**Simulation Results Report** *Testing the NES Framework on a Stroop Task*

## **1. Introduction & Goals**

The Normative Executive System (NES) posits that a norm-sensitive Comparator accumulates evidence for competing impulses, while an Assent Gate (threshold) modulated by a serotonin-like signal governs response speed vs. caution. We tested two core predictions in a simulated Stroop task:

1. A **Stroop RT cost**: incongruent trials should be slower than congruent ones.
2. A **threshold (serotonin) effect**: lowering the decision threshold speeds responses but induces errors (speed–accuracy trade-off).

## **2. Methods**

**Model Parameters**

| **Parameter** | **Value** | **Description** |
| --- | --- | --- |
| w\_s | 0.5 | Salience weight |
| w\_n | 1.0 (then 0.5 sweep) | Norm-congruence weight |
| w\_u | 0.2 | Urgency weight |
| noise\_std\_dev | 0.1 (then 0.2, 0.3) | Gaussian noise σ per timestep |
| base\_threshold | 1.0 (then 0.8, 0.5) | Assent Gate baseline threshold |
| k\_ser | 0.5 | Serotonin sensitivity → threshold modulation |
| normal\_serotonin\_level | 0.0 | “Baseline” 5HT |
| dt | 0.01 s | Simulation timestep |
| max\_time | 3.0 s | Cut-off if no decision |

**Trial Types & Conditions**

* **Congruent\_Normal**: congruent Stroop, serotonin=0.0
* **Incongruent\_Normal**: incongruent Stroop, serotonin=0.0
* **Incongruent\_Low5HT**: incongruent Stroop, serotonin=–1.0
* **Incongruent\_High5HT**: incongruent Stroop, serotonin=+1.0

We ran 500 trials per condition, recording RT, choice, and correctness.

## **3. Results**

### **3.1 Initial Simulation (w\_n = 1.0)**

| **Condition** | **Mean RT (s)** | **Median RT (s)** | **Accuracy** |
| --- | --- | --- | --- |
| Congruent\_Normal | 0.651 | 0.650 | 100 % |
| Incongruent\_Normal | 0.787 | 0.790 | 100 % |
| Incongruent\_Low5HT | 0.390 | 0.390 | 100 % |
| Incongruent\_High5HT | 1.185 | 1.180 | 100 % |

* **Stroop Effect:** 0.787 s vs. 0.651 s confirms conflict-driven slowdown.
* **Serotonin (Threshold) Effect:** Low5HT→faster (0.390 s), High5HT→slower (1.185 s).
* **Ceiling Accuracy:** 100 % in all cells—normative weight was too strong for noise to induce errors.

### **3.2 Low5HT Sweep with Weakened Norm (w\_n = 0.5)**

Focusing on **Incongruent\_Low5HT**, we varied noise ∈ {0.1, 0.2, 0.3} and threshold ∈ {1.0, 0.8, 0.5}:

| **noise\_std\_dev** | **base\_threshold** | **Mean RT (s)** | **Accuracy** |
| --- | --- | --- | --- |
| 0.1 | 1.0 | 0.9720 | 100 % |
| 0.1 | 0.8 | 0.5764 | 98.2 % |
| 0.1 | 0.5 | 0.1840 | 90.0 % |
| 0.2 | 1.0 | 0.9331 | 91.8 % |
| 0.2 | 0.8 | 0.5543 | 83.6 % |
| 0.2 | 0.5 | 0.1882 | 62.8 % |
| 0.3 | 1.0 | 0.9462 | 84.4 % |
| 0.3 | 0.8 | 0.5645 | 69.4 % |
| 0.3 | 0.5 | 0.1925 | 41.2 % |

* **Speed–Accuracy Trade-off Achieved:** Lower thresholds & higher noise produce faster RTs but progressive accuracy decline (down to 41 % at noise=0.3, threshold=0.5).
* **Role of w\_n:** Halving the norm weight brought the model into a regime where normative control no longer dominates; random fluctuations can drive incorrect “word” responses, especially under low thresholds.

## **4. Analysis**

1. **Comparator Conflict Dynamics:**
   * Incongruent trials incur a drift-rate conflict (salience for “word” vs. norm for “color”), reproducing the classic Stroop cost.
2. **Assent Gate & Serotonin:**
   * Lowering threshold (simulating low 5HT) speeds RT.
   * Under strong norm (w\_n=1.0), no errors occurred; under weaker norm (w\_n=0.5), errors emerged in Low5HT, illustrating the NES prediction that an overly permissive gate yields impulsive mistakes.
3. **Norm Weight as a Critical Lever:**
   * The balance between salience and norm-congruence is adjustable via w\_n.
   * Our sweep confirms that w\_n mediates the model’s placement on the speed–accuracy continuum: higher w\_n → high accuracy, lower w\_n → greater susceptibility to noise.

## **5. Discussion**

* **Proof-of-Concept:** These simulations validate key NES mechanisms in a canonical cognitive control task, showing both Stroop interference and threshold-induced impulsivity.
* **Model Viability:** The NES architecture (drift-diffusion Comparator + threshold-based Assent Gate) faithfully reproduces human-like effects.
* **Parameter Sensitivity:** w\_n is a pivotal parameter for normative control; noise and threshold jointly govern the speed–accuracy trade-off.
* **Next Steps:**
  1. **Fit to Human Data:** Adjust parameters (e.g., w\_n, w\_s, noise) to quantitatively match empirical Stroop distributions.
  2. **Extend to Other Tasks:** Simulate delay discounting and moral dilemmas using analogous parameter manipulations.
  3. **Report Finalization:** Incorporate these results into a full NES validation paper, then move to empirical planning.

## **6. Appendix**

* **Full Sweep Data:** Tables above.
* **Code Snippets:** See test code.py for full implementation.
* **Additional Plots:**
  + Mean RT & accuracy bar‐plots
  + RT distributions per condition

**Conclusion:** This set of simulations provides robust computational evidence that the NES framework’s core processes—integration of norms in a drift-diffusion accumulator plus a modifiable threshold—can generate hallmark cognitive control phenomena, including impulsivity under low-caution states.