# Enhancing Drive-Based NPC Behavior with Panksepp-Inspired Affective Systems

**Luís Fernando Bicalho, Augusto Baffa**[1]

[1]Departamento de Informática
Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio)
Rio de Janeiro, RJ – Brasil

lfbicalho@puc-rio.br

abaffa@inf.puc-rio.br

***Abstract. Introduction:*** *Non-player characters (NPCs) that rely on fixed scripts often exhibit repetitive and non-adaptive behavior, reducing immersion in interactive environments. Recent approaches incorporate affective and motivational models, but often lack reproducibility and robustness.*
***Objective:*** *This paper investigates whether integrating Panksepp-inspired primary affective systems into a drive-based neurochemical NPC architecture improves behavioral robustness, stability, and diversity. Survival is treated as a measure of homeostatic robustness rather than as direct evidence of player-perceived believability.*
***Method:*** *We reformulate a previous Godot-based model into a deterministic Python simulation and introduce two architectural variants incorporating Panksepp's affective systems. The models are evaluated using multi-seed experiments (30 seeds) and parameter sensitivity analysis.*
***Results:*** *The deterministic baseline improves survival from 62% to 87.87%. The Panksepp-based models further increase survival to approximately 99%, while also reducing variance across seeds. However, behavioral imbalances such as reduced aggression emerge. The inclusion of an intermediate affective layer improves robustness and interpretability, but requires multi-objective tuning to preserve behavioral diversity.*
***Keywords*** *Game AI, NPC Behavior, Affective Computing, Panksepp, Emotion Modeling, Simulation*

## 1. Introduction

Non-player characters (NPCs) play a central role in how players perceive the credibility of a game world. When NPCs react only through fixed scripts, they tend to feel repetitive, predictable, and detached from the fiction of the game, whereas more believable agents can improve immersion, support narrative coherence, and make interaction feel socially meaningful [da Silva e de Souza Ribeiro 2021]. This problem has become even more relevant as contemporary games increasingly rely on social interaction, open-ended dialogue, and persistent worlds in which players expect characters to exhibit stable personalities, emotional consistency, and context-sensitive behavior [Aljammaz et al. 2023, Armanto et al. 2025].

One productive line of research has been to ground NPC behavior in models borrowed from psychology, affective computing, and neuroscience.

Earlier game-oriented approaches linked finite-state machines to emotional or personality variables in order to diversify reactions while preserving authorial control [Baffa et al. 2017, Garavaglia et al. 2022]. More recent works expanded this direction by introducing cultural variables, neurotransmitter-inspired emotion dynamics, narrative-planning formulations of personality and emotion, or broader believable-agent architectures [Bicalho et al. 2020, de Tarso Fernandes et al. 2024, Shirvani et al. 2023, Inyang et al. 2024]. In parallel, recent character systems driven by large language models and multimodal pipelines have widened the design space of believable interactive characters, but often at the cost of higher computational complexity and reduced interpretability [Wampfler et al. 2025, Galvão et al. 2025].

Our work is inserted in this trajectory. We start from the drive-based model proposed by [Bicalho et al. 2025], which integrates needs, drives, neurotransmitters, emotions, and finite-state-machine behavior selection, and we extend it in two directions. First, we produce a reproducible Python implementation that preserves the conceptual structure of the original work while enabling systematic experimentation across seeds and parameter settings. Second, we introduce two new variants inspired by Panksepp's primary affective systems, allowing us to test whether a more explicit motivational-affective layer can improve robustness and behavioral balance [Panksepp e Biven 2012, Panksepp 1998]. Our central research question is therefore whether a Panksepp-inspired extension can improve a drive-based affective NPC architecture without sacrificing interpretability and controllability.

The contributions of this paper are threefold: (1) A deterministic and reproducible reformulation of a drive-based affective NPC model[1]; (2) The introduction of two architectural variants incorporating Panksepp-inspired primary affective systems; (3) A systematic evaluation of robustness through multi-seed experiments and parameter sensitivity analysis. This work is an incremental extension of a previously proposed drive-based architecture rather than an entirely new one; nevertheless, to the best of our knowledge, it is the first to employ Panksepp's primary affective systems as an intermediate layer within a drive- and neurochemistry-based NPC model.

This paper is organized as follows. Section 2 reviews prior work on believable NPCs, emotion and personality simulation, and recent interactive character systems. Section 3 presents the theoretical foundations of the model, covering drives, Plutchik's emotion theory, Lövheim's cube, and Panksepp's primary affective systems. Section 4 details the proposed simulation variants and the evaluation protocol. Section 5 presents the experimental results, including the comparisons across models, seeds, and parameter perturbations. Finally, Section 6 concludes the paper and discusses future work.

## 2. Related Works

Research on believable NPCs has progressed from rule-based emotional reactions toward richer architectures combining personality, social variables, physiological abstractions, and, more recently, generative interaction. A systematic review by [da Silva e de Souza Ribeiro 2021] highlights a recurring trade-off: scripted solutions

[1] Artifact availability: the simulator, the three executable model variants, configuration files, random seeds, evaluation scripts, dependencies, and reproduction instructions are publicly available at `https://github.com/Kirink212/panksepp_paper_2026_artifacts.git`.

are easy to control but shallow, whereas adaptive ones add variability at the cost of complexity. Complementing this, [Aljammaz et al. 2023] argue that believability is multidimensional, emerging from the interplay of behavior, presentation, narrative context, and player expectations. These reviews frame believability as both a design and an AI problem, but do not provide a compact computational architecture for emotion-regulated NPC behavior.

Within our own research line, [Baffa et al. 2017] combined Plutchik's emotion theory, OCEAN personality traits, and finite-state machines to modulate NPC reactions, showing that even a lightweight architecture can produce differentiated reactions and more dynamic social encounters; however, it centered on external events and personality-mediated appraisal, without a physiological or homeostatic layer. [Bicalho et al. 2020] later added cultural dimensions, prejudice, trust, and proxemics, broadening the social scope but still without explicit endogenous needs or neurochemical variables.

A step toward physiological abstraction appears in [de Tarso Fernandes et al. 2024], which linked NPC emotion to neurotransmitter-inspired variables via an extended Lövheim's cube and Plutchik-like categories, increasing internal coherence but lacking a drive layer grounded in persistent needs—so it could express affective changes without a mechanism for homeostatic pressures such as hunger, fatigue, or safety to organize action selection. Extending that line, [Pinheiro et al. 2025] show how neurotransmitter-based models can be enriched, while reinforcing that purely neurochemical mappings remain challenging when interpretable, tunable action policies are required.

The most immediate baseline is therefore [Bicalho et al. 2025], which integrated needs and drives theory with neurotransmitters, Lövheim's cube, Plutchik-inspired labeling, and an emotion-augmented finite-state machine, explicitly modeling the chain from internal deficits to behavior. Its prototype nonetheless exposed clear weaknesses—overrepresented Surprise, limited survival, and an engine-dependent implementation that hindered reproducibility. Our work preserves its core insight while addressing these limitations through deterministic equations, multi-seed analysis, parameter sweeps, and Panksepp-inspired variants.

Beyond our own models, several recent works illustrate adjacent directions. [Garavaglia et al. 2022] proposed *Moody5*, a personality-biased architecture for interactive storytelling, strong on coherent character attitudes but focused on narrative consistency rather than homeostatic regulation. [Shirvani et al. 2023] combined personality and emotion in strong-story narrative planning, showing readers find the resulting behavior more believable—a strong result for story structure rather than situated NPC survival and reactive behavior.

Other contributions highlight the breadth of the field. [Inyang et al. 2024] proposed a mental-state-modelling architecture that improved perceived believability over an FSM baseline through explicit player-perception evaluation, though without a low-level affective control loop for lightweight populations. [Kaneko e Okada 2024] survey emotion-based 3D character behavior from a graphics and animation viewpoint, and [Galvão et al. 2025] broaden the motivation toward empathetic games, while [Armanto et al. 2025] review evolutionary approaches to NPC behavior, which discover effective parameters but are typically less interpretable than theory-driven affective architectures.

A final trend is large-scale interactive character systems driven by language and multimodal foundation models. [Wampfler et al. 2025] present a platform integrating personality, memory, conversational control, speech, and animation, showing how far the field has moved toward lifelike conversational characters—but such systems are infrastructure-heavy and less suited to controlled experiments on the causal role of internal affective variables. Our contribution is more modest in scope but stronger in interpretability: how a compact, controllable simulation can benefit from Panksepp-inspired affective systems.

## 3. Theoretical Background

Our model integrates four complementary perspectives: drive theory, motivational hierarchy, discrete emotion theory, and affective neuroscience. Together, these frameworks support the combination of persistent needs, neurochemical abstractions, categorical emotions, and action-selection mechanisms [Hull 1943, Maslow 1943, Plutchik 1980, Lövheim 2012, Panksepp 1998, Panksepp e Biven 2012].

**Drives.** Drive theory [Hull 1943] models behavior as driven by internal deficits that generate motivational pressure. In computational terms, this enables mapping variables such as hunger, energy, and safety into action tendencies. Maslow's hierarchy [Maslow 1943, Maslow 1954] complements this view by organizing needs into physiological, safety, and social categories. In this work, these needs are represented as continuous variables, supporting dynamic regulation of behavior over time [Bicalho et al. 2025].

**Wheel of Emotions.** Plutchik's model [Plutchik 1980, Plutchik 2001] represents emotions as structured discrete categories with varying intensity and oppositional relations. In our architecture, it provides a categorical interpretation layer, translating internal states into readable emotional labels that support interpretability and design control [Bicalho et al. 2025].

**Cube of Emotion.** Lövheim's model [Lövheim 2012] maps emotions to combinations of dopamine, serotonin, and noradrenaline, offering a compact continuous representation of affective state. In this work, it serves as a neurochemical abstraction layer, enabling numerical updates while preserving interpretability [de Tarso Fernandes et al. 2024, Bicalho et al. 2025].

**Primary Affective Systems.** Panksepp's theory [Panksepp 1998, Panksepp e Biven 2012] introduces evolutionarily grounded affective systems such as *SEEKING*, *FEAR*, and *CARE*, which are directly associated with action readiness. These systems complement the previous layers by providing an explicit motivational-affective mechanism that biases behavior selection.

Together, these models motivate a layered architecture in which drives regulate internal pressure, neurochemical variables encode continuous affective state, emotion categories provide interpretability, and Panksepp-inspired systems introduce behavior-oriented modulation. This combination is not intended as a biologically faithful model, but as a computationally interpretable framework for generating believable NPC behavior

[Plutchik 1980, Lövheim 2012, Panksepp e Biven 2012, Bicalho et al. 2025].

## 4. Methodology

This work extends the model proposed in [Bicalho et al. 2025] by replacing the original engine-dependent prototype with a reproducible Python simulation and by introducing two new variants based on Panksepp's primary affective systems. The goal of the new methodology was not only to improve survival rates, but also to preserve behavioral balance, state diversity, and robustness under different random seeds and parameter settings. Throughout this work, survival is treated as a measure of homeostatic robustness—how effectively an agent regulates its internal needs over time—and not as direct evidence of human-like or player-perceived believability, which would require the human-centered evaluation discussed in Section 6.

The original model combined needs and drives theory, Lövheim's Cube of Emotion, Plutchik-inspired emotion labels, and an emotion-augmented finite-state machine (FSM). In that architecture, internal needs such as hunger, fatigue, socialization, and safety generated drives; these drives modulated dopamine, serotonin, and noradrenaline; the neurochemical state was then mapped to an emotional category; and the resulting emotion triggered the next FSM transition. The Godot prototype showed that the approach was viable, but it also revealed important limitations. First, Surprise accounted for 30.59% of all decisions, suggesting that moderate noradrenaline alone was an insufficient criterion for surprise. Second, the prototype reported 62% survival, which indicated that the regulation loop was still fragile. Third, the original implementation was not designed for large-scale reproducible experimentation, making systematic seed and parameter analyses difficult. To address these limitations, we reimplemented the simulation in Python and organized the comparison around four reference conditions: (1) **Original prototype**: the original Godot-based model, used here as the baseline reference; (2) **Current Python version**: a deterministic reformulation of the original architecture, preserving the needs $\rightarrow$ neurotransmitters $\rightarrow$ emotion $\rightarrow$ FSM pipeline while improving reproducibility and removing direct random perturbations from the internal equations; (3) **Panksepp version 1**: a biologically oriented extension in which drives and context first modulate neurotransmitters, and the resulting neurochemical state activates Panksepp-inspired affective systems before emotion labeling and FSM selection; (4) **Panksepp version 2**: a game-AI-oriented extension in which drives and context first activate Panksepp-inspired affective systems, which then shape the neurochemical state used by the emotion and FSM layers.

The current Python version kept the conceptual assumptions of the original paper but introduced four practical changes: (i) the internal equations became deterministic, so need growth, neurochemical updates, and non-combat action effects no longer contained direct random noise; (ii) random initialization was preserved (randomized internal states, positions, and food placement), as is appropriate for simulation studies; (iii) stochastic events were retained only where conceptually justified—random novelty events and combat outcomes; and (iv) all coefficients were moved to a configuration dictionary, enabling manual tuning and automatic experimentation.

The main conceptual extension of this paper lies in the inclusion of Panksepp-inspired affective systems. In both new variants, we modeled seven continuous systems:

*SEEKING*, *FEAR*, *RAGE*, *CARE*, *PANIC/GRIEF*, *PLAY*, and *LUST*. Each system assumes a value in $[0, 1]$ and biases the action-selection process. The difference between the two versions is architectural. In Panksepp version 1, neurochemistry remains upstream, and affective systems emerge from neurotransmitter values combined with drives and context. In Panksepp version 2, affective systems are computed first and then used to derive neurochemistry, making the model more directly action-oriented and more convenient for game-AI tuning.

**Conflict and Attack activation rules.** Because the activation limits of the aggression response are central to interpreting the results, we state them explicitly. An NPC can attack in two situations. In a *resource conflict*, two agents dispute the same food source: both must have hunger above $62\%$, be within an attack range of $2.0$ units of the contested resource, and the attacker then engages with a base probability of $30\%$. This base probability is modulated by the current affective state—Anger and *RAGE* increase it, whereas Trust, Joy, and *CARE* decrease it. In a *direct affective response*, an attack can instead be triggered by a sufficiently strong aggression signal: the deterministic baseline requires an Anger intensity above $65\%$, while the Panksepp variants use a *RAGE* activation threshold of $55\%$. In both situations, the target must lie within attack range. As discussed in Section 5, the combination of these thresholds with the abundance of food and the distance requirement makes attack a comparatively rare event in the current environment.

To formalize the proposed architecture and ensure reproducibility, we define the internal dynamics using a set of deterministic equations describing the transformation from needs to actions. Let $N_t \in \mathbb{R}^k$ be the vector of needs at time $t$, and $D_t \in \mathbb{R}^k$ the corresponding drive vector $D_t^i = \alpha_i(N_t^i - \theta_i)$ where $\alpha_i$ is the sensitivity coefficient and $\theta_i$ is the baseline threshold.

The neurochemical state $T_t \in \mathbb{R}^3$ (dopamine, serotonin, noradrenaline) is computed as $T_t = W_D D_t + W_C C_t$ where $C_t$ represents contextual inputs and $W_D$, $W_C$ are weight matrices.

For the Panksepp variants, the affective systems vector $P_t \in [0, 1]^7$ is defined as $P_t = \sigma(W_T T_t + W_D D_t + W_C C_t)$ where $\sigma(\cdot)$ is a bounded activation function. Finally, action selection is modeled as $q_{t+1} = \arg\max_a \pi(a \mid P_t, E_t)$. The deterministic baseline follows $\mathbf{N}_t \to \mathbf{D}_t \to \mathbf{T}_t \to (E_t, I_t) \to q_{t+1}$ where $\mathbf{N}_t$ is the vector of needs, $\mathbf{D}_t$ is the drive vector, $\mathbf{T}_t = (D, S, N)$ is the neurochemical state, $E_t$ is the dominant emotion, $I_t$ is emotional intensity, and $q_{t+1}$ is the next FSM state. Panksepp version 1 inserts an affective layer after neurochemistry, $\mathbf{N}_t \to \mathbf{D}_t \to \mathbf{T}_t \to \mathbf{P}_t \to (E_t, I_t) \to q_{t+1}$, whereas Panksepp version 2 uses $P_t$ before $T_t$ as follows $\mathbf{N}_t \to \mathbf{D}_t \to \mathbf{P}_t \to \mathbf{T}_t \to (E_t, I_t) \to q_{t+1}$.

All three variants were evaluated under the same protocol—20 batches, 5 agents per batch, and 100 time steps per batch—preserving the scale of the original study. At each step, living agents updated their needs, perceived their context, computed their internal variables, selected an action, and applied it. The same metrics were extracted for all variants: survival rate, state diversity, number of state changes, average reactivity lag, and action distributions, defined formally as follows.

**State Diversity:** Diversity $= |\{q_t \mid t = 1, ..., T\}|$ is used to verify the number of distinct states visited over time, measuring behavioral variety.

**State Changes:** $\text{Changes} = \sum_{t=1}^{T-1} \mathbb{I}(q_t \neq q_{t+1})$ calculate the total number of transitions between different states, capturing how frequently the system changes behavior.

**Reactivity Lag:** $\text{Lag} = \frac{1}{|S|} \sum_{s \in S} (t^s_{\text{action}} - t^s_{\text{stimulus}})$ where $S$ is the set of stimulus-response events. Estimate the average delay between a stimulus and the corresponding action, indicating responsiveness.

**Action Distribution:** $P(a) = \frac{\text{count}(a)}{\sum_{a'} \text{count}(a')}$ calculates the normalized frequency of each action, describing how actions are distributed over time.

To evaluate robustness to stochastic conditions, we first ran a multi-seed analysis with 30 seeds per model. Each seed varied the initial agent states, positions, food placement, novelty events, and combat outcomes while keeping the internal equations deterministic, letting us report means and standard deviations rather than single runs and assess stability across initializations.

We then performed a one-factor-at-a-time local parameter sweep with 10 seeds per case. The purpose of this step was not to claim biologically exact coefficients, but to identify influential parameters and verify whether the qualitative conclusions were stable under plausible local perturbations. Most coupling coefficients and need-growth rates were perturbed by $\pm 20\%$ around the baseline, that is, they were tested at $\{0.8b, b, 1.2b\}$ for a baseline value $b$. Surprise-related parameters, which are structurally more uncertain, were perturbed by $\pm 50\%$, that is, at $\{0.5b, b, 1.5b\}$. For example, a baseline hunger growth rate of $0.030$ was tested at $\{0.024, 0.030, 0.036\}$, while a baseline surprise probability of $0.080$ was tested at $\{0.040, 0.080, 0.120\}$.

Finally, we combined, for each model, the individually best survival-oriented values found in the local sweep. This combined-best test was included to measure parameter interaction effects. However, because the sweep itself was one-factor-at-a-time, this combined configuration was treated as an exploratory upper bound for survival rather than as the final recommended model.

## 5. Results

Table 1 summarizes the main comparison among the original prototype and the three executable variants. The original Godot prototype is shown as a reference row, while the three Python-based models are reported as mean $\pm$ standard deviation over 30 seeds.

**Table 1. Main comparison among the original prototype and the three executable variants.**

| Model | Survival (%) | State Div. | State Chg. | Lag | Explore (%) | SearchFood (%) | Socialize (%) | Flee (%) | Attack (%) |
|---|---|---|---|---|---|---|---|---|---|
| Paper prototype | 62.00 | – | – | – | 30.59 | 6.41 | 3.21 | 8.83 | 0.10 |
| Current Python | $87.87 \pm 3.04$ | $6.89 \pm 0.07$ | $43.45 \pm 1.09$ | $1.90 \pm 0.08$ | $7.84 \pm 0.44$ | $23.40 \pm 1.16$ | $4.40 \pm 0.17$ | $0.08 \pm 0.03$ | $0.00 \pm 0.00$ |
| Panksepp V1 | $98.70 \pm 1.12$ | $6.83 \pm 0.06$ | $44.88 \pm 0.81$ | $3.58 \pm 0.24$ | $11.17 \pm 0.37$ | $28.03 \pm 1.14$ | $32.67 \pm 1.35$ | $2.27 \pm 0.17$ | $0.00 \pm 0.00$ |
| Panksepp V2 | $98.00 \pm 1.46$ | $6.36 \pm 0.08$ | $40.17 \pm 0.71$ | $3.53 \pm 0.24$ | $3.17 \pm 0.20$ | $29.20 \pm 1.30$ | $42.85 \pm 1.52$ | $0.65 \pm 0.10$ | $0.00 \pm 0.00$ |

To validate the statistical significance of the observed differences, we performed pairwise comparisons using a two-sample t-test between the baseline and each Panksepp variant. The results indicate statistically significant improvements in survival ($p < 0.01$) for both variants compared to the deterministic baseline.

Before interpreting these numbers, we clarify how survival should be read. It measures homeostatic robustness—how effectively an agent keeps its internal needs regulated over time—and enables comparison with the original model ($62\%$). It is not, by itself, a proxy for player-perceived believability, and neither are the diversity, state-change, lag, and action-distribution metrics reported below. We therefore frame these results as evidence of robustness, stability, and behavioral balance, deferring believability claims to the human-centered studies in Section 6.

The first relevant result is that all three executable variants clearly outperform the original prototype in survival. The deterministic Python version raises survival from 62% to 87.87%, indicating that the model benefits substantially from a more stable regulation loop and from the explicit novelty filter for Surprise. This version can be interpreted as a controlled and reproducible refinement of the original proposal: it keeps the original architecture, but removes implementation noise and makes the simulation reproducible.

This $62\% \rightarrow 87.87\%$ gain, however, does not isolate determinism. The original computational logic and most preliminary tests were already in Python, with Godot used mainly for visual representation, so the gain is not an engine- or language-change artifact. Yet the reformulation altered several aspects at once—the Surprise novelty condition, the reorganization of coefficients into a configuration dictionary, and the evaluation methodology—so the comparison against $62\%$ conflates them. A matched stochastic-versus-deterministic ablation, holding all else constant, is left for future work.

The two Panksepp variants improve survival further. Version 1 reaches 98.70% while preserving high state diversity and state changes, so the affective layer improves regulation without reducing behavioral richness. Version 2 also achieves high survival (98.00%) but a more concentrated profile, with lower diversity and a larger share of *Socialize*—effective at keeping agents alive, yet overemphasizing affiliative behavior under the current coefficients.

Decomposing the total gain over the original prototype, the reformulation contributes the largest share ($+25.87$ points, $62.00\% \rightarrow 87.87\%$), while the Panksepp layer adds a smaller survival gain ($+10.83$ for V1, $+10.13$ for V2) and, more distinctively, reduces seed variance (standard deviation from $3.04$ to $1.12$ and $1.46$). Most of the mean-survival improvement therefore stems from the reformulation, whereas the affective layer mainly contributes robustness and stability.

From a behavioral standpoint, the main weakness of the original prototype becomes even clearer in comparison. In the original paper, *Explore* represented 30.59% of all decisions, far above the other states. In the deterministic Python baseline, that value drops to 7.84%, showing that the explicit novelty requirement successfully corrected the overproduction of surprise-driven exploration. At the same time, *SearchFood* rises from 6.41% in the original prototype to 23.40%, which is more coherent with a drive-based survival model. In other words, once surprise is constrained to actual novelty, the model spends more time satisfying basic needs.

The comparison between the two Panksepp variants reveals a more subtle trade-off. Version 1 produces a more balanced distribution between *SearchFood* (28.03%) and *Socialize* (32.67%), while version 2 pushes *Socialize* to 42.85% and compresses *Explore* to only 3.17%. Since both versions have almost identical survival, version 1 can be interpreted as the best overall compromise between robustness and behavioral balance, whereas version 2 is better interpreted as a highly stable but more socially biased agent policy.

Table 2 isolates the random-seed experiment, confirming that the Panksepp gains persist across 30 stochastic realizations rather than arising from a single favorable initialization: both variants pair higher mean survival with lower standard deviation than the deterministic baseline, so the affective extension not only raises average survival but also reduces sensitivity to seed changes.

**Table 2. Multi-seed robustness summary (30 seeds).**

| Model | Seeds | Survival (%) | State Div. | Lag |
|---|---|---|---|---|
| Current Python | 30 | 87.87 ± 3.04 | 6.89 ± 0.07 | 1.90 ± 0.08 |
| Panksepp V1 | 30 | 98.70 ± 1.12 | 6.83 ± 0.06 | 3.58 ± 0.24 |
| Panksepp V2 | 30 | 98.00 ± 1.46 | 6.36 ± 0.08 | 3.53 ± 0.24 |

We then ran a one-factor-at-a-time sweep with 10 seeds per case (Table 3), perturbing regular coefficients by $\{0.8b, b, 1.2b\}$ and surprise-related ones by $\{0.5b, b, 1.5b\}$ (e.g. hunger growth $0.030 \rightarrow \{0.024, 0.030, 0.036\}$; surprise probability $0.080 \rightarrow \{0.040, 0.080, 0.120\}$). The deterministic baseline was markedly more sensitive, chiefly to the hunger growth rate, the surprise-related noradrenaline boost, and the safety-to-noradrenaline coupling; both Panksepp variants were more stable but still responsive to the hunger growth rate and to SEEKING- and CARE-related coefficients. The most influential parameter was the hunger growth rate: lowering it from $0.030$ to $0.024$ raised mean survival from $88.2\%$ to $95.7\%$ in the baseline, and to $99.6\%$ and $99.7\%$ in V1 and V2.

**Table 3. Parameter-sweep summary.**

| Model | Best param. | Base | Tested range | Best Survival (%) | Combined-best (%) |
|---|---|---|---|---|---|
| Current Python | hunger_rate | 0.030 | $\{0.024, 0.030, 0.036\}$ | 95.7 ± 2.00 | 98.5 ± 1.43 |
| Panksepp V1 | hunger_rate | 0.030 | $\{0.024, 0.030, 0.036\}$ | 99.6 ± 0.70 | 99.8 ± 0.42 |
| Panksepp V2 | hunger_rate | 0.030 | $\{0.024, 0.030, 0.036\}$ | 99.7 ± 0.48 | 99.9 ± 0.32 |

However, the combined-best test shows that optimizing survival alone is not sufficient. When all individually best survival-oriented values were combined, mean survival rose to 98.5% for the current Python model, 99.8% for Panksepp version 1, and 99.9% for Panksepp version 2. These values are excellent from a survival perspective, but the behavioral distributions became less balanced. In Panksepp version 1, *Explore* rose to 43.39% and *Socialize* fell to 10.68%. In version 2, *Explore* rose to 29.97% while *Socialize* dropped to 22.12%. Therefore, the best individual values do not compose into the best global behavioral configuration.

A further limitation is the complete absence of attack behavior (0.00%) in both Panksepp variants. While it raises survival, it indicates that the current calibration underrepresents RAGE-driven aggression and may reduce realism where conflict is expected. Crucially, this absence is not exclusive to the affective layer: the deterministic baseline also produced no attacks (Table 1). Several environmental factors more likely act jointly—the restrictive resource-conflict conditions, abundant food, the Anger and RAGE thresholds of Section 4, the 2.0-unit distance requirement, and competing survival-oriented actions—and isolating each, together with adding context-sensitive aggression triggers, is left for future work. Since the original prototype still produced a small amount of attack (0.10%), this is a limitation of the present implementation rather than evidence that attack is unnecessary in the model.

These results show that optimizing survival alone is insufficient to characterize model quality: if survival were the only goal, the combined-best configurations would win. Because we seek a model that is robust, interpretable, and behaviorally balanced, the untuned multi-seed Panksepp V1 is the strongest candidate, pairing near-maximal survival with high diversity and a balanced action distribution. The deterministic baseline remains a faithful, reproducible upgrade of the original paper, while V2 shows both the potential and the social-bias risk of a more game-oriented affective controller.

## 6. Conclusion and Future Work

This paper presented a reproducible reformulation and extension of a drive-based neurochemical NPC model, showing that Panksepp-inspired affective systems improve robustness and stability under stochastic conditions. Even without Panksepp, the upgraded baseline already improved over the original Godot prototype by correcting the overproduction of Surprise and raising survival from 62% to 87.87% under a 30-seed evaluation.

The main contribution, however, is the Panksepp-inspired affective layer. Both variants reached survival close to 99% while becoming less sensitive to stochastic variation; version 1 gave the best balance of survival, diversity, and action balance, and version 2 showed that a more game-oriented formulation is also effective, though prone to social overconcentration. These results characterize homeostatic robustness and behavioral balance rather than player-perceived believability, which remains to be assessed through human evaluation.

The seed and parameter analyses qualify these conclusions: the improvements are stable rather than accidental, a few coefficients (especially hunger growth and SEEKING-related couplings) dominate behavior, and the combined-best test shows that maximizing survival alone distorts the action distribution—so tuning should be multi-objective.

A second conclusion is that the current implementations still fail to generate aggressive behavior: the absence of *Attack* indicates that the *RAGE*-, proximity-, and conflict-related thresholds remain too restrictive, a limitation to be addressed in future iterations.

Future work should proceed in five directions: (i) extend the parameter search to joint multi-objective optimization, balancing survival with behavioral balance and diversity; (ii) enrich the Panksepp layer with attachment- and social-regulation variables tied to CARE and PANIC/GRIEF; (iii) revise the conflict subsystem so that aggression emerges when the context justifies it; (iv) reintegrate the model into an interactive game engine, using the Python simulator as a calibration framework; and (v), most important for any believability claim, evaluate the model with humans—both by comparing NPC decisions against human judgments in equivalent scenarios and by measuring player-perceived believability and enjoyment in an interactive setting. The metrics reported here indicate homeostatic robustness and behavioral balance, not human-perceived believability; the study thus contributes both a stronger NPC model and a clearer methodology for evaluating emotion-driven agents.

## References


Aljammaz, R., Wardrip-Fruin, N., e Mateas, M. (2023). Towards an understanding of character believability. In *Foundations of Digital Games 2023 (FDG '23)*. ACM.

Armanto, H., Rosyid, H. A., Muladi, e Gunawan (2025). Improved non-player character (npc) behavior using evolutionary algorithm—a systematic review. *Entertainment Computing*, 52:100875.

Baffa, A., Sampaio, P., Feijó, B., e Lana, M. (2017). Dealing with the emotions of non player characters. In *Proceedings of the 16th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, pages 504–515. SBC.

Bicalho, L. F., Baffa, A., e Feijó, B. (2025). Drive-based behavior modeling for emotionally responsive npcs. In *Anais do XXIV Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames)*, pages 691–702. SBC.

Bicalho, L. F., Feijó, B., e Baffa, A. (2020). A culture model for non-player characters' behaviors in role-playing games. In *2020 19th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*, pages 9–18. IEEE.

da Silva, G. A. e de Souza Ribeiro, M. W. (2021). Development of non-player character with believable behavior: a systematic literature review. In *Proceedings of the 20th Brazilian Symposium on Computer Games and Digital Entertainment (SBGames)*. SBC.

de Tarso Fernandes, P., Baffa, A., e Feijó, B. (2024). An emotion model for non-player characters inspired by neurotransmitters. In *Anais do XXIII Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames)*, pages 547–562. SBC.

Galvão, V. F., Maciel, C., dos Santos Nunes, E. P., e da Hora Rodrigues, K. R. (2025). A framework to support the development of empathetic games. *Journal on Interactive Systems*, 16(1):25–53.

Garavaglia, F., Nobre, R. A., Ripamonti, L. A., Maggiorini, D., e Gadia, D. (2022). Moody5: Personality-biased agents to enhance interactive storytelling in video games. In *2022 IEEE Conference on Games (CoG)*, pages 175–182. IEEE.

Hull, C. L. (1943). *Principles of Behavior: An Introduction to Behavior Theory*. Appleton-Century-Crofts, New York.

Inyang, S., Hale, J., Coghlan, S., e Wyeth, P. (2024). Mental state modelling: A philosophy of mind approach to npc believability. In *Proceedings of the 36th Australian Conference on Human-Computer Interaction (OzCHI '24)*. ACM.

Kaneko, K. e Okada, Y. (2024). *Emotion-Based 3D CG Character Behaviors*, pages 645–653. Springer, Cham.

Lövheim, H. (2012). A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical Hypotheses*, 78(2):341–348.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, 50(4):370–396.

Maslow, A. H. (1954). *Motivation and Personality*. Harper & Brothers, New York.

Panksepp, J. (1998). *Affective Neuroscience: The Foundations of Human and Animal Emotions*. Oxford University Press, New York.

Panksepp, J. e Biven, L. (2012). *The Archaeology of Mind: Neuroevolutionary Origins of Human Emotions*. W. W. Norton, New York.

Pinheiro, J. R., Baffa, A., e Feijó, B. (2025). Exploring extensions of neurotransmitter-based emotion models. *Journal on Interactive Systems*, 16(1):1–24.

Plutchik, R. (1980). *Emotion: A Psychoevolutionary Synthesis*. Harper & Row, New York.

Plutchik, R. (2001). The nature of emotions. *American Scientist*, 89(4):344–350.

Shirvani, A., Ware, S. G., e Baker, L. J. (2023). Personality and emotion in strong-story narrative planning. *IEEE Transactions on Games*, 15(4):669–682.

Wampfler, R., Yang, C., Elste, D., Kovačević, N., Witzig, P., e Gross, M. (2025). A platform for interactive ai character experiences. In *SIGGRAPH Conference Papers '25*. ACM.