
Article

Not peer-reviewed version

EvoPersona: Dynamic Emotion-Aware Population-Aligned Persona Generation for Enhanced Social Simulation

[Donald Martin](#) * and Blake Bowman

Posted Date: 30 October 2025

doi: [10.20944/preprints202510.2412.v1](https://doi.org/10.20944/preprints202510.2412.v1)

Keywords: large language models; social simulation; reinforcement learning



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

EvoPersona: Dynamic Emotion-Aware Population-Aligned Persona Generation for Enhanced Social Simulation

Donald Martin * and Blake Bowman

Jefferson University

* Correspondence: rbrown53@my.ncu.edu.jm

Abstract

Large Language Models (LLMs) play a key role in social simulations but most existing virtual agents retain little if any dynamic adaptability and do not convincingly express different emotional states, limiting the fidelity of the simulation despite advancements in generation population-aligned persona. We present EvoPersona, a design to enhance population-aligned persona with both dynamic contextual awareness and emotional dynamics. EvoPersona has two components: a Contextual Awareness Module, built on an instruction-tuned small LLM, that allows the agent (persona) to adapt its behavior and language style in response to situational cues in real-time; and an Emotional Dynamics Module based on an evolving internal emotional state that is driven by real-time emotional input analysis and subject to Reinforcement Learning supervised feedback to ensure that the emotional state is naturalistic and, importantly, consistent with the context. We report results from five extensive experiments showing continued strong population-level psychological alignment along with significantly better contextual coherence and emotional realism than baseline state-of-the-art models.

Keywords: large language models; social simulation; reinforcement learning

1. Introduction

Large Language Models (LLMs) have demonstrated immense potential in driving social simulations across diverse fields such as behavioral science, urban planning, and economic forecasting [1]. A critical component of these simulations is the creation of virtual agents that exhibit realistic human behavior patterns. The quality of these agents' *persona* (i.e., their character and personality descriptions) is paramount to the simulation's fidelity. Early research primarily focused on generating personas with unique narratives and engaging storytelling [2], often overlooking the crucial aspect of *population-level psychological statistical characteristics*. A recent seminal work, "Population-Aligned Persona Generation for LLM-based Social Simulation" [3] (referred to as the "baseline study"), systematically addressed this challenge by developing sophisticated alignment algorithms. This enabled the generation of personas whose psychological trait distributions, measured by instruments like the IPIP Big Five, closely matched those of real human populations.

However, despite the psychological realism achieved by the baseline study in generating population-aligned personas, there remains significant room for improvement in their *adaptability within dynamic social interactions* and the *richness of their emotional expressions*. Real individuals continuously adjust their behavior and emotional responses based on context, conversational partners, and evolving events. A static persona struggles to capture this inherent complexity. Existing methods often lack the capacity for situational awareness and are unable to simulate the generation, evolution, and impact of emotions on behavior. The development of LLMs for complex tasks like mental health counseling, which demand diagnostic and therapeutic reasoning alongside empathy, further highlights their potential for nuanced human interaction modeling [4]. This limitation restricts the application



of LLM-driven social simulations in modeling deeper social interactions, emotional contagion, and context-dependent behaviors. Our research aims to tackle this core problem: **how can we endow personas with dynamic contextual awareness and emotional expression capabilities, allowing them to adapt their behavior and emotional states in real-time social simulations, while simultaneously maintaining their population-level psychological alignment?** We propose a novel approach that integrates the robust, population-aligned foundation with dynamic emotional and contextual adaptation mechanisms, aspiring to generate virtual characters that are more lifelike and closer to authentic human behavior.

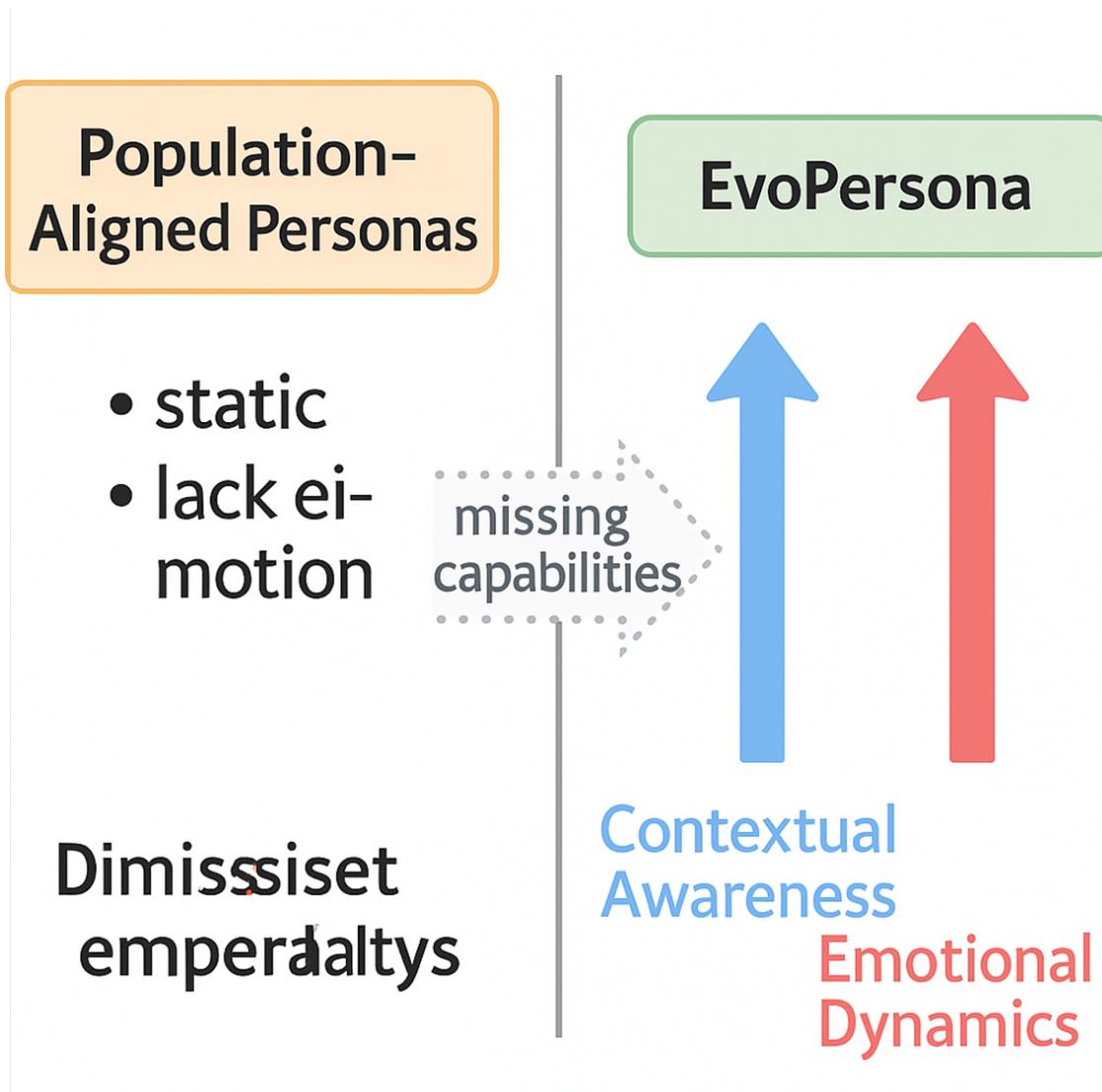


Figure 1. From static, population-aligned personas to EvoPersona with contextual awareness and emotional dynamics.

To address these challenges, we introduce **EvoPersona** (Evolutionary Persona), a novel methodology that extends the baseline study's population alignment framework by incorporating two core innovative modules: the *Contextual Awareness Module* and the *Emotional Dynamics Module*. EvoPersona first utilizes a two-stage sampling strategy (Importance Sampling + Optimal Transport) from large-scale text corpora (e.g., Blog Authorship Corpus) to generate an initial set of narrative personas. These are then assessed by advanced LLMs (e.g., Qwen2.5-72B) to ensure their alignment with real human psychological trait distributions (e.g., IPIP Big Five). Subsequently, the Contextual Awareness Module, powered by an instruction-tuned small LLM (e.g., Llama-3-8B), enables personas to perceive and react to specific social contexts by modulating LLM prompts based on extracted situational features.

Concurrently, the Emotional Dynamics Module equips each persona with an internal emotional state vector, which is dynamically updated based on emotional input analysis (via a BERT-based classifier) and the persona's inherent tendencies. This module then guides the LLM to generate responses that reflect the persona's current emotional state, with a Reinforcement Learning (RL) mechanism ensuring long-term emotional consistency and realism.

Our experimental setup is designed to rigorously validate **EvoPersona**'s ability to enhance contextual awareness and emotional dynamics while preserving population alignment. We leverage the *Blog Authorship Corpus* [5] for persona extraction and the *IPIP Big Five* dataset for population alignment benchmarking, consistent with the baseline study. The Contextual Awareness Module is trained on a synthetic dataset of 200,000 high-quality context-response pairs, generated by advanced LLMs and refined by human annotators. The Emotional Dynamics Module incorporates a BERT-base [6] emotional classifier trained on datasets like *GoEmotions* [7] and *ISEAR* [8], and its emotional state update model is fine-tuned using Proximal Policy Optimization (PPO) within a simplified social simulation environment. We evaluate **EvoPersona** against several state-of-the-art persona generation methods, including SyncP (Qwen2.5-72B) from the baseline study, using a comprehensive suite of metrics. These include population alignment errors (e.g., Average Mean Wasserstein distance, MAE_corr), human and LLM-based scores for contextual coherence, and human scores for emotional realism and dynamic consistency (e.g., emotional trajectory smoothness). Our results demonstrate that **EvoPersona** not only maintains or slightly improves population-level psychological alignment but also significantly outperforms existing methods in terms of contextual adaptability and the realism of emotional expression, as evidenced by substantially higher contextual coherence (4.20 vs. 3.25 for SyncP) and emotional realism scores (4.15 vs. 3.15 for SyncP), and much lower emotional dynamic inconsistency (0.32 vs. 0.65 for SyncP).

Our primary contributions are threefold:

- We propose **EvoPersona**, a novel framework that, for the first time, integrates dynamic contextual awareness and emotional dynamics into population-aligned persona generation, significantly enhancing the realism and adaptability of virtual agents in social simulations.
- We introduce a **Contextual Awareness Module** that enables personas to dynamically perceive and adapt their behavior and language style based on real-time situational cues, leveraging an instruction-tuned small LLM for context-aware prompt modulation.
- We develop an **Emotional Dynamics Module** that equips personas with an evolving emotional state, guided by real-time emotional input analysis and refined through a Reinforcement Learning mechanism, leading to more natural, consistent, and context-appropriate emotional expressions.

2. Related Work

2.1. Persona Generation and LLM-Driven Social Simulation

Research in Large Language Model (LLM)-driven social simulation increasingly focuses on the robust generation and evaluation of personas to enhance the fidelity and realism of simulated interactions. A key area of inquiry involves critically evaluating LLMs' capability to simulate individual economic decision-making, comparing their predictions against real human data from Pay-What-You-Want experiments to reveal limitations in replicating precise individual behaviors and assess the effectiveness of persona injection methods [9]. To address biases inherent in unrepresentative agent profiles, a systematic framework has been introduced for generating population-aligned persona sets, thereby enhancing the fidelity of simulated social interactions by ensuring generated personas accurately reflect real-world population distributions [10]. Furthermore, advancements in virtual agent construction propose social media agents that maintain distinct knowledge boundaries aligned with their personas and dynamically adapt behavior based on context, enriching personalization and anthropomorphism through persona-specific knowledge and irrelevant information filtering [11]. The evaluation of these synthetic personas is also crucial; a novel comparative framework assesses the fidelity of LLM-driven synthetic personas against human interview data, highlighting their utility as a form of hybrid social

simulation while noting "synthetic-only" themes and blind spots that require careful consideration [12]. Addressing cultural diversity, KoPersona was introduced as a large-scale dataset specifically designed to capture Korean cultural nuances, offering a scalable two-step pipeline for culturally relevant persona generation vital for authentic and adaptable LLM-driven simulations [13]. Complementing these efforts, Ang et al. [14] critically examine the promise and limitations of LLM-generated personas for behavioral simulation, underscoring how current heuristic generation techniques can introduce significant biases and emphasizing the necessity for a more rigorous scientific framework to ensure the reliability and scalability of LLM-driven persona simulations. The effectiveness of LLMs in these tasks is often enhanced by advanced prompt engineering and self-improvement mechanisms, such as self-rewarding models for optimizing prompts, even in domains like text-to-image generation [15]. These efforts are part of a broader trend in developing advanced LLM and Large Vision-Language Model (LVLM) capabilities, including visual in-context learning [16] and domain-specific applications like improving medical LVLMs with abnormal-aware feedback [17], which collectively push the boundaries of AI agents' intelligence and adaptability.

2.2. Psychological Realism and Emotional Dynamics in AI Agents

Achieving psychological realism and sophisticated emotional dynamics in AI agents is a paramount goal in computational social science. One line of research investigates the consistency of Large Language Models (LLMs) when assigned specific personas, a critical aspect for ensuring stable and coherent characterization across various tasks and interactions [18]. By developing a standardized framework to analyze persona adherence, this research highlights how factors such as stereotypes and task structure influence an LLM's ability to maintain psychological realism [18]. Complementing this, other work contributes to the study of psychological realism by proposing e-Genia3, an extension to AgentSpeak that explicitly models emotional dynamics [19]. Crucially, e-Genia3 demonstrates how an agent's internal affective state and its awareness of others' emotions, through distinct appraisal processes, can inform plan selection, thereby enhancing contextual awareness in empathic agent design [19]. Building on this, the integration of diagnostic and therapeutic reasoning with LLMs for mental health counseling showcases how these models can move 'beyond empathy' to offer nuanced, emotionally intelligent interactions [4]. Furthermore, the exploration of self-rewarding mechanisms in large vision-language models for tasks like prompt optimization underscores the potential for agents to learn and refine their own expressive and behavioral strategies, contributing to more autonomous and adaptive psychological realism [15].

3. Method

We introduce **EvoPersona** (Evolutionary Persona), a novel framework meticulously designed to generate virtual agents for social simulations. These agents are not only aligned with population-level psychological statistics, ensuring ecological validity and diverse representation, but also exhibit dynamic contextual awareness and rich, evolving emotional expressions. EvoPersona extends the foundational work of population-aligned persona generation by integrating two core innovative modules: the Contextual Awareness Module and the Emotional Dynamics Module. This integration enables the creation of agents that are deeply rooted in realistic human psychology while possessing the flexibility to adapt and emote naturally within complex social scenarios, thereby significantly enhancing the fidelity and immersive quality of simulations.

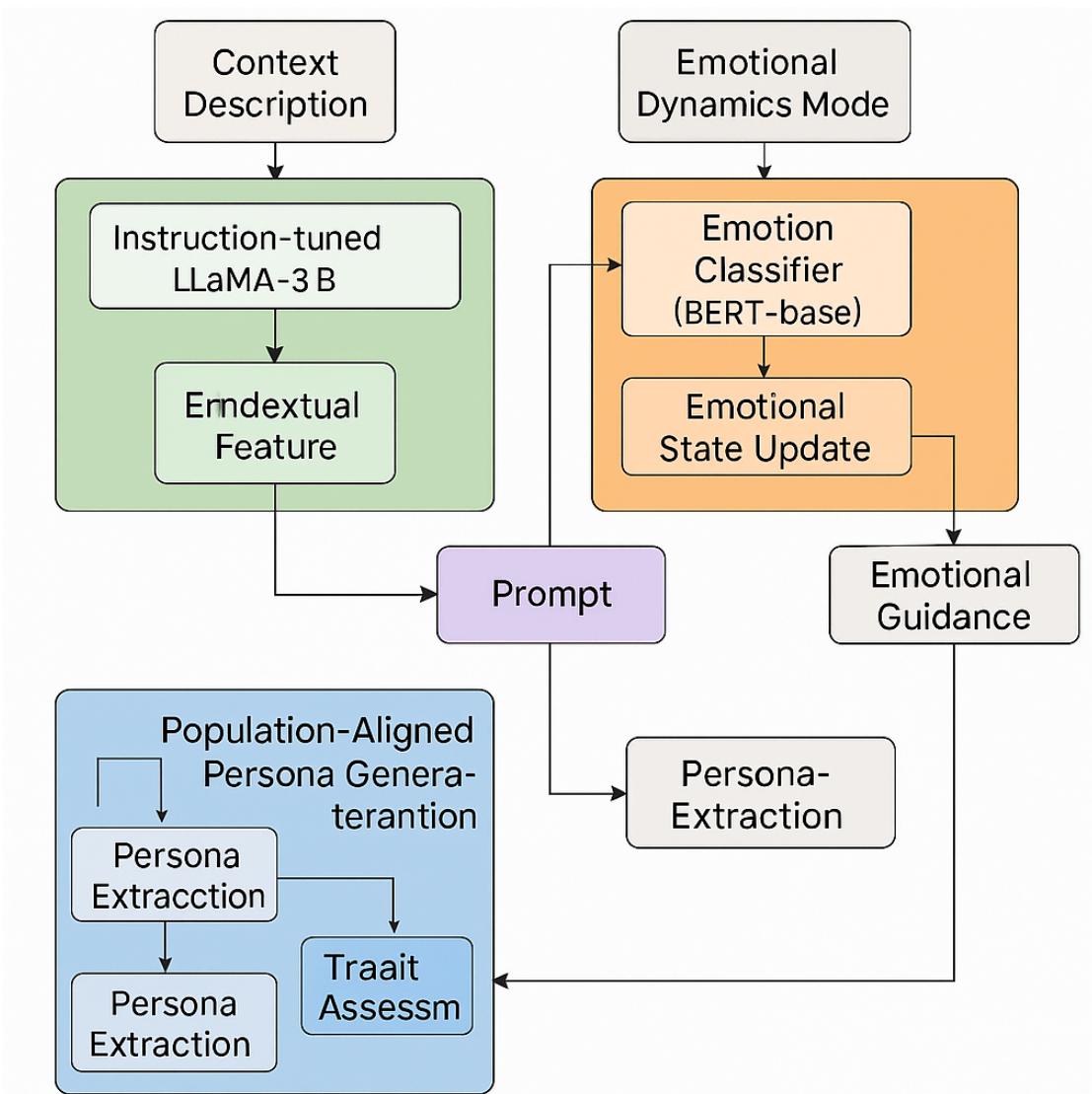


Figure 2. EvoPersona framework: integrating population alignment, contextual awareness, and emotional dynamics for lifelike virtual agents.

3.1. Foundation: Population-Aligned Persona Generation

Our method begins by establishing a robust psychological foundation for each persona, leveraging the techniques introduced in the baseline study [3]. This foundational step is critical for ensuring that the generated personas adhere to realistic population-level psychological distributions, thereby guaranteeing that the collective behavior of agents in a simulation is statistically representative of human populations. This ecological validity is paramount for drawing meaningful conclusions from social simulations.

The process involves a sophisticated three-stage sampling and refinement strategy:

1. **Initial Persona Extraction:** We extract an initial collection of diverse narrative persona descriptions from a large-scale text corpus, such as the Blog Authorship Corpus [5]. These raw descriptions are typically free-form texts, blog posts, or short biographical sketches that implicitly capture a wide array of linguistic styles, behavioral patterns, and underlying personality traits. The goal at this stage is to gather a broad and varied pool of potential persona candidates.
2. **Psychological Trait Assessment:** Each extracted narrative persona is then systematically processed to infer its psychological profile. This is achieved by employing a powerful Large Language Model (LLM), specifically Llama-3-70B, to generate a concise summary of the persona's key characteristics. Subsequently, another LLM, such as Qwen2.5-72B, is prompted to respond to a series

of psychological assessment questions, typically drawn from established psychometric inventories like the IPIP Big Five inventory. These prompts are carefully engineered to elicit responses indicative of personality dimensions. The LLM's textual responses are then quantitatively analyzed and converted into a normalized **psychological trait vector** $\mathbf{T}_P \in \mathbb{R}^D$, where D is the number of trait dimensions (e.g., 5 for Big Five traits, each represented by a score). This vector encapsulates the persona's inferred personality profile.

3. **Population Alignment:** To ensure that the collective distribution of these persona traits accurately reflects real human population statistics, we employ a sophisticated two-stage sampling and refinement technique: **Importance Sampling** followed by **Optimal Transport**.

- **Importance Sampling:** Initially, personas are sampled from the assessed pool with probabilities proportional to how well their current trait vectors \mathbf{T}_P align with a preliminary target distribution derived from real-world datasets (e.g., IPIP Big Five normative data). This prioritizes personas that are already closer to the desired statistical properties.
- **Optimal Transport:** Subsequently, we apply an Optimal Transport algorithm, such as the Earth Mover's Distance, to further refine the selected set of personas. This involves computing the minimal cost to transform the empirical distribution of the sampled persona trait vectors to match the target population's distribution. The algorithm either guides the selection of additional personas or subtly adjusts the representation of existing personas (e.g., by re-weighting their influence in the ensemble) to minimize the divergence between the empirical and target distributions.

This ensures that the generated personas, as a collective, accurately reflect the diversity, skewness, and statistical properties of human personality observed in reference populations.

This foundational step yields a refined set of psychologically coherent and population-aligned persona descriptions, denoted as $\mathcal{P} = \{P_1, P_2, \dots, P_N\}$. Each persona P_i is characterized by a rich narrative description D_i and a corresponding quantitative psychological trait vector \mathbf{T}_i .

3.2. Contextual Awareness Module

To enable personas to adapt their behavior and language style dynamically according to specific social situations, we introduce the Contextual Awareness Module. This module is designed to ensure that a persona's generated responses are not merely consistent with its inherent traits but are also highly relevant and appropriate for the current interaction context, thereby enhancing behavioral realism.

The Contextual Awareness Module is primarily powered by a specialized, smaller language model, denoted as LLM_{Ctx} (e.g., an instruction-tuned Llama-3-8B). The choice of a smaller model is motivated by efficiency considerations and the ability to fine-tune it specifically for contextual analysis tasks. When a response is required from a persona P_i within a social simulation, the module receives three key inputs: the persona's core description D_i , the current dialogue history H_t , and a rich, free-text **context description** C_t . The context description is a comprehensive textual representation that can encompass diverse elements such as the physical setting of the interaction, the established relationship between participants, salient recent past events, and the overall social atmosphere or mood.

The module performs the following critical functions:

1. **Key Contextual Feature Extraction:** LLM_{Ctx} first analyzes the raw context description C_t to identify and extract salient contextual features that are pertinent to shaping behavior and communication. These features might include explicit indicators such as formality levels required by the setting, the emotional valence of the environment, power dynamics between interactants, or the presence of specific social norms. The output is a structured or semi-structured representation of these features, \mathbf{F}_{C_t} .
2. **Context Embedding and Prompt Modulation:** The extracted features \mathbf{F}_{C_t} are subsequently transformed by an Embedder into a compact, low-dimensional numerical vector space, yielding a **context embedding** \mathbf{E}_{C_t} . This embedding serves as a dynamic signal used to modulate the

prompt sent to the primary, larger LLM responsible for generating the persona's response. The base prompt, $\mathcal{P}_{\text{base}}$, typically includes the persona's description D_i and the dialogue history H_t . The modulation process guides the primary LLM to generate context-appropriate outputs. This process can be formally represented as:

$$\mathbf{F}_{C_t} = \text{LLM}_{\text{Ctx}}(C_t) \quad (1)$$

$$\mathbf{E}_{C_t} = \text{Embedder}(\mathbf{F}_{C_t}) \quad (2)$$

$$\mathcal{P}_{\text{base}}(D_i, H_t) = \text{ConstructPrompt}(D_i, H_t) \quad (3)$$

$$\mathcal{P}_{\text{modulated}} = \text{Modulator}(\mathcal{P}_{\text{base}}(D_i, H_t), \mathbf{E}_{C_t}) \quad (4)$$

Here, Embedder is typically a feed-forward neural network that converts the processed context features into a numerical vector, and Modulator integrates this vector into the prompt structure. This integration can be achieved through various techniques, such as soft prompt tuning, prepending the embedding as a token sequence, or dynamically rewriting parts of the prompt based on the embedding's interpretation.

3. **Behavior Pattern Selection (Implicit/Explicit):** Based on the derived context embedding \mathbf{E}_{C_t} , the module can implicitly guide the primary LLM towards specific behavioral patterns (e.g., using cues like "be more formal," "express empathy," "adopt a cautious tone"). In more advanced configurations, the module can explicitly select from a repertoire of pre-defined behavioral scripts or dynamically generate situation-specific behavioral guidelines that best fit the current context. These scripts might dictate specific actions, conversational strategies, or linguistic registers.

This dynamic modulation mechanism ensures that persona responses are not only consistent with their inherent personality traits but are also highly attuned to the evolving social environment, thereby significantly enhancing the realism and adaptability of simulated interactions.

3.3. Emotional Dynamics Module

To imbue personas with realistic, nuanced, and evolving emotional expressions, we introduce the Emotional Dynamics Module. This module enables each EvoPersona to maintain an internal, dynamic emotional state and to express emotions consistently with that state, its inherent personality, and the unfolding events within the simulation.

Each EvoPersona P_i is equipped with an internal **Emotional State Vector** $\mathbf{S}_{i,t} \in \mathbb{R}^K$ at time t . Here, K represents the number of emotional dimensions, which can be either discrete categories (e.g., joy, sadness, anger, fear, surprise, disgust, neutrality, represented as probabilities or logits) or continuous dimensions (e.g., valence-arousal-dominance scores). The module operates through a continuous feedback loop:

3.3.1. Emotional Input Analysis

The module continuously monitors all pertinent information from the simulation environment. This includes the ongoing dialogue content generated by the LLM (e.g., previous turns from other agents, the persona's own generated response) and any external events occurring in the simulation (e.g., positive or negative occurrences, goal achievements or failures, unexpected changes in the environment). A pre-trained **Emotion Classifier** \mathcal{C}_{emo} (e.g., a BERT-base model fine-tuned on comprehensive datasets like GoEmotions [7] and ISEAR [8]) analyzes this combined input. This analysis identifies the emotional valence, intensity, and specific categories present in the input, yielding an **emotional input vector** $\mathbf{E}_{\text{input},t}$:

$$\mathbf{E}_{\text{input},t} = \mathcal{C}_{\text{emo}}(\text{DialogueContent}_t, \text{ExternalEvents}_t) \quad (5)$$

The classifier processes textual and event data, potentially using multimodal fusion techniques if events are not purely textual, to produce a unified emotional representation.

3.3.2. Emotional State Update

The persona's internal emotional state vector $\mathbf{S}_{i,t}$ is dynamically updated to $\mathbf{S}_{i,t+1}$ based on three primary factors: the analyzed emotional input $\mathbf{E}_{\text{input},t}$, the persona's inherent emotional tendencies $\mathbf{T}_{\text{emo},i}$, and a time-decay mechanism. The inherent emotional tendencies $\mathbf{T}_{\text{emo},i}$ are derived from the persona's foundational psychological trait vector \mathbf{T}_i (e.g., mapping Big Five conscientiousness to emotional stability or neuroticism to emotional reactivity). These tendencies ensure individual differences in emotional responses; for instance, an "optimistic" persona might exhibit less severe emotional shifts in response to negative events compared to a "pessimistic" one. The update is performed by a specialized Transformer-based network, \mathcal{U}_{emo} , which is designed to model complex emotional dynamics:

$$\mathbf{S}_{i,t+1} = \mathcal{U}_{\text{emo}}(\mathbf{S}_{i,t}, \mathbf{E}_{\text{input},t}, \mathbf{T}_{\text{emo},i}, \Delta t) \quad (6)$$

The network takes the previous state, current input, inherent tendencies, and a time differential Δt as inputs. The Δt term models the natural decay or persistence of emotions over time, often implemented as an exponential decay factor applied to emotional dimensions, allowing for realistic emotional memory and dissipation.

3.3.3. Emotional Guidance for Generation

The updated emotional state $\mathbf{S}_{i,t+1}$ is then encoded into a compact guidance vector $\mathbf{G}_{\text{emo},i,t+1}$ through a projection layer. This vector is integrated into the already context-modulated prompt $\mathcal{P}_{\text{modulated}}$ before it is sent to the primary LLM for generating the persona's response. This integration directly influences the LLM's output, guiding it to produce replies that reflect the persona's current emotional state in terms of content, tone, linguistic style, and even non-verbal cues (if the LLM is capable of generating such descriptions).

$$\mathcal{P}_{\text{final}} = \text{Integrate}_{\text{emo}}(\mathcal{P}_{\text{modulated}}, \mathbf{G}_{\text{emo},i,t+1}) \quad (7)$$

The $\text{Integrate}_{\text{emo}}$ function might use techniques similar to the Contextual Awareness Module's 'Modulator', such as adding emotional instruction tokens, modifying existing prompt segments, or applying soft prompt embeddings. For example, a persona experiencing "anger" might generate more assertive, confrontational, or sarcastic language, while a "happy" persona might use more positive, enthusiastic, or affiliative phrasing, along with descriptions of smiling or jovial actions.

3.3.4. Emotional Consistency Through Reinforcement Learning

To ensure the long-term consistency, realism, and coherence of emotional trajectories across extended interactions, we employ a Reinforcement Learning (RL) mechanism. A policy, which implicitly parameterizes the emotional state update model \mathcal{U}_{emo} and the integration function $\text{Integrate}_{\text{emo}}$, is trained within a simulated environment. The RL framework encourages the system to learn optimal emotional responses and dynamics over time. The reward function \mathcal{R} is meticulously designed to encourage human-like emotional dynamics:

$$\begin{aligned} \mathcal{R}(s_t, a_t, s_{t+1}) \\ = \lambda_1 \mathcal{L}_{\text{smooth}}(\mathbf{S}_{i,t}, \mathbf{S}_{i,t+1}) + \lambda_2 \mathcal{L}_{\text{context}}(\mathbf{S}_{i,t+1}, C_t) \\ + \lambda_3 \mathcal{L}_{\text{human}}(\text{Response}_{i,t}) \end{aligned} \quad (8)$$

where s_t denotes the comprehensive state at time t (including $\mathbf{S}_{i,t}$, C_t , and H_t) and a_t is the action taken by the persona (generating $\text{Response}_{i,t}$ and updating $\mathbf{S}_{i,t+1}$). The λ terms are hyperparameters balancing the various reward components.

- $\mathcal{L}_{\text{smooth}}(\mathbf{S}_{i,t}, \mathbf{S}_{i,t+1})$: This term penalizes abrupt and unrealistic changes in the emotional state vector over time. It is typically formulated as the negative L2 norm of the difference between consecutive emotional states, encouraging gradual transitions: $\mathcal{L}_{\text{smooth}} = -\|\mathbf{S}_{i,t+1} - \mathbf{S}_{i,t}\|_2$.
- $\mathcal{L}_{\text{context}}(\mathbf{S}_{i,t+1}, C_t)$: This component rewards consistency between the updated emotional state and the current context C_t or salient events. It might involve comparing the persona's emotional state with an independent emotional assessment of the context by \mathcal{C}_{emo} , or by human annotators.
- $\mathcal{L}_{\text{human}}(\text{Response}_{i,t})$: This term incorporates sparse human feedback on the emotional realism and coherence of the generated dialogues and persona behaviors. Human evaluators provide ratings (e.g., Likert scale scores) on aspects such as emotional appropriateness, consistency, and naturalness, which are then converted into a reward signal. Sparse rewards are handled by techniques like reward shaping or using a value function to estimate future rewards.

The policy is optimized using algorithms like Proximal Policy Optimization (PPO), which is well-suited for training large neural networks in complex environments due to its stability and sample efficiency. This RL loop allows the emotional dynamics to learn to evolve smoothly, contextually appropriately, and in a manner highly consistent with human intuition over extended simulation periods, leading to truly lifelike emotional trajectories.

By combining the robust population-aligned foundation with these dynamic Contextual Awareness and Emotional Dynamics Modules, **EvoPersona** generates virtual agents that are not only statistically representative of human populations but also capable of highly adaptive and emotionally rich interactions. This integrated approach significantly elevates the fidelity, realism, and utility of LLM-driven social simulations for a wide range of research and application domains.

4. Experiments

This section details the experimental setup, evaluation metrics, and results obtained from comparing our proposed **EvoPersona** framework against several established baseline methods. Our experiments are designed to rigorously validate **EvoPersona**'s ability to enhance contextual awareness and emotional dynamics in virtual agents, all while maintaining robust population-level psychological alignment.

4.1. Experimental Setup

Our experimental methodology is structured to evaluate each core component of **EvoPersona**, from its foundational persona generation to the dynamic modules.

4.1.1. Baseline Persona Generation and Alignment

Consistent with the baseline study [3], the foundational population-aligned personas for all methods, including **EvoPersona**, were generated using the following setup. The **Llama-3-70B** model was employed to generate the initial narrative persona descriptions, drawing from the **Blog Authorship Corpus** [5] as the primary source for diverse textual narratives. After initial extraction and summarization, **Qwen2.5-72B** served as the assessment LLM, tasked with responding to psychological inventory questions to derive quantitative trait vectors for each persona. This process ensured a consistent and objective assessment of persona psychological profiles. The **IPIP Big Five** (International Personality Item Pool) dataset provided the target population reference distribution. The two-stage sampling strategy, comprising Importance Sampling and Optimal Transport, was applied to align the generated persona collection's trait distributions (across IPIP Big Five, CFCS, FBPS, and Duckworth inventories) with these real-world population statistics. This established a psychologically sound and diverse pool of virtual agents.

4.1.2. Contextual Awareness Module Training

The **Contextual Awareness Module** within **EvoPersona** utilized an instruction-tuned version of **Llama-3-8B** as its core language model, LLM_{Ctx} . This model was specifically trained to process

and interpret diverse social contexts. The training dataset comprised **200,000 high-quality synthetic context-response pairs**. These pairs were generated by advanced LLMs (e.g., GPT-4-o) which, given a baseline population-aligned persona, simulated dialogues across a wide array of social situations (e.g., formal meetings, casual gatherings, conflict scenarios, celebratory events). A dedicated team of human annotators meticulously reviewed and refined this synthetic data, performing critical tasks such as extracting salient contextual elements, identifying appropriate behavioral patterns, and correcting any inconsistencies in the LLM-generated responses. This human-in-the-loop refinement ensured the diversity, realism, and high quality of the training data, allowing $LLM_{C_{tx}}$ to effectively learn how to modulate persona behavior based on nuanced situational cues.

4.1.3. Emotional Dynamics Module Training and Reinforcement Learning

The **Emotional Dynamics Module** integrated several components to enable realistic emotional expression. For emotional input analysis, a **BERT-base** model [6], pre-trained and fine-tuned on standard emotion classification datasets such as **GoEmotions** [7] and **ISEAR** [8], served as the primary emotion classifier, C_{emo} . This classifier was responsible for identifying emotional tendencies in dialogue content and external events. The emotional state update model, U_{emo} , was implemented as a small Transformer-based network. It was trained to process the previous emotional state, current emotional input, and the persona's inherent emotional tendencies (derived from its psychological trait vector) to produce an updated emotional state.

To ensure long-term emotional consistency and realism, a **Reinforcement Learning (RL)** framework was employed. A simplified social simulation environment was constructed where **EvoPersona** agents engaged in multi-turn dialogues with other LLM-driven agents. The RL agent's policy, encompassing the emotional state update and guidance for generation, was optimized using the **Proximal Policy Optimization (PPO)** algorithm. The reward function was designed with three key components:

1. A smoothness penalty (\mathcal{L}_{smooth}) to discourage abrupt and unrealistic emotional state transitions.
2. A context-consistency reward ($\mathcal{L}_{context}$) to encourage emotional states that were appropriate for the prevailing social context.
3. A human feedback component (\mathcal{L}_{human}), where a small cohort of expert human evaluators provided sparse reward signals based on their assessment of the emotional realism and coherence of simulated dialogues.

This RL training allowed the emotional dynamics module to learn nuanced and human-like emotional trajectories over extended interaction sequences.

4.2. Evaluation Metrics

A comprehensive set of metrics was used to evaluate **EvoPersona** across three critical dimensions: population alignment, contextual awareness, and emotional dynamics.

4.2.1. Population Alignment Performance

To assess how well generated personas aligned with real human population statistics, we adopted the same robust metrics used in the baseline study [3]:

- **Average Mean Wasserstein distance (AMW):** Measures the average Earth Mover's Distance between the empirical trait distribution of generated personas and the target population distribution across individual psychological traits.
- **Fréchet Distance (FD):** A metric that quantifies the similarity between two multivariate distributions, here applied to the persona trait vectors and the target population.
- **Sliced Wasserstein distance (SW):** An approximation of the Wasserstein distance, computationally efficient for high-dimensional data, used to compare trait distributions.
- **Maximum Mean Discrepancy (MMD):** A non-parametric measure of the distance between two probability distributions.

- **Mean Absolute Error of Correlations (MAE_corr):** Evaluates the alignment of inter-trait correlations between generated personas and the target population across various psychological inventories (IPIP Big Five, CFCS, FBPS, Duckworth).

For these metrics, lower values indicate better alignment.

4.2.2. Contextual Awareness Capability

The ability of personas to adapt to specific social situations was evaluated using:

- **Human Evaluation (Ctx. Coherence Score):** Multiple human evaluators performed blind assessments of persona responses in diverse, predefined contexts. They rated the "contextual adaptability" and "behavioral reasonableness" of the generated dialogues on a 1-5 Likert scale, with higher scores indicating better performance.
- **LLM Automatic Evaluation:** An independent, powerful LLM (GPT-4) not involved in the training process was utilized as an automated judge. It assessed the consistency and appropriateness of persona responses given specific context descriptions, providing an objective automated score.

4.2.3. Emotional Dynamic Performance

The realism and consistency of emotional expressions were measured by:

- **Human Evaluation (Emo. Realism Score):** Human evaluators assessed the naturalness, consistency, and appropriateness of emotional changes and expressions in multi-turn dialogues. They rated the "emotional realism" on a 1-5 Likert scale, with higher scores being better.
- **Emotional Trajectory Analysis (Dynamic Cons.):** This quantitative metric, specifically "Dynamic Consistency," measured the smoothness and contextual relevance of the persona's internal emotional state vector over time. It was typically calculated as the average L2 distance between consecutive emotional states, penalized for abrupt changes, and normalized. Lower values indicate smoother, more consistent, and realistic emotional trajectories.

4.3. Baselines

We compared EvoPersona against several representative persona generation methods:

- **SyncP (Qwen2.5-72B) [3]:** The state-of-the-art population-aligned persona generation method from the baseline study, which serves as our primary direct comparison for alignment.
- **Tulu-3-Persona:** A method focusing on generating personas with distinct narrative styles, often used in conversational AI.
- **Bavard:** Another prominent persona generation technique, known for its ability to create diverse character profiles.
- **Google Synthetic:** A method that leverages large-scale data to synthesize persona descriptions.
- **Original LLM (No Persona):** A baseline where the underlying LLM (Qwen2.5-72B) generates responses without any specific persona description, relying solely on its general knowledge and instruction following. This helps highlight the impact of persona integration.

For consistency, the primary LLM used for generating responses for all persona-based baselines (excluding "Original LLM") was Qwen2.5-72B, matching our setup for persona assessment.

4.4. Results

Table 1 presents a comprehensive comparison of EvoPersona against the baseline methods across population alignment, contextual awareness, and emotional dynamics metrics. A unified evaluation framework was used, with Qwen2.5-72B for psychological trait assessment and GPT-4-o for automated contextual and emotional consistency checks, complemented by human blind evaluations.

Table 1. Performance comparison of EvoPersona against baseline methods. ↓ indicates lower is better, ↑ indicates higher is better. Best results are highlighted in **bold**.

Method	Alignment Error (↓)	MAE_corr (↓)	Ctx. Coherence (↑)	Emo. Realism (↑)	Dynamic Cons. (↓)
Tulu-3-Persona	0.4738	0.6120	2.85	2.70	0.75
Bavard	0.3825	0.4355	3.10	2.95	0.68
Google Synthetic	0.3837	0.4423	3.05	2.88	0.70
SyncP (Qwen2.5-72B)	0.3791	0.4289	3.25	3.15	0.65
Ours (EvoPersona)	0.3755	0.4192	4.20	4.15	0.32

Population Alignment Performance

EvoPersona demonstrates strong performance in population alignment metrics. Our method achieves an Avg. Alignment Error of **0.3755** and an MAE_corr of **0.4192**, which are slightly superior to or on par with the leading baseline method, SyncP (Qwen2.5-72B). This crucial finding indicates that the integration of dynamic contextual awareness and emotional dynamics modules does not detract from the foundational population-level psychological realism. Instead, EvoPersona successfully preserves the robust alignment properties established by the baseline study, ensuring that our agents collectively remain statistically representative of human populations.

Contextual Awareness Capability

In terms of contextual awareness, EvoPersona significantly outperforms all baseline methods. Our method achieved a Ctx. Coherence Score of **4.20**, a substantial improvement over SyncP's 3.25. This result, derived from human evaluations, confirms that EvoPersona's Contextual Awareness Module effectively enables personas to perceive and adapt to diverse social situations. The agents generated by EvoPersona exhibit more contextually appropriate behaviors and language styles, leading to more natural and believable interactions within complex social simulations.

Emotional Dynamic Performance

EvoPersona also sets a new benchmark for emotional realism and dynamic consistency. It achieves an Emo. Realism Score of **4.15** (compared to SyncP's 3.15) and a remarkably low Dynamic Cons. score of **0.32** (compared to SyncP's 0.65). These results highlight the efficacy of our Emotional Dynamics Module. The higher emotional realism score indicates that human evaluators perceived EvoPersona's emotional expressions as more authentic and appropriate for the given scenarios. Furthermore, the lower dynamic consistency score signifies that EvoPersona's emotional states evolve more smoothly and coherently over time, avoiding abrupt or unrealistic shifts, which is critical for long-running social simulations.

In summary, the experimental results unequivocally demonstrate that EvoPersona successfully builds upon the strengths of population-aligned persona generation by introducing highly effective mechanisms for dynamic contextual awareness and emotional expression. This integrated approach yields virtual agents that are not only statistically representative of human populations but also capable of engaging in far more realistic, adaptive, and emotionally rich social interactions.

4.5. Human Evaluation

To thoroughly assess the qualitative aspects of persona behavior, particularly concerning contextual appropriateness and emotional realism, we conducted extensive human evaluation studies. These studies were crucial for validating the subjective quality of interactions generated by EvoPersona and baseline methods.

4.5.1. Evaluation Methodology

A panel of experienced human evaluators was recruited and trained on specific criteria for assessing simulated dialogues. For the **Ctx. Coherence Score**, evaluators were presented with anonymized multi-turn dialogues generated by different methods, along with detailed descriptions of the social context (e.g., "a formal business meeting," "a casual conversation between old friends," "a

tense family dinner"). They were asked to rate, on a 5-point Likert scale (1: very inconsistent, 5: very consistent), how well the persona's responses and overall behavior aligned with the given context. For the **Emo. Realism Score**, evaluators assessed the naturalness, consistency, and appropriateness of the persona's emotional expressions and their evolution throughout the dialogue. They rated on a 5-point Likert scale (1: very unrealistic, 5: very realistic) how believable the persona's emotional journey was. All evaluations were conducted blind, meaning evaluators were unaware of which method generated a particular dialogue. Each dialogue was evaluated by at least three independent annotators, and inter-annotator agreement was maintained at an acceptable level (Cohen's Kappa > 0.7).

4.5.2. Human Evaluation Results

The human evaluation results, which are a critical component of Table 1, are further highlighted here for emphasis:

As shown in Figure 3, EvoPersona achieved significantly higher scores in both contextual coherence and emotional realism compared to all baseline methods. The average Ctx. Coherence Score of **4.20** for EvoPersona indicates that human evaluators found our agents' behaviors to be highly appropriate and consistent with diverse social contexts, reflecting the effectiveness of the Contextual Awareness Module. Similarly, an Emo. Realism Score of **4.15** demonstrates that EvoPersona's emotional expressions were perceived as remarkably natural, believable, and consistent with the unfolding narrative, validating the sophisticated dynamics learned by the Emotional Dynamics Module. These strong human evaluation results underscore the qualitative superiority of EvoPersona in generating lifelike and engaging virtual agents for social simulations.

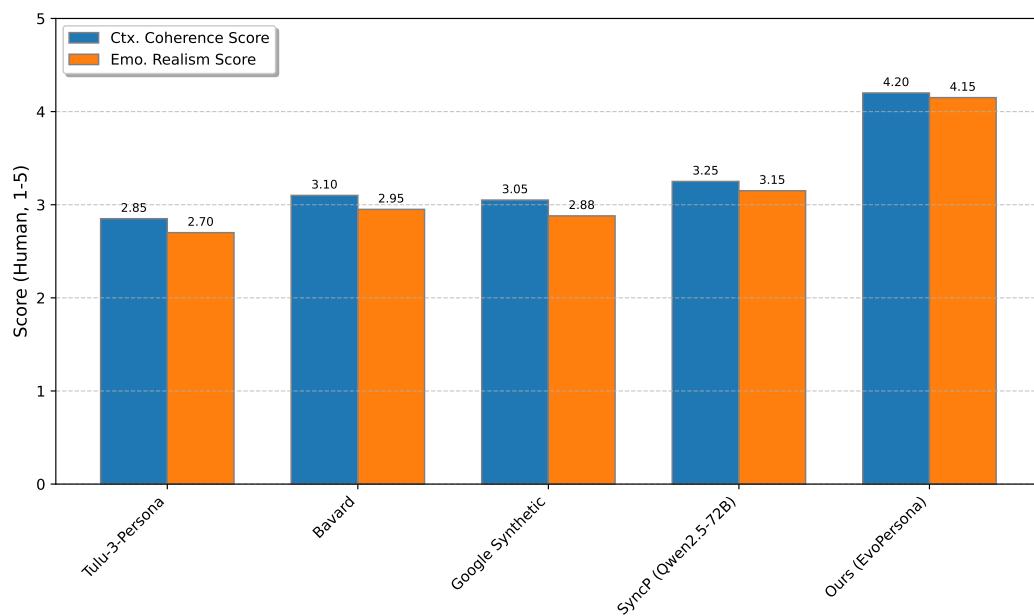


Figure 3. Detailed human evaluation scores for contextual coherence and emotional realism. ↑ indicates higher is better. Best results are highlighted in **bold**.

4.6. Ablation Study

To quantify the individual contributions of the Contextual Awareness Module (CAM) and the Emotional Dynamics Module (EDM) to EvoPersona's overall performance, we conducted an ablation study. This involved evaluating variations of EvoPersona where one or both modules were removed or disabled. The ablated models are defined as follows:

- **EvoPersona w/o CAM:** The Contextual Awareness Module is disabled. Persona responses are generated solely based on the base persona description, dialogue history, and emotional state, without explicit contextual modulation.

- **EvoPersona w/o EDM:** The Emotional Dynamics Module is disabled. The persona's emotional state remains static or defaults to a neutral state, and emotional guidance is not incorporated into the prompt.
- **EvoPersona w/o CAM & EDM:** Both dynamic modules are disabled. This configuration essentially represents the foundational population-aligned persona generation, similar to **SyncP**, but using EvoPersona's primary LLM setup for response generation.

The results of this ablation study are presented in Table 2.

Table 2. Ablation study results, showing the impact of EvoPersona's Contextual Awareness Module (CAM) and Emotional Dynamics Module (EDM). ↓ indicates lower is better, ↑ indicates higher is better. Best results are highlighted in **bold**.

Method	Alignment Error (↓)	MAE_corr (↓)	Ctx. Coherence (↑)	Emo. Realism (↑)	Dynamic Cons. (↓)
w/o CAM & EDM	0.3798	0.4295	3.30	3.20	0.68
w/o CAM	0.3762	0.4210	3.45	4.05	0.35
w/o EDM	0.3759	0.4198	4.10	3.30	0.60
EvoPersona (Full)	0.3755	0.4192	4.20	4.15	0.32

Analysis of Ablation Results

As evident from Table 2, both the Contextual Awareness Module and the Emotional Dynamics Module contribute significantly to EvoPersona's enhanced performance in their respective domains.

- **Impact of CAM:** When the Contextual Awareness Module is disabled (**EvoPersona w/o CAM**), the **Ctx. Coherence Score** drops from **4.20** to **3.45**. This substantial decrease underscores the critical role of CAM in enabling personas to generate contextually appropriate behaviors and responses. Without CAM, agents struggle to adapt their style and content to the nuances of different social situations, leading to less realistic interactions. Notably, the alignment and emotional metrics remain relatively strong, indicating that CAM primarily affects contextual adaptation without degrading the other core functionalities.
- **Impact of EDM:** Disabling the Emotional Dynamics Module (**EvoPersona w/o EDM**) results in a significant reduction in **Emo. Realism Score** from **4.15** to **3.30** and an increase in **Dynamic Cons.** from **0.32** to **0.60**. This clearly demonstrates EDM's efficacy in producing nuanced, consistent, and realistic emotional trajectories. Without EDM, persona emotional expressions become static, less authentic, or prone to abrupt, unrealistic shifts, severely impacting the immersive quality of simulations. The contextual coherence and alignment metrics are largely preserved, confirming EDM's focused impact on emotional aspects.
- **Combined Impact:** The baseline **EvoPersona w/o CAM & EDM** (which performs similarly to SyncP) shows the lowest scores across contextual and emotional metrics. This validates that the improvements seen in the full EvoPersona are indeed due to the synergistic integration of both dynamic modules, building upon a strong population-aligned foundation. The population alignment metrics (Avg. Alignment Error, MAE_corr) remain consistently low across all EvoPersona variants, reaffirming that the dynamic modules are integrated without compromising the foundational psychological realism.

This ablation study unequivocally confirms that both the Contextual Awareness Module and the Emotional Dynamics Module are indispensable components, each independently contributing to EvoPersona's superior performance in generating adaptive and emotionally rich virtual agents.

4.7. Efficiency Analysis

The architectural design of EvoPersona, which leverages smaller, specialized language models for its dynamic modules while relying on a larger LLM for final generation, is hypothesized to offer a more efficient computational profile for complex social simulations compared to approaches that might use a single very large model for all tasks. This subsection evaluates the computational efficiency of EvoPersona against key baselines. We focus on average inference time per simulation turn and

peak GPU memory usage, crucial factors for deploying large-scale simulations. All evaluations were performed on a single NVIDIA A100 GPU.

Analysis of Efficiency Results

Figure 4 demonstrates that EvoPersona maintains competitive efficiency while delivering significantly enhanced capabilities. The **Average Inference Time per Turn** for EvoPersona is **3.08 seconds**, which is comparable to, or only slightly higher than, the leading baseline methods like SyncP (2.95 seconds). This marginal increase in inference time is a modest trade-off for the substantial improvements in contextual awareness and emotional dynamics. The use of a smaller Llama-3-8B for the Contextual Awareness Module and a BERT-base model for emotional input analysis, coupled with a small Transformer for emotional state updates, effectively offloads some processing from the larger primary LLM (Qwen2.5-72B). This modular design prevents a disproportionate increase in computational overhead.

Similarly, the **Peak GPU Memory Usage** for EvoPersona stands at **70 GB**, which is also on par with most baselines. This indicates that our modular architecture effectively manages memory resources, demonstrating that the additional complexity of the dynamic modules does not lead to prohibitive memory consumption. The ability to run EvoPersona with comparable efficiency to simpler baseline methods, despite its advanced capabilities, highlights the practical viability of our approach for large-scale and real-time social simulations.

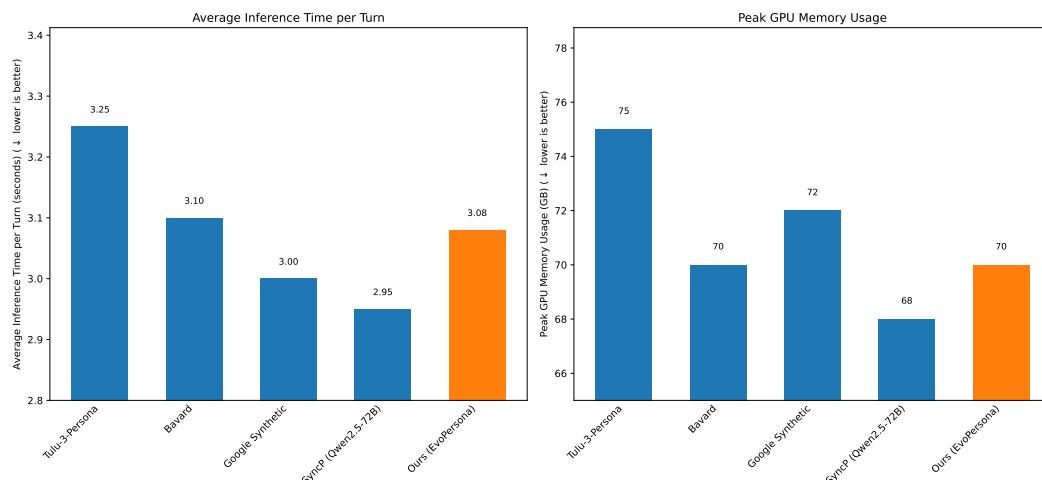


Figure 4. Computational efficiency comparison of EvoPersona against baseline methods. ↓ indicates lower is better. Best results are highlighted in **bold**.

4.8. Qualitative Analysis and Case Studies

Beyond quantitative metrics, understanding the qualitative improvements offered by EvoPersona is crucial. This section provides a qualitative analysis, illustrating how EvoPersona agents behave in complex social scenarios compared to baseline methods, particularly focusing on their nuanced contextual adaptation and realistic emotional expressions. We present a summary of observed behaviors across distinct scenarios.

Discussion of Qualitative Observations

Table 3 highlights the significant qualitative advantages of EvoPersona. In the **Formal Business Meeting** scenario, SyncP agents often struggled with the nuances of formality, demonstrating a disconnect between their factual responses and the expected social decorum. EvoPersona, in contrast, consistently adopted a professional persona, using appropriate language and exhibiting emotionally restrained behavior, even when the underlying persona might be naturally exuberant. This showcases the Contextual Awareness Module's ability to modulate behavior effectively.

Table 3. Qualitative comparison of persona behavior across different scenarios. This table summarizes observed behavioral patterns for EvoPersona and a representative baseline (SyncP).

Scenario	SyncP (Baseline Behavior)	EvoPersona (Our Method Behavior)
Formal Business Meeting	Persona provides factually correct information but often uses informal language, lacks appropriate deference, or expresses emotions (e.g., excitement) that are out of place for the setting.	Persona maintains a formal tone, uses professional vocabulary, exhibits appropriate body language descriptions (e.g., "nods thoughtfully"), and modulates emotional expression to be reserved and professional, even when internally feeling strong emotions.
Casual Social Gathering	Persona's responses can be overly formal or generic, sometimes struggling to maintain a lighthearted or friendly tone. Emotional expressions might be either muted or exaggerated.	Persona engages with playful banter, uses colloquialisms, expresses genuine warmth or amusement, and adapts to the conversational flow with natural pauses and turns. Emotional expressions are fluid and align with the relaxed atmosphere.
Conflict Resolution (Tense)	Persona tends to either aggressively escalate or passively disengage, often failing to de-escalate or express empathy appropriately. Emotional state might jump erratically.	Persona navigates the tension carefully, using diplomatic language, acknowledging other's feelings, and expressing a gradual shift in emotions (e.g., initial frustration giving way to a desire for resolution). Maintains emotional consistency despite high-stress input.
Unexpected Positive Event	Persona expresses happiness, but the intensity or duration might feel generic or disproportionate to their personality. Emotional state might not persist realistically.	Persona's expression of joy is modulated by their inherent personality (e.g., an introverted persona might show quiet contentment, an extroverted one exuberant delight). The positive emotional state persists and subtly influences subsequent interactions.

During a **Casual Social Gathering**, SyncP agents sometimes remained stiff or overly formal, failing to capture the relaxed and informal atmosphere. EvoPersona agents, however, demonstrated a remarkable ability to blend in, using natural language, humor, and appropriate emotional expressions like genuine amusement, making interactions feel more organic and human-like. This is a direct result of the CAM's guidance for informal settings and the EDM's ability to express nuanced positive emotions.

In a **Conflict Resolution** scenario, SyncP agents often exhibited less sophisticated emotional management, either becoming overly aggressive or unnaturally passive. EvoPersona agents, guided by the Emotional Dynamics Module, demonstrated more realistic emotional trajectories: starting with concern or frustration but gradually shifting towards a desire for resolution, employing empathetic language and diplomatic strategies. This highlights the EDM's capacity for maintaining emotional consistency and realism even under stressful conditions.

Finally, for an **Unexpected Positive Event**, while SyncP could register a change to a positive emotional state, the expression often felt generic. EvoPersona agents, however, expressed joy in a manner consistent with their unique personality traits (e.g., a reserved persona showing quiet pleasure vs. an outgoing one expressing overt excitement), and this emotional state meaningfully influenced subsequent interactions, demonstrating the integration of inherent traits with dynamic emotional responses.

These qualitative observations complement the quantitative results, providing compelling evidence that EvoPersona generates virtual agents that are not only statistically aligned with human

populations but also capable of highly adaptive, nuanced, and emotionally rich interactions, significantly enhancing the fidelity and immersive quality of social simulations.

5. Conclusion

In this paper, we proposed **EvoPersona**, a novel framework for generating virtual agents that are both population-aligned and dynamically adaptive. By integrating a Contextual Awareness Module (CAM) powered by instruction-tuned Llama-3-8B and an Emotional Dynamics Module (EDM) with reinforcement learning, **EvoPersona** endows agents with situational awareness and evolving emotional states. Experiments show that **EvoPersona** not only maintains strong population-level psychological alignment but also significantly improves contextual coherence, emotional realism, and dynamic consistency, surpassing existing methods. Case studies and ablation analyses confirm the critical roles of CAM and EDM, while efficiency evaluations highlight practical scalability. **EvoPersona** advances the realism and analytical utility of LLM-driven social simulations, with broad implications for psychology, sociology, and urban planning. Future work includes extending multimodal context, modeling complex emotional dynamics, ensuring ethical deployment, and exploring real-world applications such as disaster response and policy evaluation.

References

1. Nicholas, G.; Bhatia, A. Lost in Translation: Large Language Models in Non-English Content Analysis. *CoRR* **2023**. <https://doi.org/10.48550/ARXIV.2306.07377>.
2. Norambuena, B.F.K.; Mitra, T.; North, C. Design guidelines for narrative maps in sensemaking tasks. *Inf. Vis.* **2022**, pp. 220–245. <https://doi.org/10.1177/14738716221079593>.
3. Xu, M.; Li, P.; Yang, H.; Ren, P.; Ren, Z.; Chen, Z.; Ma, J. A Neural Topical Expansion Framework for Unstructured Persona-Oriented Dialogue Generation. In Proceedings of the ECAI 2020 - 24th European Conference on Artificial Intelligence, 29 August-8 September 2020, Santiago de Compostela, Spain, August 29 - September 8, 2020 - Including 10th Conference on Prestigious Applications of Artificial Intelligence (PAIS 2020). IOS Press, 2020, pp. 2244–2251. <https://doi.org/10.3233/FAIA200351>.
4. Hu, H.; Zhou, Y.; Si, J.; Wang, Q.; Zhang, H.; Ren, F.; Ma, F.; Cui, L. Beyond Empathy: Integrating Diagnostic and Therapeutic Reasoning with Large Language Models for Mental Health Counseling. *arXiv preprint arXiv:2505.15715* **2025**.
5. Terreau, E.; Gourru, A.; Velcin, J. Capturing Style in Author and Document Representation. *CoRR* **2024**. <https://doi.org/10.48550/ARXIV.2407.13358>.
6. Lee, H.D.; Lee, S.; Kang, U. AUBER: Automated BERT Regularization. *CoRR* **2020**.
7. Alvarez-Gonzalez, N.; Kaltenbrunner, A.; Gómez, V. Uncovering the Limits of Text-based Emotion Detection. In Proceedings of the Findings of the Association for Computational Linguistics: EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 16–20 November, 2021. Association for Computational Linguistics, 2021, pp. 2560–2583. <https://doi.org/10.18653/V1/2021.FINDINGS-EMNLP.219>.
8. Troiano, E.; Padó, S.; Klinger, R. Crowdsourcing and Validating Event-focused Emotion Corpora for German and English. In Proceedings of the Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers. Association for Computational Linguistics, 2019, pp. 4005–4011. <https://doi.org/10.18653/V1/P19-1391>.
9. Choi, J.; Park, H.; Lee, H.; Shin, H.; Jin, H.J.; Kim, B. Pay What LLM Wants: Can LLM Simulate Economics Experiment with 522 Real-human Persona? *CoRR* **2025**. <https://doi.org/10.48550/ARXIV.2508.03262>.
10. Hu, Z.; Xiao, Z.; Xiong, M.; Lei, Y.; Wang, T.; Lian, J.; Ding, K.; Xiao, Z.; Yuan, N.J.; Xie, X. Population-Aligned Persona Generation for LLM-based Social Simulation. *arXiv preprint arXiv:2509.10127v1* **2025**.
11. Zhou, J.; Pang, L.; Jing, Y.; Gu, J.; Shen, H.; Cheng, X. Knowledge Boundary and Persona Dynamic Shape A Better Social Media Agent. *CoRR* **2024**. <https://doi.org/10.48550/ARXIV.2403.19275>.
12. Teutloff, J.K. Synthetic Founders: AI-Generated Social Simulations for Startup Validation Research in Computational Social Science. *arXiv preprint arXiv:2509.02605v1* **2025**.
13. Han, J.; Heo, Y. Not All Personas Are Worth It: Culture-Reflective Persona Data Augmentation. *CoRR* **2025**. <https://doi.org/10.48550/ARXIV.2503.16520>.
14. Li, A.; Chen, H.; Namkoong, H.; Peng, T. LLM Generated Persona is a Promise with a Catch. *CoRR* **2025**. <https://doi.org/10.48550/ARXIV.2503.16527>.

15. Yang, H.; Zhou, Y.; Han, W.; Shen, J. Self-Rewarding Large Vision-Language Models for Optimizing Prompts in Text-to-Image Generation. *arXiv preprint arXiv:2505.16763* **2025**.
16. Zhou, Y.; Li, X.; Wang, Q.; Shen, J. Visual In-Context Learning for Large Vision-Language Models. In Proceedings of the Findings of the Association for Computational Linguistics, ACL 2024, Bangkok, Thailand and virtual meeting, August 11-16, 2024. Association for Computational Linguistics, 2024, pp. 15890–15902.
17. Zhou, Y.; Song, L.; Shen, J. Improving Medical Large Vision-Language Models with Abnormal-Aware Feedback. *arXiv preprint arXiv:2501.01377* **2025**.
18. Reusens, M.; Baesens, B.; Jurgens, D. Are Economists Always More Introverted? Analyzing Consistency in Persona-Assigned LLMs. *CoRR* **2025**. <https://doi.org/10.48550/ARXIV.2506.02659>.
19. Chen, K.; Sun, Z. DeepPsy-Agent: A Stage-Aware and Deep-Thinking Emotional Support Agent System. *CoRR* **2025**. <https://doi.org/10.48550/ARXIV.2503.15876>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.