFPT(v, a, t, z) with a probability density function specified as follows:

We constructed hierarchical Bayesian DDMs to estimate average and subject-specific drift rates (v), boundary separations (a), non-decision times (t) and biases (z). Additionally, we examined whether individuals in the three clinical groups (HC, rMDD and SI) differ significantly in DDM parameters by performing a model comparison between a model that included trial type (congruent vs. incongruent) as the only covariate (M1), and one that included the main effects of group and trial type as well as their interaction (M3). Further, we included a model that had group (HC, rMDD and SI) as the only covariate (M2). In all models, we regressed v, a and t on the covariate(s). We used weakly informative priors for all model parameters. Priors for the most complex model (i.e., the model with the interaction term, M3) are shown below:

In the above equations, *i* denotes trials and *j* denotes subjects.

The posterior distributions of DDM and regression parameters were obtained using Hamiltonian Monte Carlo (HMC) sampling with the No U-Turn (NUTS) sampler. Sampling was conducted with 4 chains and 15000 iterations for each chain (the first 5000 of which were discarded as warm-ups). Convergence of the HMC chains were assessed by examining the rank histograms and of the HMC chains, the trace plots of the posterior samples and the statistics (Vehtari et al., 2021). As shown in Figures 2 and 5, the rank histograms of the HMC chains were relatively uniform for all parameters except those related to t (e.g., the intercept of t). Similarly, the trace plots (Figures 3 and 6) suggested good mixing of the chains for all DDM parameters except those related to t. The statistics were <1.01 for all model parameters except t-related parameters. The effective sample sizes were > 400 for all parameters to ensure accurate approximation of the posterior distributions. Overall, convergence diagnostics suggested that non-decision times may not have been estimated properly.

**Results**

Model comparison (see Tables 1 and 2) suggested that a model that included the main effects of group and trial type and their interaction (M3) did not outperform a model that included trial type as the only covariate (M1), as the differences between M1’s and M3’s LOOIC and WAIC values were not larger than their standard errors. This indicates that incorporating clinical group as a predictor did not significantly improve the prediction of DDM parameters. We presented results from both M1 and M3 here.

Table 3 summarizes the coefficient estimates of M1. We found evidence that non-decision times and boundary separations were greater in the incongruent trials than in the congruent trials (β2t = 0.075, 80% HDI = [0.072, 0.078]; β2a = 0.013, 80% HDI = [0.005, 0.020]). The posterior predictive checks (Figure 4) show that the model predictions were overall consistent with the observed data.

Table 4 summarizes the coefficient estimates of M3. Similar to what M1 indicated, we found evidence that non-decision times and boundary separations were greater in the incongruent trials than in the congruent trials (β2t = 0.057, 80% HDI = [0.049, 0.066]; β2a = 0.028, 80% HDI = [0.013, 0.043]). Further, there was a positive main effect of being in the recent SI group on boundary separations (β1a\_SI = 0.048, 80% HDI = [0.023, 0.075]). Analyses of the posterior draws of each of the six combinations of trial type and group (Tables 5 and 6) revealed that individuals with recent SI showed greater boundary separations than HCs in the congruent trials (HC a congruent - SI a congruent 80% HDI = [-0.075, -0.023]). Moreover, individuals with rMDD showed greater non-decision times than HCs in incongruent trials (HC t incongruent - rMDD t incongruent 80% HDI = [-0.075, -0.024]).

**Discussion**

The present study sought to examine differences in drift rate, boundary separation and non-decision time in the flanker task between healthy individuals, individuals with rMDD and individuals with recent SI. Overall, we did not find evidence for such group differences, as including group as a predictor did not significantly improve the prediction of DDM parameters beyond what a model with trial type as the only covariate could do. However, further examination of a fully specified model that included both predictors revealed significant differences between HCs and the other two clinical groups in boundary separation and non-decision time. Specifically, compared to HCs, young adults with rMDD displayed greater non-decision times, suggesting that people with rMDD spent longer time on non-decision-making behavior such as perception and motor execution. In addition, individuals with recent SI showed greater boundary separation than HCs, indicating that greater amount of evidence was needed for individuals with recent SI to reach a decision threshold.

There are two possible reasons for a lack of overall effects of clinical groups. First, both the level-one and the level-two sample sizes were small. All three clinical groups, especially the SI group, were relatively small in size (HC = 51, rMDD = 51, SI = 29). Moreover, the flanker task consisted of 48 trials only, with 24 trials in each condition. Although Bayesian methods are known to be robust with smaller samples, we did observe convergence issues for parameters relating to non-decision time (e.g., the intercept of t), which might suggest that a higher number of iterations was needed to stably estimate certain DDM parameters. Additionally, the smaller sample sizes might have contributed to wider posterior distributions of the DDM parameters and regression coefficients than they would have been if we had more participants per group and more trials, thereby further limiting our ability to detect group differences in DDM parameters. As reported by Rappaport and colleagues (2024), DDM parameters fit to a flanker task with 330 trials exhibited low reliability with NIH Toolbox flanker scores but substantially better reliability with other neuropsychological measures of cognitive control (e.g., electroencephalogram data), suggesting that the NIH Toolbox flanker task may not contain sufficient number of trials to reliably index cognitive control. Second, a lack of group differences in behavioral characteristics in the flanker task is also consistent with an increasing emphasis on a dimensional approach to examining psychopathology in the field of clinical psychology (Cuthbert, 2014). As opposed to treating mental disorders as all-or-none entities, a dimensional perspective of psychopathology views mental health as a continuum with no clear qualitative demarcations in between. Moreover, putative mechanisms of psychopathology, such as cognitive control, may also be viewed as continuous in nature and cut across diagnostic categories. Thus, individuals with and without a mental disorder (e.g., MDD) may differ in cognitive control by degree instead of by kind, leading to a difficulty detecting group differences in flanker task behavior.

Despite not outperforming a more parsimonious model, a model that included group and trial type as predictors of DDM parameters revealed a longer non-decision time (i.e., time spent on perception, motor preparation and response execution) among individuals with recent SI than HCs. This is consistent with the fact that psychomotor retardation is a canonical symptom of MDD, which could manifest as slower motor movement in the flanker task (Buyukdura et al., 2011). Further, we also found evidence for wider decision boundaries among individuals with rMDD than HCs, which implies that individuals with rMDD may have required more evidence to reach a decision threshold at the cost of speed than healthy individuals. This finding bears resemblance to the theory that characterizes MDD etiology as an overly cautious cognitive style that prioritizes accuracy over speed (Andrews & Thomson, 2009; Robinson et al., 2007). Indeed, despite using a different task, Lawlor and colleagues (2019) reported wider decision boundaries in a probabilistic reward task among individuals with current MDD than never-depressed individuals. The current study builds on the existing literature on MDD and decision-making behavior by suggesting that this deliberative cognitive style may be relatively independent from the active state of depression. Although the current study does not provide conclusive evidence for this possibility, and future studies with greater numbers of trials and participants are needed to replicate the result, the current study nevertheless may raise the question of whether such deliberative cognitive style may be a vulnerability factor that predisposes individuals to depression, or it may reflect a residual phenomenon from past depression. This question could be answered, for instance, with longitudinal studies that capture decision-making behavior before and after the onset of MDD, which would allow researchers to resolve the temporal order of a deliberative cognitive/response style and MDD.

**References**

Andrews, P. W., & Thomson, J. A. (2009). The bright side of being blue: Depression as an adaptation for analyzing complex problems. *Psychological Review*, *116*(3), 620–654. https://doi.org/10.1037/a0016242

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Badre, D. (2024). *Cognitive Control*. https://doi.org/10.1146/annurev-psych-022024-103901

Buyukdura, J. S., McClintock, S. M., & Croarkin, P. E. (2011). Psychomotor retardation in depression: Biological underpinnings, measurement, and treatment. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *35*(2), 395–409. https://doi.org/10.1016/j.pnpbp.2010.10.019

Cuthbert, B. N. (2014). The RDoC framework: Facilitating transition from ICD/DSM to dimensional approaches that integrate neuroscience and psychopathology. *World Psychiatry*, *13*(1), 28–35.

Dillon, D. G., Wiecki, T., Pechtel, P., Webb, C., Goer, F., Murray, L., Trivedi, M., Fava, M., McGrath, P. J., Weissman, M., Parsey, R., Kurian, B., Adams, P., Carmody, T., Weyandt, S., Shores-Wilson, K., Toups, M., McInnis, M., Oquendo, M. A., … Pizzagalli, D. A. (2015). A computational analysis of flanker interference in depression. *Psychological Medicine*, *45*(11), 2333–2344. https://doi.org/10.1017/S0033291715000276

Dubal, S., & Jouvent, R. (2004). Time-on-task effect in trait anhedonia. *European Psychiatry*, *19*(5), 285–291.

Dubal, S., Pierson, A., & Jouvent, R. (2000). Focused attention in anhedonia: A P3 study. *Psychophysiology*, *37*(5), 711–714.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17*(2), 124.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*(1), 143–149.

Fengler, A., Xu, P., Bera, K., Omar, A., Frank, M.J. (in preparation). HSSM: A generalized toolbox for hierarchical bayesian estimation of computational models in cognitive neuroscience.

Hall, N. T., Schreiber, A. M., Allen, T. A., & Hallquist, M. N. (2021). *Disentangling cognitive processes in externalizing psychopathology using drift diffusion modeling: Antagonism, but not disinhibition, is associated with poor cognitive control*. https://doi.org/10.1111/jopy.12628

Haller, S. P., Stoddard, J., Pagliaccio, D., Bui, H., MacGillivray, C., Jones, M., & Brotman, M. A. (2021). Computational Modeling of Attentional Impairments in Disruptive Mood Dysregulation and Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *60*(5), 637–645. https://doi.org/10.1016/j.jaac.2020.08.468

Kramer, J. H., Mungas, D., Possin, K. L., Rankin, K. P., Boxer, A. L., Rosen, H. J., Bostrom, A., Sinha, L., Berhel, A., & Widmeyer, M. (2014). NIH EXAMINER: conceptualization and development of an executive function battery. *Journal of the International Neuropsychological Society*, *20*(1), 11–19.

Lawlor, V. M., Webb, C. A., Wiecki, T. V., Frank, M. J., Trivedi, M., Pizzagalli, D. A., & Dillon, D. G. (2020). Dissecting the impact of depression on decision-making. *Psychological Medicine*, *50*(10), 1613–1622. https://doi.org/10.1017/S0033291719001570

Navarro, D. J., & Fuss, I. G. (2009). Fast and accurate calculations for first-passage times in Wiener diffusion models. *Journal of Mathematical Psychology*, *53*(4), 222–230. https://doi.org/10.1016/j.jmp.2009.02.003

Pe, M. L., Vandekerckhove, J., & Kuppens, P. (2013). A diffusion model account of the relationship between the emotional flanker task and rumination and depression. *Emotion*, *13*(4), 739–747. https://doi.org/10.1037/a0031628

Posner, K., Brown, G. K., Stanley, B., Brent, D. A., Yershova, K. V., Oquendo, M. A., Currier, G. W., Melvin, G. A., Greenhill, L., Shen, S., & Mann, J. J. (2011). The Columbia–Suicide Severity Rating Scale: Initial Validity and Internal Consistency Findings From Three Multisite Studies With Adolescents and Adults. *The American Journal of Psychiatry*, *168*(12), 1266–1277. https://doi.org/10.1176/appi.ajp.2011.10111704

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Riesel, A., Kathmann, N., & Klawohn, J. (2019). Flexibility of error-monitoring in obsessive–compulsive disorder under speed and accuracy instructions. *Journal of Abnormal Psychology*, *128*(7), 671–677. https://doi.org/10.1037/abn0000463

Robinson, M. D., Meier, B. P., Wilkowski, B. M., & Ode, S. (2007). Introversion, inhibition, and displayed anxiety: The role of error reactivity processes. *Journal of Research in Personality*, *41*(3), 558–578.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2021). Rank-normalization, folding, and localization: An improved R ̂ for assessing convergence of MCMC (with discussion). *Bayesian Analysis*, *16*(2), 667–718.

**Appendix**

**Table 1**

*Model Comparison Results (by LOO)*

A screenshot of a number

Description automatically generated

**Table 2**

*Model Comparison Results (by WAIC)*

A screenshot of a computer

Description automatically generated

**Table 3**

*Regression Coefficient Estimates of M1*

A table with numbers and text

Description automatically generated

**Table 4**

*Regression Coefficient Estimates of M3*

A table of numbers and letters

Description automatically generated

**Table 5**

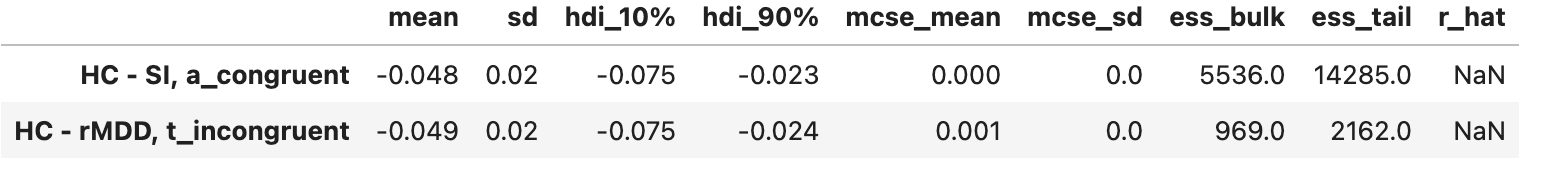
*Summary Statistics of Posterior Draws from the Six Combinations of Trial Type and Group in M3*

A table of numbers and symbols

Description automatically generated

**Table 6**

*Summary Statistics of the Significant Group Comparisons in M3*

**

**Figure 1**

*Schematics of a Drift Process*

A diagram of a response boundary

Description automatically generated

**Figure 2**

*Rank Histograms of the HMC Chains for all DDM Parameters in M1*

A group of colorful bars

Description automatically generated

**Figure 3**

*Trace Plots of the DDM Parameters in M1*

**A screenshot of a graph

Description automatically generated**

**Figure 4**

*Posterior Predictive Checks of M1*

A graph of a number of graphs

Description automatically generated with medium confidence

**Figure 5**

*Rank Histograms of**the HMC Chains for all DDM Parameters in M3*

**A group of colorful bars

Description automatically generated**

**Figure 6**

*Trace Plots of M3*

A group of graphs showing different types of data

Description automatically generated

**Figure 7**

*Posterior Predictive Checks of M3*

A graph of a number of graphs

Description automatically generated with medium confidence