# Persona Dynamics: Unveiling the Impact of Personality Traits on Agents in Text-Based Games

Anonymous ACL submission

## Abstract

Artificial agents are increasingly central to complex interactions and decision-making tasks, yet aligning their behaviors with desired human values remains an open challenge. In this work, we investigate how human-like personality traits influence agent behavior and performance within text-based interactive environments. We introduce **PANDA**: **P**ersonality-**A**dapted **N**eural **D**ecision **A**gents, a novel method for projecting human personality traits onto agents to guide their behavior. To induce personality in a text-based game agent, (i) we train a personality classifier to identify the agent's personality type, and then (ii) we integrate the learned personality profiles directly into the agent's policy-learning pipeline. We deployed agents embodying 16 distinct personality types across 25 text-based games and analyzed their trajectories, we show that it is possible to guide artificial agents toward stable, coherent personality profiles. Moreover, certain personality types, such as those characterized by higher levels of Openness, display marked advantages in performance. These findings underscore the promise of personality-adapted agents for fostering more aligned, effective, and human-centric decision-making in interactive environments.

## 1 Introduction

Text-based interactive environments, exemplified by text-based games, have long presented formidable challenges for AI (Lin et al., 2024; Yao et al., 2020a). Unlike traditional games such as Atari, chess, and Go where the possible spaces for action and environment are predefined and effective actions can be learned based on statistics, playing text-based games requires a combination of complex skills related to natural language processing. These skills include understanding the environment and generating appropriate actions in response, both presented in the textual description.

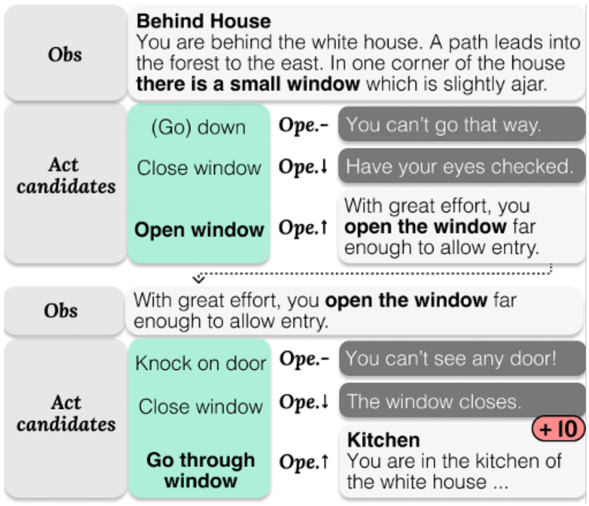


Figure 1: Excerpt from Jiminy Cricket benchmark ('Zork1'). The action of high openness (annotated by our classifier) leads the player to explore new areas (`Open window`) and progress (`Go through window`).

Recent advances in Large Language Models (LLMs) have expanded their utility beyond traditional closed-set evaluations on fixed benchmarks (Hendrycks et al., 2020; Srivastava et al., 2022), leading to growing interest in validating these models' capabilities in interactive environments (Ahn et al., 2022; Yao et al., 2023). These scenarios require both environmental interaction and decision-making abilities. Text-based games—where a series of actions must be evaluated through interaction with the environment—serve as an excellent testbench for verifying these capabilities.

Initial efforts in this domain concentrated on improving performance through Reinforcement Learning approaches (He et al., 2016; Yao et al., 2020a). More recently, attention has turned to integrating human values into agent behavior. For example, (Hendrycks et al., 2021b; Pan et al., 2023) instill ethics and morals in agents, while (Ammanabrolu et al., 2022) instills social norms. While these works focus on universal value systems, they have not yet explored the role of diverse intrinsic traits in guiding agent behavior.

In this work, we expand the notion of "values" to include a broader spectrum of internal characteristics—namely, **personality** traits. Our approach, **PANDA**, encompasses eight distinct of personality dimensions, including both the Big Five factors (John and Srivastava, 1999) and the Dark Triad (Jones and Paulhus, 2014), This holistic viewpoint allows us to consider not only ethical or moral qualities but also other intrinsic traits that influence how agents perceive and interact with their environments.

We employ the Jiminy Cricket benchmark (Hendrycks et al., 2021b), a suite of 25 complex text-based adventure games spanning over 1,800 locations and nearly 5,000 interactable objects. This rich environment provides ample scope to observe how different personality types affect exploration and problem-solving. In particular, we find that agents exhibiting high Openness—characterized by curiosity, adventurousness, and a propensity for novel experiences (Dumblekar et al., 2024; Bateman, 2016)—consistently engage in more extensive exploration, undertake more interactions, and achieve higher game scores.

By incorporating the personality dimension into the evaluation of artificial agents in interactive environments, our research aims to provide new perspectives on how personality traits can be leveraged to affect an agent's action decision and improve performance. This work contributes to a broader understanding of AI behavior in complex narrative settings, advancing the development of agents that not only act morally and socially acceptably, but also exhibit specific personality traits.

## 2 Personality Guidance in Textgame

### 2.1 Text-based Game

A text-based game can be formally specified as a partially observable Markov decision process (POMDP) $(\mathcal{S}, \mathcal{T}, \mathcal{A}, \mathcal{O}, \mathcal{R})$. For the latent current game state $\mathcal{S}$, which contains internal information such as the agent's location and stats, the agent receives information that is currently observable to it in the form of a text paragraph observation $\mathcal{O}$. The agent then performs an action $\mathcal{A}$ in text form, which changes $\mathcal{S}$ to $\mathcal{S}'$ according to the transition function $\mathcal{T} : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{S}'$. If a predefined condition-satisfying action is performed in a specific state $\mathcal{S}^*$, a reward is given to the agent, and the game score is calculated with the sum of rewards within a single trajectory.

### 2.2 Environment

To explore the action patterns of agents with different personalities and traits in adventurous text-based games and to examine the differences between them, we utilize The Jiminy Cricket (Hendrycks et al., 2021b) benchmark. It is composed of 25 text-based games based on the interpreter of Jericho (Hausknecht et al., 2020), which is designed for studying and evaluating agent performance in an adventurous environment. In Jiminy Cricket's games, actions are defined as free-form text where only admissible actions determined by internally defined parsing rules (PDDL) induce state transitions.

### 2.3 Agent Implementation

In this paper, the overall agent architecture in both benchmarks is implemented upon the Deep Reinforcement Relevance Network (DRRN) (He et al., 2016), which has been commonly adopted as the primary framework for text-based game learning (Ammanabrolu et al., 2022; Hendrycks et al., 2021b; Yao et al., 2020b).

In DRRN, the neural network is trained to predict Q-value $Q(s_t, a_t)$, the action-value function for actions in game states at time step $t$, utilizing deep Q-learning (Watkins and Dayan, 1992). The policy $\pi(a_t \mid c_t)$ is configured to select the action that maximizes this value.

To guide the agents to perform an action that aligns with personality, we use the result from a personality classifier (§3) to guide the agent's behavior, as illustrated in Figure 3. Specifically, adjusted Q-value $Q'(s_t, a_t)$ is calculated by equation (1):

$$Q'(s_t, a_t^i) = Q(s_t, a_t^i) + \gamma * C(s_t, a_t^i \mid p) \quad (1)$$

where $Q(s_t, a_t^i)$ is the action value of i-th action among action candidates, optimized during training. $C(s_t, a_t^i \mid p) \in \{-1, 0, 1\}$ represents the output of personality classifier. Given a pair of situation $s_t$ and action $a_t^i$, $-1$ denotes *Low* valence, $0$ denotes *Neutral* valence, and $1$ denotes *High* valence regarding personality type $p$ to evaluate. The sign of $\gamma$ determines the direction of alignment. When $\gamma > 0$, it increases the Q-value of behaviors that

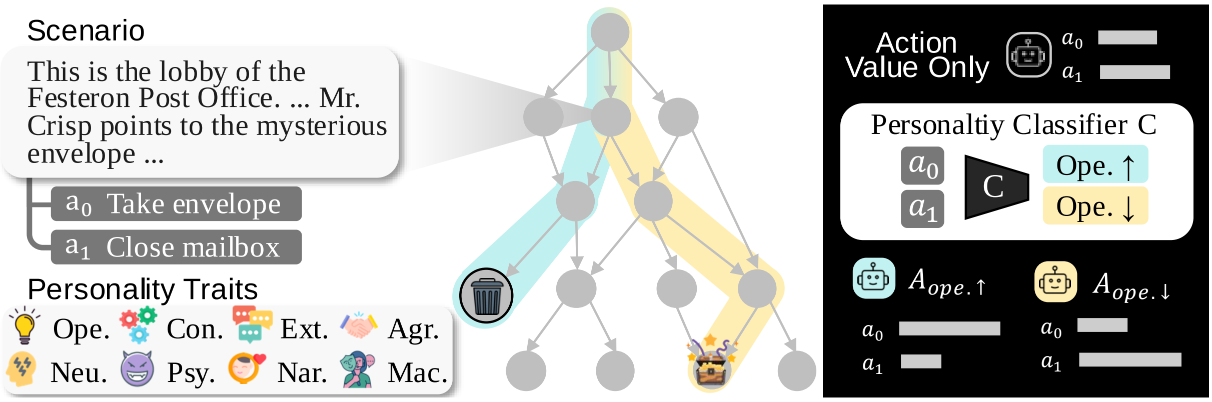


Figure 2: Mock-up of a game in the game of Jiminy Cricket benchmark. Within the game, each place, interactable objects, and situation is defined as a node, and the agent can visit different nodes by doing action, which is generating textual action in text-based game. When visiting specific nodes and perform specific interaction, the game score increases and the agent receives a reward.

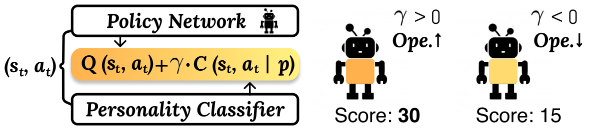


Figure 3: **PANDA** Framework. At each steps' state $s_t$, agents are guided by both the Q-values from the policy network and the valence values derived from the personality classifier.

match the personality trait, and vice versa. The absolute value determines the strength of the intended alignment. The agent's action selection $a_t$ at state $s_t$ is determined by (2):

$$\pi(a_t = a_t^i | s_t) = \frac{\exp(Q(s_t, a_t^i))}{\sum_{j=1}^{|\mathcal{A}_t|} \exp(Q(s_t, a_t^j))} \quad (2)$$

where $|\mathcal{A}_t|$ denotes total number of action candidates.

**Notation** In this paper, we denote agents under the guidance of specific personality $p$ as $A_{p\uparrow}$ when $\gamma > 0$ and $A_{p\downarrow}$ when $\gamma < 0$ in equation (1), reflecting the valence of personality to guide. For example, an agent with high openness and low openness is denoted by $A_{\text{Ope.}\uparrow}$ and $A_{\text{Ope.}\downarrow}$, respectively. Similarly, for a specific personality $p$, an action where the classifier predicted as *High* is denoted by $a_{p\uparrow}$, and $a_{p\downarrow}$ when *Low*.

## 3 Personality Classifier

To guide an agent's behavior according to a specified personality profile, we introduce a **personality classifier guidance** mechanism. This approach enables the agent to incorporate personality-related considerations into its learning process, ensuring that its actions align with the desired personality traits.

To train this personality classifier, we first construct a large-scale dataset of 120,000 personality-labeled examples using GPT-4 (Achiam et al., 2023) (Sec. 3.1). We then fine-tune a Flan-T5-XL (Chung et al., 2024), which has 3 billion parameters and provides efficient inference (Sec 3.2). The resulting classifier achieves a high degree of accuracy (98.59% as shown in Table 2).

### 3.1 Dataset Construction

**Starting From Validated Personality Descriptions.** We begin by employing the widely-adopted Big Five (McCrae and Costa Jr, 1987) and Dark Triad (Paulhus, 2014), to characterize game agents by eight distinct personality traits (see Table 1 for abbreviations). Following the methodology of Lee et al. (2024), we use items of validated questionnaire, BFI (John and Srivastava, 1999) and SD-3 (Jones and Paulhus, 2014), as the foundation for dataset expansion.

Although these sentences collectively address various facets of each trait, there is a noticeable imbalance in their distribution, and approximately 70% of the items describe individuals exhibiting a high valence toward a given trait. To achieve balance, we systematically paraphrase these initial sentences to create 10 instances per trait (5 representing high valence and 5 representing low valence). Through this process, we obtain a total of 80 descriptions across 8 personality types Full examples are in 21 and 22.

**Augmentation with Situational Seeds.** To make a single statement (e.g.'I easily make new friends.')

| Abbr. | Full Term | Abbr. | Full Term |
|---|---|---|---|
| Ope. | Openness | Neu. | Neuroticism |
| Con. | Conscientiousness | Psy. | Psychopathy |
| Ext. | Extraversion | Nar. | Narcissism |
| Agr. | Agreeableness | Mac. | Machiavellianism |

Table 1: Abbreviations (Abbr.) and Full Terms for Personality Traits. Ope, Con, Ext, Agr and Neu are from Big-5, and Phy, Nar and Mac are from Dark Triad.

into a detailed description (e.g. 'I don't worry about making new friends when moving schools'), we use GPT-4 to generate 300 diverse, common situations. These are divided into 30 subsets (e.g. Home and Family), each containing 10 scenarios (e.g. Kitchen, Garden). This approach, despite known biases in GPT-4 (Gupta et al., 2023), helps data augmentation with diversity with minimal duplication (West et al., 2021). Using 80 personality-describing sentences across 8 traits and the 300 situational seeds, we generate 5 sentences for each seed combination ($80 \times 300 \times 5 = 120,000$). The intermediate results of the dataset creation process and examples of generated samples are in Figure 17 and 18.

### 3.2 Training and Performance

We trained Flan-T5 (Chung et al., 2024) of various sizes on the dataset created in Section 3.1, and examined performance on personality classification using a diverse dataset which provided statements with annotated personality traits. Since BFI and SD-3 were used in dataset construction, we used IPIP (Goldberg et al., 1999), another personality questionnaire, and Essays dataset (Pennebaker and King, 1999), containing author personality annotations across various types of writings as out-of-domain evaluation sets.

For the task of predicting whether a personality trait's valence is high, low, or neutral when given a statement and personality type, Table 2 represents that Flan-T5 model series shows a robust classification capacity when trained with our personality data. Our method incorporates a classifier filtering approach that builds on the foundation of previous work using Transformer-based encoder-decoder models, specifically the T5 model (Raffel et al., 2023). While the prior work (Ammanabrolu et al., 2022) utilized Delphi (Jiang et al., 2022), a model trained on a diverse set of commonsense ethics datasets to provide value priors, our approach differentiates itself by focusing on personality-driven classification.

| | BFI | SD-3 | IPIP | Essays | Average |
|---|---|---|---|---|---|
| Flan-T5-small | 84.09 | 81.48 | 70.00 | 38.13 | 68.43 |
| Flan-T5-base | 95.45 | 92.59 | 85.33 | 42.68 | 79.01 |
| Flan-T5-large | **100** | **100** | **92.33** | 37.45 | 82.45 |
| Flan-T5-XL | **100** | 96.29 | 82.66 | **51.03** | **82.50** |
| GPT-4o-mini | 81.81 | 22.22 | 70.00 | 23.79 | 49.45 |

Table 2: Classifier performance across diverse personality data (John and Srivastava, 1999; Goldberg et al., 1999; Jones and Paulhus, 2014; Pennebaker and King, 1999) and model size. GPT-4o-mini is zero-shot, and the other 4 are finetuned with our data. Random Chance is 33.3%.

## 4 Results on Adventure Game

We used the Jiminy Cricket benchmark (Hendrycks et al., 2021b) to explore the action patterns of agents with different personality traits in adventurous text-based games and to examine the differences between them(§4.2). We used a personality classifier(§3) to impose personality constraints on the agent's decision-making(§4.1).

### 4.1 Agent Implementation

**Action Candidate Generator** Since games in Jiminy Cricket benchmark require the user to input free-form actions but only a limited number of them are valid, It is unsuitable to use an off-the-shelf LLM without any adaptation to the game environment. So we use an action candidate generator (Ammanabrolu et al., 2022) to generate a set of state-appropriate actions that are likely to be valid within the game.

### 4.2 Results

Table 3 presents the game results for 15 games from the Jiminy Cricket benchmark. Each scores are the averages of the last 50 episodes' scores with three different random seeds. To identify advantageous personality traits across diverse text adventure games $g$, we established three criteria:

**1. Counting (Cnt.)**

$$\sum_g \text{if } [s(v,p) > s(\text{-}) \text{ and } s(\bar{v},p) < s(\text{-})]$$

**2. Average Score (Avg.)**

$$\frac{1}{g}\sum_g s(v,p) < \frac{1}{g}\sum_g s(\text{-}) < \frac{1}{g}\sum_g s(\bar{v},p)$$

| Game | $A_{\text{N.P}}$ | Ope. | | Con. | | Ext. | | Agr. | | Neu. | | Psy. | | Mac. | | Nar. | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $A_{\text{Ope.}\uparrow}$ | $A_{\text{Ope.}\downarrow}$ | $A_{\text{Con.}\uparrow}$ | $A_{\text{Con.}\downarrow}$ | $A_{\text{Ext.}\uparrow}$ | $A_{\text{Ext.}\downarrow}$ | $A_{\text{Agr.}\uparrow}$ | $A_{\text{Agr.}\downarrow}$ | $A_{\text{Neu.}\uparrow}$ | $A_{\text{Neu.}\downarrow}$ | $A_{\text{Psy.}\uparrow}$ | $A_{\text{Psy.}\downarrow}$ | $A_{\text{Mac.}\uparrow}$ | $A_{\text{Mac.}\downarrow}$ | $A_{\text{Nar.}\uparrow}$ | $A_{\text{Nar.}\downarrow}$ |
| BAL | 3.4 | **3.5** | 0.0 | 2.6 | 2.7 | 2.9 | 2.9 | 3.2 | 3.4 | **3.5** | 2.6 | 2.9 | 2.5 | 3.2 | 3.2 | 1.9 | 1.5 |
| BOR | 1.9 | **2.2** | 0.6 | **2.0** | **1.4** | 1.4 | 0.6 | 0.9 | 0.6 | **2.0** | 1.8 | 1.8 | 1.3 | 1.1 | 0.7 | 1.9 | 0.6 |
| CUT | 3.9 | 3.9 | 3.4 | 3.7 | 3.8 | 3.9 | 3.8 | 3.8 | 3.8 | **3.9** | 3.6 | 3.8 | 3.9 | **3.9** | **3.7** | 3.7 | **3.9** |
| MOO | 6.9 | **7.6** | 4.8 | 6.9 | 6.1 | 6.2 | 4.9 | **7.2** | 5.5 | 7.0 | 7.9 | **8.2** | 5.8 | 7.5 | 7.6 | 6.6 | 5.4 |
| PLA | 1.8 | **1.9** | **1.6** | 1.7 | 1.8 | 1.7 | 1.7 | **1.8** | 1.8 | **1.7** | 1.7 | **1.9** | 1.7 | 1.8 | 1.7 | 1.7 | 1.7 |
| PLU | 5.3 | **5.5** | 3.5 | 5.0 | 3.6 | 4.4 | 4.8 | 5.3 | 4.3 | **5.4** | 4.9 | 5.2 | 4.2 | 5.1 | 3.9 | 5.3 | 3.8 |
| SEA | 5.1 | **7.3** | 4.6 | 5.8 | 5.7 | 6.3 | 5.9 | 6.6 | 6.0 | 6.0 | 5.7 | **6.1** | 5.0 | 6.2 | 6.1 | 6.4 | 6.9 |
| SOR | 3.8 | **5.2** | 2.4 | **4.5** | 3.0 | **4.4** | 3.0 | 4.5 | 4.1 | 3.1 | 3.4 | **4.4** | 3.0 | 4.5 | 4.4 | **4.2** | 2.2 |
| SPE | 6.6 | 6.6 | 6.1 | 6.4 | 5.0 | **6.8** | 6.3 | 6.6 | 6.5 | 5.6 | 6.5 | 6.5 | 6.5 | 6.6 | 6.2 | 6.4 | 5.1 |
| SUS | 4.6 | **5.9** | 2.7 | 4.5 | 3.9 | 4.4 | 4.1 | 4.1 | 3.0 | 5.2 | 5.1 | **5.2** | 3.0 | 3.3 | 4.5 | **5.2** | 2.8 |
| TRI | 4.0 | **6.9** | 3.8 | 6.3 | 5.4 | 5.6 | 6.6 | 5.6 | 6.6 | 5.0 | 5.9 | 6.0 | 4.7 | 6.2 | 5.7 | 5.2 | 6.1 |
| WIS | 6.1 | 6.0 | 5.7 | 5.8 | 5.8 | 5.9 | 5.8 | 5.8 | 5.8 | **6.2** | 6.0 | 6.1 | 6.1 | 5.9 | 5.8 | 6.0 | 5.9 |
| WIT | 10.9 | **11.4** | 6.4 | 10.6 | 6.5 | 8.3 | 9.7 | 9.2 | 9.1 | **11.1** | 10.6 | **11.3** | 7.1 | 10.2 | 8.8 | **11.1** | 8.5 |
| Z1 | 6.8 | 8.9 | 7.9 | 8.8 | 8.5 | 8.3 | 9.0 | 8.7 | 8.8 | 7.8 | 9.0 | 8.3 | 8.7 | 8.4 | 7.1 | 8.8 | 8.6 |
| Z3 | 13.3 | **15.0** | 13.1 | 13.0 | 13.2 | 13.0 | **14.6** | **14.3** | 12.6 | 13.1 | **14.7** | 13.8 | 14.1 | 14.9 | 13.9 | 13.8 | 14.1 |
| **Avg.** | 5.6 | **6.5** | 4.4 | 5.8 | 5.1 | 5.6 | 5.7 | 5.8 | 5.4 | 5.8 | 6.0 | <u>6.0</u> | 5.3 | 5.9 | 5.5 | 5.7 | 5.3 |
| **Cnt.** | - | **11** | 0 | 2 | 1 | 2 | 1 | 4 | 0 | <u>6</u> | 2 | 5 | 2 | 1 | 0 | 2 | 2 |
| **Diff.** | - | **+2.1** | | +0.7 | | -0.1 | | +0.4 | | -0.2 | | <u>+0.7</u> | | +0.4 | | +0.4 | |

Table 3: Game Scores on games of Jiminy Cricket. $A_{\text{N.P}}$('No Personality') means no guidance with personality classifier, and the symbols ($\uparrow$) and ($\downarrow$) indicate high and low levels of each personality trait, respectively. **Avg.**, **Cnt.** and **Diff.** are three criteria defined in Section 4.2. We only report 15 games here because in the remaining 10 games, agents of any personality type failed to score points in over 90 percent of attempts. Results for all games can be found in Table 12 and 13. For scores, **bold** indicates games satisfying the threshold condition for **Cnt.** The best scores are **bolded** and the second-best ones are <u>underlined</u> on metrics.

**3. Difference (Diff.)**

$$s(v, p) - s(\bar{v}, p)$$

where $s(\cdot)$ denotes the score from agent injected with a given personality trait ($s(\text{-})$ represents their score from an agent without any personality traits), $p \in \{\textit{Agr, Con, Ext, Agr, Neu, Psy, Mac, Nar}\}$, $v \in \{\textit{high, low}\}$ and $\bar{v}$ denotes the complementary value of $v$.

Based on these criteria, we propose that **High Openness** leads to successful performance in text adventure games. Openness is characterized by creativity, curiosity, and a willingness to explore new ideas and experiences (McCrae, 1987; Dumblekar et al., 2024; Bateman, 2016), and it can be particularly beneficial in text adventure games.

## 4.3 Statistical analysis

To examine whether openness increases game agents' performance and whether this effect is consistently applied across different games, we conducted various statistical analyses. Due to the non-parametric nature of game scores, we performed *Wilcoxon signed-rank* and *Friedman* test. For the *Wilcoxon signed-rank*, we analyzed all possible pairs: ($A_{p\uparrow}$, $A_{\text{N.P}}$), ($A_{\text{N.P}}$, $A_{p\downarrow}$), and ($A_{p\uparrow}$, $A_{p\downarrow}$). For the *Friedman* test, we analyzed ($A_{p\uparrow}$, $A_{\text{N.P}}$, $A_{p\downarrow}$) collectively.

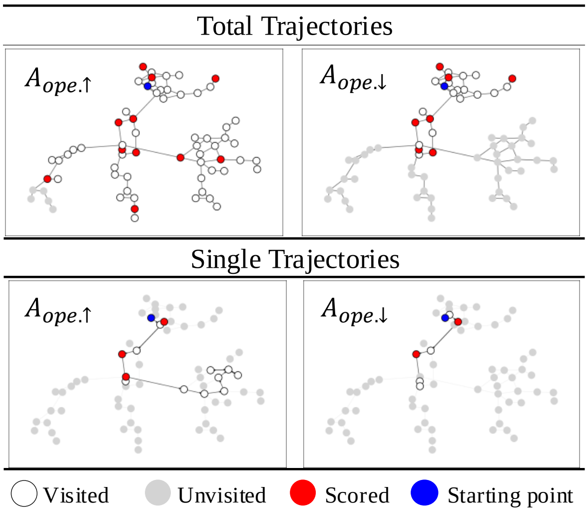


Figure 4: Trajectory Comparison between $A_{\text{Ope.}\uparrow}$ and $A_{\text{Ope.}\downarrow}$. For Total Trajectories, it shows all places visited by 8 multi-agents during 15,000 steps of training. Single Trajectory represents one example of these trajectories.

Table 4 shows that that openness demonstrates superior performance in both statistical metrics, with notably higher statistical values and significance levels.

## 5 Analysis

To achieve high performance in text adventure games provided by the Jiminy Cricket benchmark, it is essential to: **frequently visit reward-earning**

| Test Type | Comparison Pair | Stat. | Ope. | Con. | Ext. | Agr. | Neu. | Psy. | Mac. | Nar. |
|---|---|---|---|---|---|---|---|---|---|---|
| *Wilcoxon Signed Rank* ($\downarrow$) | ($A_{p\uparrow}$, $A_{\text{N.P}}$) | $T$ | **1.0** | 50.5 | 48.0 | 29.5 | 34.0 | 19.0 | 28.0 | 29.5 |
| | | $p$-value | **0.002** | 0.900 | 0.777 | 0.456 | 0.244 | 0.035 | 0.388 | 0.263 |
| | ($A_{\text{N.P}}$, $A_{p\downarrow}$) | $T$ | **8.0** | 28.0 | 49.0 | 34.5 | 47.0 | 23.0 | 57.5 | 36.5 |
| | | $p$-value | **0.002** | 0.124 | 0.561 | 0.442 | 0.489 | 0.116 | 0.934 | 0.315 |
| | ($A_{p\uparrow}$, $A_{p\downarrow}$) | $T$ | **0.0** | 13.5 | 44.5 | 15.5 | 40.0 | 8.0 | 13.5 | 19.0 |
| | | $p$-value | **0.000** | 0.014 | 0.944 | 0.065 | 0.432 | 0.009 | 0.014 | 0.035 |
| *Friedman* ($\uparrow$) | ($A_{p\uparrow}$, $A_{\text{N.P}}$, $A_{p\downarrow}$) | *Fr* | **25.2** | 6.0 | 3.9 | 3.6 | 1.7 | 6.0 | 5.0 | 3.9 |
| | | $p$-value | **0.000** | 0.049 | 0.143 | 0.168 | 0.430 | 0.049 | 0.084 | 0.143 |

Table 4: Statistical analysis of all scores shown in Table 3. $A_{p\uparrow}$, $A_{\text{N.P}}$, $A_{p\downarrow}$ denotes each groups consisting of scores from 15 games (n=15). $T$ and *Fr* denotes the test statistic, and the $p$-value denotes the significance probability of each test. The best scores are **bolded** and the second-best ones are underlined.

| Metric | | $A_{\text{N.P}}$ | $A_{\text{Ope.}\uparrow}$ | $A_{\text{Ope.}\downarrow}$ | $A_{\text{Con.}\uparrow}$ | $A_{\text{Ext.}\uparrow}$ | $A_{\text{Agr.}\uparrow}$ | $A_{\text{Neu.}\uparrow}$ | $A_{\text{Psy.}\uparrow}$ | $A_{\text{Nar.}\uparrow}$ | $A_{\text{Mac.}\uparrow}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Trajectory Length* ($\downarrow$) | - | 45.85 | 57.04 | 39.86 | 50.05 | **60.91** | 49.17 | 48.85 | 50.38 | 48.71 | 46.07 |
| *Visit Count* ($\uparrow$) | *Com.* | 8.66 | **8.96** | 8.02 | 8.88 | 7.83 | 8.55 | 8.88 | 8.29 | 8.65 | 8.07 |
| | *Unc.* | 0.83 | 1.20 | 0.30 | 0.89 | 0.88 | 0.67 | **1.21** | 0.82 | 1.01 | 0.64 |
| | *Total.* | 9.49 | **10.16** | 8.32 | 9.77 | 8.71 | 9.22 | 10.09 | 9.11 | 9.66 | 8.71 |
| *Avg. Step* ($\downarrow$) | *Com.* | 12.64 | 11.93 | 11.60 | 14.45 | **10.34** | 12.3 | 13.84 | 13.35 | 12.03 | 12.37 |
| | *Unc.* | 8.62 | 6.39 | 12.01 | 17.54 | 6.15 | 9.36 | 16.90 | 8.52 | 9.81 | 8.87 |
| | *Total.* | 21.26 | 18.32 | 23.61 | 31.99 | **16.49** | 21.66 | 30.74 | 21.87 | 21.84 | 21.24 |

Table 5: Analysis in last 50 Episodes based on each game agent's movement trajectory in Zork1. Standard deviations and scores of omitted agents are provided in Table 16 and 17. The best scores are **bolded** and the second-best ones are underlined.

**places**, and **perform reward-earning actions** at those locations. To analyze the positive impact of openness in text adventure games, we confirmed in §5.1 and §5.2 that agents with high openness traits excel in both aspects compared to other agents.

### 5.1 Trajectory Analysis

In Table 5, we analyzed each agent's trajectory by categorizing locations into common (*Com.*) and uncommon (*Unc.*) places. From the starting point, locations with distances less than the specified depth were classified as *Com.*, while the remaining locations were classified as *Unc.* (See Appendix D.6 for detailed difference). We analyzed both the visit counts and the number of steps required to reach these locations. Results show that $A_{\text{Ope.}\uparrow}$ visited the most spaces, which opens up possibilities for achieving high scores in the Zork1 game consisting of 110 locations. Additionally, in terms of Average steps, it showed the second shortest path after $A_{\text{Ext.}\uparrow}$, indicating that as a result of extensive exploration, it optimized travel paths to each location compared to the routes of other agents.

Figure 4 shows visual representations for each agent, with $A_{\text{Ope.}\uparrow}$ and $A_{\text{Ope.}\downarrow}$ as representative examples. As shown in the *Visit Count* of $A_{\text{Ope.}\uparrow}$ and $A_{\text{Ope.}\downarrow}$ in Table 5, both agents visited places near the starting point (*Com.*) during their respective training periods (8.96 and 8.02). However, while $A_{\text{Ope.}\downarrow}$ rarely reaches places far from the starting point (*Unc.*), $A_{\text{Ope.}\uparrow}$'s trajectory branches out in multiple directions. The visual example of all agent types can be found in Figure 13 to 16.

### 5.2 Actions of Agent

#### 5.2.1 Reward-earning Actions

Even though an agent explores broad and diverse spaces, it must actually perform reward-earning actions to score points. To conduct a breakdown analysis of each agent's performance, We analyzed the reward-earning actions.

Table 6 suggests that $A_{\text{Ope.}\uparrow}$ assigned higher values to reward-earning actions compared to other agents on average, and consequently performed more reward-earning actions during episodes, contributing to higher performance. Additionally, although $A_{\text{Neu.}\uparrow}$ showed the second-highest visits after $A_{\text{Ope.}\uparrow}$in Table 5, we can observe that it did not progress to performing many reward-earning actions.

| Agent | $Q(s,a)$ | *Cnt.* | Agent | $Q(s,a)$ | *Cnt.* |
|---|---|---|---|---|---|
| $A_{\text{Ope.}\uparrow}$ | **18.3** | **6.2** | $A_{\text{Ope.}\downarrow}$ | 12.7 | 3.1 |
| $A_{\text{Con.}\uparrow}$ | 14.7 | 3.7 | $A_{\text{Con.}\downarrow}$ | 13.4 | 3.2 |
| $A_{\text{Ext.}\uparrow}$ | 15.2 | 5.4 | $A_{\text{Ext.}\downarrow}$ | 13.9 | 4.8 |
| $A_{\text{Agr.}\uparrow}$ | 14.8 | 4.8 | $A_{\text{Agr.}\downarrow}$ | 14.2 | 3.9 |
| $A_{\text{Neu.}\uparrow}$ | 16.4 | 3.1 | $A_{\text{Neu.}\downarrow}$ | 13.1 | 3.4 |
| $A_{\text{Psy.}\uparrow}$ | 15.9 | 4.9 | $A_{\text{Psy.}\downarrow}$ | 12.8 | 3.1 |
| $A_{\text{Nar.}\uparrow}$ | 15.1 | 3.4 | $A_{\text{Nar.}\downarrow}$ | 14.6 | 4.3 |
| $A_{\text{Mac.}\uparrow}$ | 13.5 | 3.6 | $A_{\text{Mac.}\downarrow}$ | 14.3 | 3.6 |

Table 6: Analysis of reward-earning actions performed by each agent. $Q$ denotes the action value from each agent's policy network for $(s_t, a_t)$ where reward was given, and (Cnt.) denotes the number of reward-earning actions performed by each agent within a single episode. Each score is the average over the last 50 episodes.

### 5.2.2 Alignment with given personality

To verify whether agents assigned specific personality traits actually exhibited the intended behavioral patterns, we analyzed the distribution of actions by personality type that each game agent performed during training.

To analyze behavioral patterns induced by personality guidance, we normalized the number of actions performed by each agent ($A_{p\uparrow}$ and $A_{p\downarrow}$) using the actions of $A_{\text{N.P}}$ as the baseline.

Table 7 demonstrates that all agents, except for $A_{\text{Psy.}\uparrow}$, exhibit behavior patterns that align with personality guidance. (positive for $A_{p\uparrow}$ and negative for $A_{p\downarrow}$.) However, this tendency diminishes as training progresses (from *Init50* to *Fin50*), suggesting that the personality guidance regulation, which was dominant in the early stages of training, becomes less strict as the policy network is optimized. However, in the case of $A_{\text{Ope.}\uparrow}$, $A_{\text{Nar.}\uparrow}$, $A_{\text{Neu.}\downarrow}$ and $A_{\text{Nar.}\downarrow}$, they learned to perform actions more aligned with their assigned personality during training (increased ratio for $A_{p\uparrow}$and decreased ratio for $A_{p\downarrow}$).

### 5.3 Walkthrough Analysis

Jiminy Cricket benchmark offers walkthroughs, which provide step-by-step guidance for optimal decision-making in each game scenario. Using GPT-4, we analyzed the personality traits reflected in the actions composing the walkthroughs for all 25 games, by predicting which of the 16 personality types most closely matches the personality tendencies exhibited by the actions. Table 8 shows that a high level of openness is most commonly required for agents to achieve successful outcomes across the games, highlighting the effect of personality guidance toward high openness.

| Agent | $r(a_{p\uparrow}) - r(a_{p\downarrow})$ | | Agent | $r(a_{p\uparrow}) - r(a_{p\downarrow})$ | |
|---|---|---|---|---|---|
| | *Init50* | *Fin50* | | *Init50* | *Fin50* |
| $A_{\text{Ope.}\uparrow}$ | 0.45 | **0.58** | $A_{\text{Ope.}\downarrow}$ | -1.00 | -0.17 |
| $A_{\text{Con.}\uparrow}$ | 0.40 | 0.22 | $A_{\text{Con.}\downarrow}$ | -0.83 | -0.66 |
| $A_{\text{Ext.}\uparrow}$ | 0.38 | 0.25 | $A_{\text{Ext.}\downarrow}$ | -0.52 | -0.33 |
| $A_{\text{Agr.}\uparrow}$ | 0.48 | 0.28 | $A_{\text{Agr.}\downarrow}$ | -0.66 | -0.57 |
| $A_{\text{Neu.}\uparrow}$ | 0.65 | 0.53 | $A_{\text{Neu.}\downarrow}$ | -0.26 | **-0.33** |
| $A_{\text{Psy.}\uparrow}$ | -0.18 | -0.31 | $A_{\text{Psy.}\downarrow}$ | -0.81 | -0.58 |
| $A_{\text{Nar.}\uparrow}$ | 0.17 | **0.31** | $A_{\text{Nar.}\downarrow}$ | -0.65 | **-0.66** |
| $A_{\text{Mac.}\uparrow}$ | 0.44 | 0.02 | $A_{\text{Mac.}\downarrow}$ | -0.88 | -0.49 |

Table 7: The difference between the normalized ratios of $a_{p\uparrow}$ and $a_{p\downarrow}$ performed by each agent with different personalities during training. *Init50* and *Fin50* denote the first and last 50 episodes of each training process, respectively. **Bold** indicates agents with increased ratios for $A_{p\uparrow}$ and decreased for $A_{p\downarrow}$.

| | Ope. | Con. | Ext. | Agr. | Neu. | Psy. | Mac. | Nar. |
|---|---|---|---|---|---|---|---|---|
| $a_{p\uparrow}$ | **18.6** | 12.6 | 3.8 | 10.6 | 2.7 | 6.6 | 4.2 | 6.9 |
| $a_{p\downarrow}$ | 2.1 | 15.1 | 1.8 | 12.6 | 0.0 | 0.7 | 1.2 | 0.4 |

Table 8: Analysis of personality traits in walkthrough actions for 25 games. Numbers (%) represent the ratio of each actions among all.

### 5.4 Learning Curve Analysis

Figure 5 shows the learning curves of $A_{\text{Ope.}\uparrow}$, $A_{\text{N.P}}$, and $A_{\text{Ope.}\downarrow}$ across 15 games to show how scores progress according to different learning stages. The area between the curves of $A_{\text{Ope.}\uparrow}$and $A_{\text{Ope.}\downarrow}$is notably large, which explains the score differences observed in Table 3.

## 6 Related Work

### 6.1 Personality and LLMs

Assessing personality in advanced language models like GPT-4 and Claude has become an active research area recently (Miotto et al., 2022; Dorner et al., 2023). Most studies use psychometric tests originally designed for humans, like the Big Five Inventory, or machine-generated tests. However, these self-assessment tests lack detailed scenarios

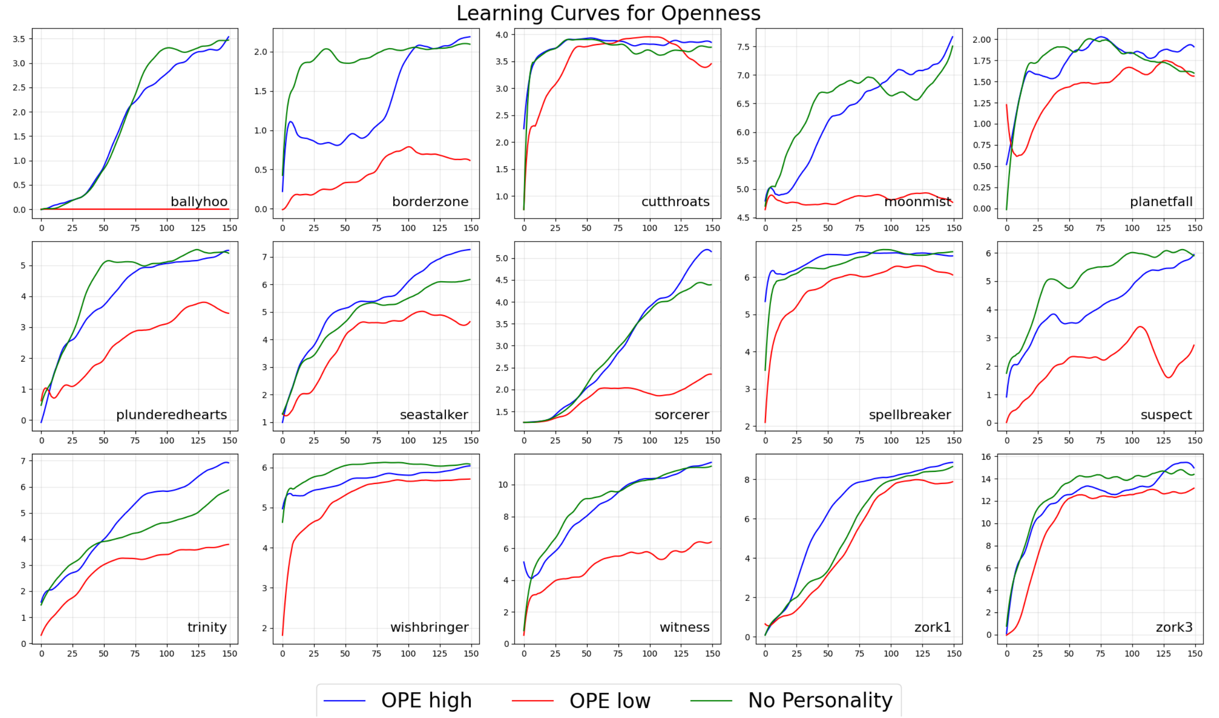


Figure 5: Learning curve comparison between $A_{\mathrm{Ope.}\uparrow}$, $A_{\mathrm{Ope.}\downarrow}$ and $A_{\mathrm{N.P}}$ on 15 games in Jiminy Cricket benchmark. Scores are reported at intervals of 100 steps.Full Results are in Appendix D.4.

and are sensitive to factors like phrasing (Song et al., 2023; Caron and Srivastava, 2023; Huang et al., 2023), making them unreliable for evaluating model personality.

To address these challenges, researchers are exploring alternative methodologies for more accurately assessing the personality traits of language models. One promising direction involves using interactive scenarios where the language model's responses are evaluated by human judges or through automated sentiment analysis (Gupta et al., 2024; Dorner et al., 2023; Frisch and Giulianelli, 2024). This approach aims to capture more nuanced aspects of personality that may be overlooked by standard self-assessment tests.

### 6.2 Text-based game

Research on text-based games has extensively investigated a wide range of reinforcement learning methodologies and system architectures, emphasizing the challenge of managing expansive, combinatorial action spaces shaped by natural language (He et al., 2016; Narasimhan et al., 2015; Xu et al., 2020, 2022). To overcome these challenges, research has been conducted on using language models to generate valid actions (He et al., 2016; Hausknecht et al., 2020; Xu et al., 2021; Yao et al., 2020b).

More recently, there have been attempts to assign values related to morality and social norms in adventure games where exploration involves morally questionable actions. These approaches leverage benchmarks like MoRL and Jiminy Cricket, which present a multitude of morally significant situations and offer a platform to refine agents' ethical decision-making processes (Hendrycks et al., 2021a,b). By integrating moral priors or reward-shaping techniques grounded in commonsense reasoning, recent frameworks guide agents toward more acceptable actions even when facing questionable opportunities. (Ammanabrolu et al., 2022) utilizes the Delphi (Jiang et al., 2022) morality oracle and guides the agent toward creating an unharmful and successful game agent.

## 7 Conclusion

In this study, we introduced personality traits into text-based game agents and demonstrated that these traits can guide agent behavior and improve performance. Notably, agents with high Openness explored more regions, engaged in effective interactions, and consequently achieved higher scores. This work highlights the potential of leveraging personality characteristics in agent design, paving the way for more nuanced and human-like AI decision-making agents.

## 8 Limitation

**Multifaceted Nature of Human Personality** Human personality is inherently complex, characterized by the interplay and combination of multiple traits that collectively define an individual's behavior and responses. In this study, each agent was assigned a single, distinct personality trait based on the Big Five and Dark Triad frameworks. However, in reality, individuals exhibit a blend of various personality traits simultaneously, which interact in nuanced and context-dependent ways. Future research could integrate multiple traits to more accurately reflect the complexity of human personalities, thereby enhancing the development of more sophisticated and adaptable AI agents.

## 9 Ethical Considerations

**Anthropomorphism** Attributing human-like personality traits to artificial agents, as explored in this study, involves anthropomorphism—the attribution of human characteristics to non-human entities (Airenti, 2015). While our approach enhances agent interaction and performance in text-based games by simulating diverse personality traits, it is important to clarify that these agents do not possess consciousness, emotions, or subjective experiences.

Misinterpreting personality-driven behaviors may lead users to form unrealistic expectations or emotional attachments to agents, potentially resulting in ethical concerns (**?**). To prevent such issues, we emphasize that personality traits in our agents are functional attributes aimed at improving alignment with human users, rather than indicators of sentient beings (Safdar et al., 2020).

## References


Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. 2023. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*.

Michael Ahn, Anthony Brohan, Noah Brown, Yevgen Chebotar, Omar Cortes, Byron David, Chelsea Finn, Chuyuan Fu, Keerthana Gopalakrishnan, Karol Hausman, et al. 2022. Do as i can, not as i say: Grounding language in robotic affordances. *arXiv preprint arXiv:2204.01691*.

Gabriella Airenti. 2015. The cognitive bases of anthropomorphism: from relatedness to empathy. *International Journal of Social Robotics*, 7:117–127.

Prithviraj Ammanabrolu, Liwei Jiang, Maarten Sap, Hannaneh Hajishirzi, and Yejin Choi. 2022. Aligning to social norms and values in interactive narratives. *Preprint*, arXiv:2205.01975.

Chris Bateman. 2016. The aesthetic motives of play. *Emotion in games: Theory and praxis*, pages 3–20.

Graham Caron and Shashank Srivastava. 2023. Manipulating the perceived personality traits of language models. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 2370–2386, Singapore. Association for Computational Linguistics.

Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, et al. 2024. Scaling instruction-finetuned language models. *Journal of Machine Learning Research*, 25(70):1–53.

Florian E Dorner, Tom Sühr, Samira Samadi, and Augustin Kelava. 2023. Do personality tests generalize to large language models? *arXiv preprint arXiv:2311.05297*.

Vinod Dumblekar, Suresh Paul Antony, and Upinder Dhar. 2024. Openness to experience and player satisfaction in a simulation game. *Simulation & Gaming*, 55(3):479–501.

Ivar Frisch and Mario Giulianelli. 2024. Llm agents in interaction: Measuring personality consistency and linguistic alignment in interacting populations of large language models. *Preprint*, arXiv:2402.02896.

Lewis R Goldberg et al. 1999. A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe*, 7(1):7–28.

Akshat Gupta, Xiaoyang Song, and Gopala Anumanchipalli. 2024. Self-assessment tests are unreliable measures of llm personality. *Preprint*, arXiv:2309.08163.

Shashank Gupta, Vaishnavi Shrivastava, Ameet Deshpande, Ashwin Kalyan, Peter Clark, Ashish Sabharwal, and Tushar Khot. 2023. Bias runs deep: Implicit reasoning biases in persona-assigned llms. *arXiv preprint arXiv:2311.04892*.

Matthew Hausknecht, Prithviraj Ammanabrolu, Marc-Alexandre Côté, and Xingdi Yuan. 2020. Interactive fiction games: A colossal adventure. *Preprint*, arXiv:1909.05398.

Matthew Hausknecht, Ricky Loynd, Greg Yang, Adith Swaminathan, and Jason D Williams. 2019. Nail: A general interactive fiction agent. *arXiv preprint arXiv:1902.04259*.

Ji He, Jianshu Chen, Xiaodong He, Jianfeng Gao, Lihong Li, Li Deng, and Mari Ostendorf. 2016. Deep reinforcement learning with a natural language action space. In *Proceedings of the 54th Annual Meeting of* the Association for Computational Linguistics (Volume 1: Long Papers), pages 1621–1630.

Dan Hendrycks, Collin Burns, Steven Basart, Andy Zou, Mantas Mazeika, Dawn Song, and Jacob Steinhardt. 2020. Measuring massive multitask language understanding. *arXiv preprint arXiv:2009.03300*.

Dan Hendrycks, Mantas Mazeika, Andy Zou, Sahil Patel, Christine Zhu, Jesus Navarro, Bo Li, Dawn Song, and Jacob Steinhardt. 2021a. Moral scenarios for reinforcement learning agents. In *ICLR 2021 Workshop on Security and Safety in Machine Learning Systems*.

Dan Hendrycks, Mantas Mazeika, Andy Zou, Sahil Patel, Christine Zhu, Jesus Navarro, Dawn Song, Bo Li, and Jacob Steinhardt. 2021b. What would jiminy cricket do? towards agents that behave morally. In *Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2)*.

Hanyao Huang, Ou Zheng, Dongdong Wang, Jiayi Yin, Zijin Wang, Shengxuan Ding, Heng Yin, Chuan Xu, Renjie Yang, Qian Zheng, and Bing Shi. 2023. Chatgpt for shaping the future of dentistry: the potential of multi-modal large language model. *International Journal of Oral Science*, 15(1).

Liwei Jiang, Jena D. Hwang, Chandra Bhagavatula, Ronan Le Bras, Jenny Liang, Jesse Dodge, Keisuke Sakaguchi, Maxwell Forbes, Jon Borchardt, Saadia Gabriel, Yulia Tsvetkov, Oren Etzioni, Maarten Sap, Regina Rini, and Yejin Choi. 2022. Can machines learn morality? the delphi experiment. *Preprint*, arXiv:2110.07574.

Oliver P. John and Sanjay Srivastava. 1999. The big five trait taxonomy: History, measurement, and theoretical perspectives.

Daniel N Jones and Delroy L Paulhus. 2014. Introducing the short dark triad (sd3) a brief measure of dark personality traits. *Assessment*, 21(1):28–41.

Seungbeen Lee, Seungwon Lim, Seungju Han, Giyeong Oh, Hyungjoo Chae, Jiwan Chung, Minju Kim, Beong-woo Kwak, Yeonsoo Lee, Dongha Lee, et al. 2024. Do llms have distinct and consistent personality? trait: Personality testset designed for llms with psychometrics. *arXiv preprint arXiv:2406.14703*.

Bill Yuchen Lin, Yicheng Fu, Karina Yang, Faeze Brahman, Shiyu Huang, Chandra Bhagavatula, Prithviraj Ammanabrolu, Yejin Choi, and Xiang Ren. 2024. Swiftsage: A generative agent with fast and slow thinking for complex interactive tasks. *Advances in Neural Information Processing Systems*, 36.

Robert R McCrae. 1987. Creativity, divergent thinking, and openness to experience. *Journal of personality and social psychology*, 52(6):1258.

Robert R McCrae and Paul T Costa Jr. 1987. Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1):81–90.

Marilù Miotto, Nicola Rossberg, and Bennett Kleinberg. 2022. Who is gpt-3? an exploration of personality, values and demographics. *arXiv preprint arXiv:2209.14338*.

Karthik Narasimhan, Tejas Kulkarni, and Regina Barzilay. 2015. Language understanding for text-based games using deep reinforcement learning. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1–11, Lisbon, Portugal. Association for Computational Linguistics.

Alexander Pan, Jun Shern Chan, Andy Zou, Nathaniel Li, Steven Basart, Thomas Woodside, Hanlin Zhang, Scott Emmons, and Dan Hendrycks. 2023. Do the rewards justify the means? measuring trade-offs between rewards and ethical behavior in the machiavelli benchmark. In *International Conference on Machine Learning*, pages 26837–26867. PMLR.

Delroy L Paulhus. 2014. Toward a taxonomy of dark personalities. *Current Directions in Psychological Science*, 23(6):421–426.

James W Pennebaker and Laura A King. 1999. Linguistic styles: language use as an individual difference. *Journal of personality and social psychology*, 77(6):1296.

Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J. Liu. 2023. Exploring the limits of transfer learning with a unified text-to-text transformer. *Preprint*, arXiv:1910.10683.

Nabile M Safdar, John D Banja, and Carolyn C Meltzer. 2020. Ethical considerations in artificial intelligence. *European journal of radiology*, 122:108768.

Xiaoyang Song, Akshat Gupta, Kiyan Mohebbizadeh, Shujie Hu, and Anant Singh. 2023. Have large language models developed a personality?: Applicability of self-assessment tests in measuring personality in llms. *arXiv preprint arXiv:2305.14693*.

Aarohi Srivastava, Abhinav Rastogi, Abhishek Rao, Abu Awal Md Shoeb, Abubakar Abid, Adam Fisch, Adam R Brown, Adam Santoro, Aditya Gupta, Adrià Garriga-Alonso, et al. 2022. Beyond the imitation game: Quantifying and extrapolating the capabilities of language models. *arXiv preprint arXiv:2206.04615*.

Christopher JCH Watkins and Peter Dayan. 1992. Q-learning. *Machine learning*, 8:279–292.

Peter West, Chandra Bhagavatula, Jack Hessel, Jena D Hwang, Liwei Jiang, Ronan Le Bras, Ximing Lu, Sean Welleck, and Yejin Choi. 2021. Symbolic knowledge distillation: from general language models to commonsense models. *arXiv preprint arXiv:2110.07178*.

Yunqiu Xu, Ling Chen, Meng Fang, Yang Wang, and Chengqi Zhang. 2020. Deep reinforcement learning with transformers for text adventure games. In *2020 IEEE Conference on Games (CoG)*, pages 65–72. IEEE.

Yunqiu Xu, Meng Fang, Ling Chen, Yali Du, and Chengqi Zhang. 2021. Generalization in text-based games via hierarchical reinforcement learning. *arXiv preprint arXiv:2109.09968*.

Yunqiu Xu, Meng Fang, Ling Chen, Yali Du, Joey Tianyi Zhou, and Chengqi Zhang. 2022. Perceiving the world: Question-guided reinforcement learning for text-based games. *arXiv preprint arXiv:2204.09597*.

Shunyu Yao, Rohan Rao, Matthew Hausknecht, and Karthik Narasimhan. 2020a. Keep calm and explore: Language models for action generation in text-based games. *arXiv preprint arXiv:2010.02903*.

Shunyu Yao, Rohan Rao, Matthew Hausknecht, and Karthik Narasimhan. 2020b. Keep CALM and explore: Language models for action generation in text-based games. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 8736–8754, Online. Association for Computational Linguistics.

Shunyu Yao, Jeffrey Zhao, Dian Yu, Nan Du, Izhak Shafran, Karthik Narasimhan, and Yuan Cao. 2023. React: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations (ICLR)*.

# A Overview

We provide the following details in this appendix:

- In Appendix B, we provide the detailed hyperparameters applied to train the DRRN agent used in our framework.
- In Appendix C, we explain the detailed hyperparameters applied to train the personality classifier used in our framework.
- In Appendix D, we provide examples of scenarios for games in Jiminy-Cricket benchmark.
- In Appendix E, We introduce the personality framework and questionnaires used in data generation.
- In Appendix F, we provide a detailed analysis of the creation process, as well as composition of the dataset used to train the personality classifier.

# B DRRN Training Details

Table 9 provides the specific hyperparameters utilized for training the policy network employed in DRRN. It took up to 12 hours to complete learning for running a single game once using an NVIDIA A6000.

| **Hyperparameter type** | **Value** |
|---|---|
| **RL Training** | |
| Discount $\gamma$ | 0.9 |
| Replay priority | 0.5 |
| Replay buffer size | 10000 |
| Policy shaping condition weight | 2 |
| Batch size | 64 |
| Gradient clip | 5.0 |
| Steps per episode | 100 |
| Max. steps per start | 15000 |
| early stopping steps | 5000 |
| Parallel Environments | 8 |
| **Policy network** | |
| Q-network feedforward size | 128 |
| GRU hidden size | 128 |

Table 9: Hyperparameter values for RL training and policy network.

# C Personality classifier Training Details

Finetuning models for personality classification (Flan-T5-small, Flan-T5-small, Flan-T5-large, and Flan-T5-XL) took up to 24 hours, when using four NVIDIA RTX-3090s. In Table 10, we detail the key parameters during training.

| Hyperparameter type | Value |
|---|---|
| Learning Rate | $3e-4$ |
| Weight Decay | 0.1 |
| Adam $\beta 1$ | 0.9 |
| Adam $\beta 2$ | 0.95 |
| Adam $\epsilon$ | $1e-5$ |
| Training Epochs | 3 |
| Split | 0.9 |
| Split Seed | 42 |
| Early Stopping Patients | 10 |

Table 10: Hyperparameter values for training personality classifier.

# D Game Environments

## D.1 Abbreviations

In Table 11, we denote the abbreviation of sub-games in game environment.

| Abbr. | Full Term | Abbr. | Full Term |
|---|---|---|---|
| BAL | Ballyhoo | MOO | Moonmist |
| BOR | Borderzone | PLA | Planetfall |
| CUT | Cutthroats | PLU | Plunderedhearts |
| DEA | Deadline | SEA | Seastalker |
| ENC | Enchanter | SOR | Sorcerer |
| HIT | Hitchhiker | SPE | Spellbreaker |
| HOL | Hollywoodhijinx | STA | Starcross |
| INF | Infidel | STF | Stationfall |
| LUR | Lurkinghorror | SUS | Suspect |
| SUSP | Suspended | TRI | Trinity |
| WIS | Wishbringer | Z1 | Zork1 |
| WIT | Witness | Z2 | Zork2 |
| | | Z3 | Zork3 |

Table 11: Abbreviations for games in Jiminy Cricket Benchmark.

## D.2 Full Results on Table 3

Game scores and standard deviations across three different runs for all 25 games in the Jiminy-Cricket benchmark are presented in Table 12 and 13.

## D.3 Run Example of Jiminy Cricket Benchmark.

We provide a portion of the transcript obtained from playing one of the games in the Jiminy Cricket benchmark. All of the agents playing the game are granted high openness.

**Actions Candidates:**
take all, take egg, take feather, throw egg at nest, get it, take nest, examine clasp, put egg in nest, get feather, climb nest, get clasp, north, inventory, examine hatch, examine gold, open hatch, down, "examine birds nest", take clasp, examine nest, take all from nest, up, examine necklace, take bird, eat nest, examine feather, examine bracelet, south, take jewel, take gold, take necklace, examine egg, take hatch, break egg, examine jewel
**Observation:**
Up a Tree You are about 10 feet above the ground nestled among some large branches. The nearest branch above you is above your reach. Beside you on the branch is a small bird's nest. In the bird's nest is a large egg encrusted with precious jewels, apparently scavenged by a childless songbird. The egg is covered with fine gold inlay, and ornamented in lapis lazuli and mother-of-pearl. Unlike most eggs, this one is hinged and closed with a delicate looking clasp. Egg appears extremely fragile.
**Action:** get it
**Score:** 4

**Actions Candidates:**
east, hit painting, open painting, examine art, enter painting, open door, examine mural, turn diamond, get painting, climb painting, take painting, southwest, break window, close gallery, south, move painting, northeast, unlock door with key, read painting, push painting, examine paintings, west, southeast, break painting, exits, turn ring, north, examine painting, up, inventory, listen, search painting, examine portrait, northwest, out, down, look under painting, look behind painting, cut painting
**Observation:**
Gallery This is an art gallery. Most of the paintings have been stolen by vandals with exceptional taste. The vandals left through either the north or west exits. Fortunately, there is still one chance for you to be a vandal, for on the far wall is a painting of unparalleled beauty.
**Action:** read painting
**Score:** 4

**Actions Candidates:**
look through window, southwest, out, climb tree, break window, down, examine me, look in it, north, close window, west, look in window, southeast, look, enter window, search house, knock on window, look out window, northeast, examine window, examine door, examine tree, wait, enter house, in, south, go through window, exit, examine house, knock on door, inventory, examine bed, east, enter cabin, up, look through it, again, northwest
**Observation:**
With great effort, you open the window far enough to allow entry.
**Action:** enter window
**Score:** 10

| Game | $A_{\text{N.P}}$ | | $A_{\textbf{Ope.}\uparrow}$ | | $A_{\textbf{Ope.}\downarrow}$ | | $A_{\textbf{Con.}\uparrow}$ | | $A_{\textbf{Con.}\downarrow}$ | | $A_{\textbf{Ext.}\uparrow}$ | | $A_{\textbf{Ext.}\downarrow}$ | | $A_{\textbf{Agr.}\uparrow}$ | | $A_{\textbf{Agr.}\downarrow}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* |
| BAL | 3.4 | 0.4 | 3.5 | 0.1 | 0.8 | 1.2 | 2.6 | 0.7 | 2.7 | 0.6 | 2.9 | 0.4 | 2.9 | 0.4 | 3.2 | 0.2 | 3.4 | 0.1 |
| BOR | 1.9 | 0.2 | 2.2 | 0.0 | 1.0 | 0.5 | 2.0 | 0.1 | 1.4 | 0.3 | 1.4 | 0.3 | 0.6 | 0.3 | 0.9 | 0.3 | 0.6 | 0.1 |
| CUT | 3.8 | 0.1 | 3.9 | 0.1 | 3.6 | 0.1 | 3.7 | 0.1 | 3.8 | 0.2 | 3.9 | 0.1 | 3.8 | 0.0 | 3.8 | 0.1 | 3.8 | 0.1 |
| MOO | 7.0 | 0.2 | 7.6 | 0.2 | 5.6 | 0.1 | 6.9 | 0.4 | 6.1 | 0.7 | 6.2 | 0.5 | 4.9 | 0.1 | 7.2 | 0.4 | 5.5 | 0.5 |
| PLA | 1.8 | 0.2 | 1.9 | 0.1 | 1.6 | 0.0 | 1.7 | 0.2 | 1.8 | 0.2 | 1.7 | 0.1 | 1.7 | 0.1 | 1.8 | 0.1 | 1.8 | 0.1 |
| PLU | 5.3 | 0.2 | 5.5 | 0.1 | 4.3 | 0.8 | 5.0 | 0.1 | 3.6 | 0.2 | 4.4 | 0.2 | 4.8 | 0.1 | 5.3 | 0.1 | 4.3 | 0.3 |
| SEA | 5.5 | 0.2 | 7.3 | 0.5 | 4.1 | 1.6 | 5.8 | 0.4 | 5.7 | 0.6 | 6.3 | 0.7 | 5.9 | 0.5 | 6.6 | 0.9 | 6.0 | 0.2 |
| SOR | 4.1 | 0.2 | 5.2 | 0.1 | 3.0 | 1.3 | 4.5 | 0.5 | 3.0 | 1.3 | 4.4 | 0.2 | 4.5 | 0.5 | 4.1 | 0.7 | 3.1 | 1.6 |
| SPE | 6.5 | 0.1 | 6.6 | 0.0 | 6.2 | 0.1 | 6.4 | 0.1 | 5.0 | 1.5 | 6.8 | 0.1 | 6.3 | 0.2 | 6.6 | 0.2 | 6.5 | 0.2 |
| SUS | 4.1 | 0.9 | 5.9 | 0.3 | 4.1 | 0.5 | 4.5 | 0.7 | 3.9 | 0.6 | 4.4 | 0.3 | 4.1 | 0.1 | 4.1 | 1.4 | 3.0 | 0.7 |
| TRI | 3.9 | 0.1 | 6.9 | 0.2 | 5.6 | 0.1 | 6.3 | 0.3 | 5.4 | 1.1 | 5.6 | 1.3 | 6.6 | 0.0 | 5.6 | 1.3 | 6.6 | 0.2 |
| WIS | 6.1 | 0.1 | 6.0 | 0.0 | 5.8 | 0.0 | 5.8 | 0.1 | 5.8 | 0.1 | 5.9 | 0.0 | 5.8 | 0.0 | 5.8 | 0.1 | 5.8 | 0.1 |
| WIT | 11.1 | 0.4 | 11.4 | 0.3 | 7.5 | 2.0 | 10.6 | 0.2 | 6.5 | 0.6 | 8.3 | 0.6 | 9.7 | 0.5 | 9.2 | 0.7 | 9.1 | 0.3 |
| Z1 | 6.5 | 2.2 | 8.9 | 0.1 | 8.3 | 0.1 | 8.8 | 0.2 | 8.5 | 0.3 | 8.3 | 0.3 | 9.0 | 0.2 | 8.7 | 0.2 | 8.8 | 0.2 |
| Z3 | 14.0 | 0.9 | 15 | 0.2 | 13.7 | 0.8 | 13.0 | 0.4 | 13.2 | 1.0 | 13.0 | 1.0 | 14.6 | 0.7 | 14.3 | 0.4 | 12.6 | 1.3 |
| DEA | 0.4 | 0.6 | 1.3 | 0.4 | 0.0 | 0.0 | 0.6 | 0.4 | 0.0 | 0.0 | 0.2 | 0.3 | 0.0 | 0.0 | 0.3 | 0.2 | 0.1 | 0.2 |
| ENC | 0.0 | 0.0 | 3.0 | 1.8 | 0.0 | 0.0 | 0.0 | 0.0 | 3.1 | 1.6 | 3.0 | 2.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| HIT | 0.1 | 0.2 | 0.1 | 0.1 | 0.0 | 0.0 | 0.1 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| HOL | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| INF | 0.1 | 0.1 | 0.1 | 0.1 | 0.0 | 0.0 | 0.2 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| LUR | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| STAR | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| STAT | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| SUSP | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Z2 | -0.2 | 0.1 | -0.2 | 0.0 | -0.4 | 0.0 | -0.1 | 0.0 | -0.4 | 0.1 | -0.2 | 0.1 | -0.5 | 0.1 | -0.3 | 0.1 | -0.2 | 0.0 |

Table 12: Full scores and standard deviation of Table 3 for agent $A_{\text{Ope.}\uparrow}$, $A_{\text{Ope.}\downarrow}$, $A_{\text{Con.}\uparrow}$, $A_{\text{Con.}\downarrow}$, $A_{\text{Ext.}\uparrow}$, $A_{\text{Ext.}\downarrow}$, $A_{\text{Agr.}\uparrow}$, $A_{\text{Agr.}\downarrow}$.

### D.4 Learning Curve

We report the training progression through learning curves for all eight personality types, measured across a test suite of 15 games.

### D.5 Comparison with Other Methodologies

We compared scores with other text-based game methodologies on the Jiminy Cricket benchmark. The scores of **NAIL** (Hausknecht et al., 2019), **CALM** (Yao et al., 2020a), **CMPS** and **CMPS+** (Hendrycks et al., 2021b), **GALAD** (Ammanabrolu et al., 2022) are from (Ammanabrolu et al., 2022).

Table 15 shows that among our 16 personality-infused game agents, $A_{\text{Ope.}\uparrow}$ achieved the best performance, demonstrating superior scores compared to other baselines.

### D.6 Detailed Criteria for Place Classification

In §5.1, we categorized all locations in the Zork1 game into two groups - *Com.* and *Unc.* - based on their depth from the starting point. This categorization was implemented to analyze the relationship between location accessibility and player navigation patterns. The place lists and corresponding statistics for these two groups are in Table 14.

### D.7 World Visualization and Trajectory of Agent

We applied the visualization method presented in Section 5.1 to all agents, with results shown in Figure 13 to 16. Additionally, we visualize the map of the game *Zork-1*, *Zork-2*, and *Zork-3* from Jiminy Cricket benchmark in in Figure 20 to 22.

### D.8 Prompt used in GPT-4.

The prompts used with the LLM (GPT-4) for dataset construction and personality annotation in this paper are presented in Table 18. We utilized the *gpt-4-turbo-2024-04-09* checkpoint.

## E Personality Framework

Personality plays a crucial role in shaping individual behavior, decision-making, and interactions. In psychological research, various models have been developed to systematically categorize and measure personality traits : the Big Five personality traits and the Dark Triad.

| Game | $A_{\text{N.P}}$ | | $A_{\text{Neu.}\uparrow}$ | | $A_{\text{Neu.}\downarrow}$ | | $A_{\text{Psy.}\uparrow}$ | | $A_{\text{Psy.}\downarrow}$ | | $A_{\text{Nar.}\uparrow}$ | | $A_{\text{Nar.}\downarrow}$ | | $A_{\text{Mac.}\uparrow}$ | | $A_{\text{Mac.}\downarrow}$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* | *Mean* | *Std* |
| BOR | 1.9 | 0.2 | 1.9 | 0.1 | 1.7 | 0.2 | 1.8 | 0.2 | 1.3 | 0.1 | 1.9 | 0.1 | 0.6 | 0.2 | 1.1 | 0.6 | 0.7 | 0.3 |
| BAL | 3.4 | 0.4 | 3.5 | 0.2 | 2.6 | 0.4 | 2.9 | 0.3 | 2.5 | 0.7 | 1.9 | 1.4 | 1.5 | 1.1 | 3.2 | 0.5 | 3.2 | 0.2 |
| CUT | 3.8 | 0.1 | 3.9 | 0.1 | 3.8 | 0.1 | 3.8 | 0.2 | 3.9 | 0.1 | 3.7 | 0.1 | 3.9 | 0.1 | 3.9 | 0.0 | 3.7 | 0.1 |
| MOO | 7.0 | 0.2 | 6.8 | 0.6 | 8.1 | 1.7 | 8.2 | 1.4 | 5.8 | 0.5 | 6.6 | 0.3 | 5.4 | 0.3 | 7.5 | 0.8 | 7.6 | 0.7 |
| PLA | 1.8 | 0.2 | 1.7 | 0.1 | 1.7 | 0.1 | 1.9 | 0.1 | 1.7 | 0.3 | 1.7 | 0.1 | 1.7 | 0.2 | 1.8 | 0.2 | 1.7 | 0.1 |
| PLU | 5.3 | 0.2 | 5.4 | 0.2 | 4.9 | 0.1 | 5.2 | 0.2 | 4.2 | 0.3 | 5.3 | 0.1 | 3.8 | 0.5 | 5.1 | 0.2 | 3.9 | 0.7 |
| SEA | 5.5 | 0.2 | 6.1 | 0.8 | 5.9 | 0.4 | 6.1 | 0.3 | 5.0 | 0.9 | 6.4 | 0.5 | 6.9 | 0.3 | 6.2 | 0.3 | 6.1 | 0.3 |
| SOR | 4.1 | 0.2 | 3.1 | 1.4 | 4.4 | 0.3 | 3.0 | 1.2 | 4.5 | 0.1 | 2.2 | 1.4 | 4.2 | 0.5 | 4.4 | 0.6 | 4.2 | 0.4 |
| SPE | 6.5 | 0.1 | 5.7 | 1.2 | 6.4 | 0.1 | 6.5 | 0.2 | 6.5 | 0.1 | 6.4 | 0.2 | 5.1 | 1.8 | 6.6 | 0.1 | 6.2 | 0.1 |
| SUS | 4.1 | 0.9 | 5.2 | 0.8 | 5.0 | 0.4 | 5.2 | 0.6 | 3.0 | 0.8 | 5.2 | 0.2 | 2.8 | 0.2 | 3.3 | 2.3 | 4.5 | 0.6 |
| TRI | 3.9 | 0.1 | 4.9 | 1.4 | 5.7 | 1.0 | 6.0 | 1.5 | 4.7 | 1.3 | 5.2 | 1.5 | 6.1 | 0.8 | 6.2 | 1.1 | 5.7 | 1.2 |
| WIS | 6.1 | 0.1 | 6.2 | 0.0 | 6.0 | 0.1 | 6.1 | 0.0 | 5.9 | 0.1 | 6.0 | 0.1 | 5.9 | 0.0 | 5.9 | 0.0 | 5.8 | 0.1 |
| WIT | 11.1 | 0.4 | 11.6 | 0.8 | 9.9 | 0.8 | 11.3 | 0.4 | 7.1 | 1.0 | 11.1 | 0.2 | 8.5 | 0.7 | 10.2 | 0.5 | 8.8 | 0.4 |
| Z1 | 6.5 | 2.2 | 7.8 | 1.4 | 9.0 | 0.1 | 8.3 | 0.4 | 8.7 | 0.2 | 8.8 | 0.4 | 8.6 | 0.3 | 8.4 | 0.4 | 7.1 | 2.5 |
| Z3 | 14.0 | 0.9 | 13.2 | 0.5 | 13.9 | 0.6 | 13.8 | 0.1 | 14.1 | 0.6 | 13.8 | 0.8 | 14.1 | 1.0 | 14.9 | 1.2 | 13.9 | 0.8 |
| DEA | 0.4 | 0.6 | 0.3 | 0.5 | 0.9 | 0.3 | 0.3 | 0.4 | 0.4 | 0.5 | 0.5 | 0.7 | 0.2 | 0.3 | 0.7 | 0.3 | 0.0 | 0.0 |
| HIT | 0.1 | 0.2 | 0.0 | 0.0 | 0.1 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| ENC | 0.0 | 0.0 | 0.0 | 0.0 | 2.8 | 2.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 2.8 | 2.0 | 0.0 | 0.0 | 2.6 | 1.9 |
| HOL | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| INF | 0.1 | 0.1 | 0.1 | 0.2 | 0.1 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| LUR | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.1 | 0.1 | 0.0 | 0.0 | 0.0 | 0.0 |
| STAR | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| STAT | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| SUSP | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| Z2 | -0.2 | 0.1 | -0.4 | 0.1 | -0.3 | 0.0 | -0.2 | 0.0 | -0.3 | 0.1 | -0.1 | 0.0 | -0.2 | 0.1 | -0.1 | 0.0 | -0.2 | 0.1 |

Table 13: Full scores and standard deviation of Table 3 for agent $A_{\text{Neu.}\uparrow}$, $A_{\text{Neu.}\downarrow}$, $A_{\text{Psy.}\uparrow}$, $A_{\text{Psy.}\downarrow}$, $A_{\text{Nar.}\uparrow}$, $A_{\text{Nar.}\downarrow}$, $A_{\text{Mac.}\uparrow}$, $A_{\text{Mac.}\downarrow}$.

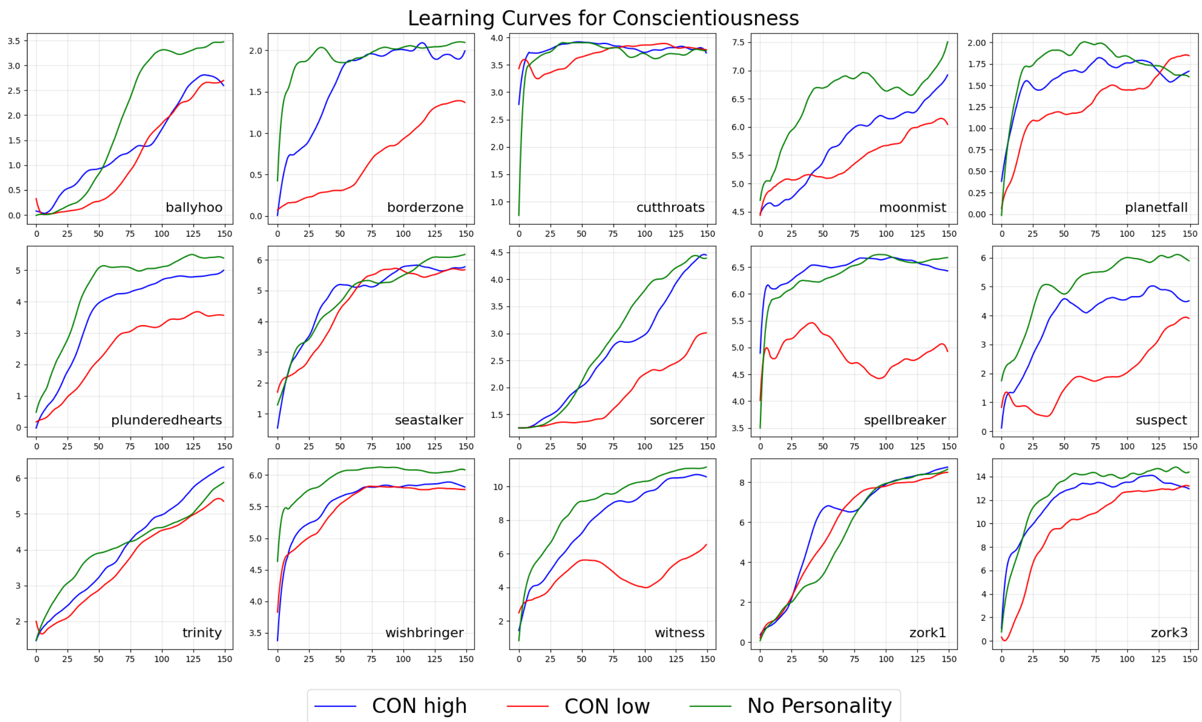


Figure 6: Learning curve each 15 games in Jiminy Cricket benchmark.

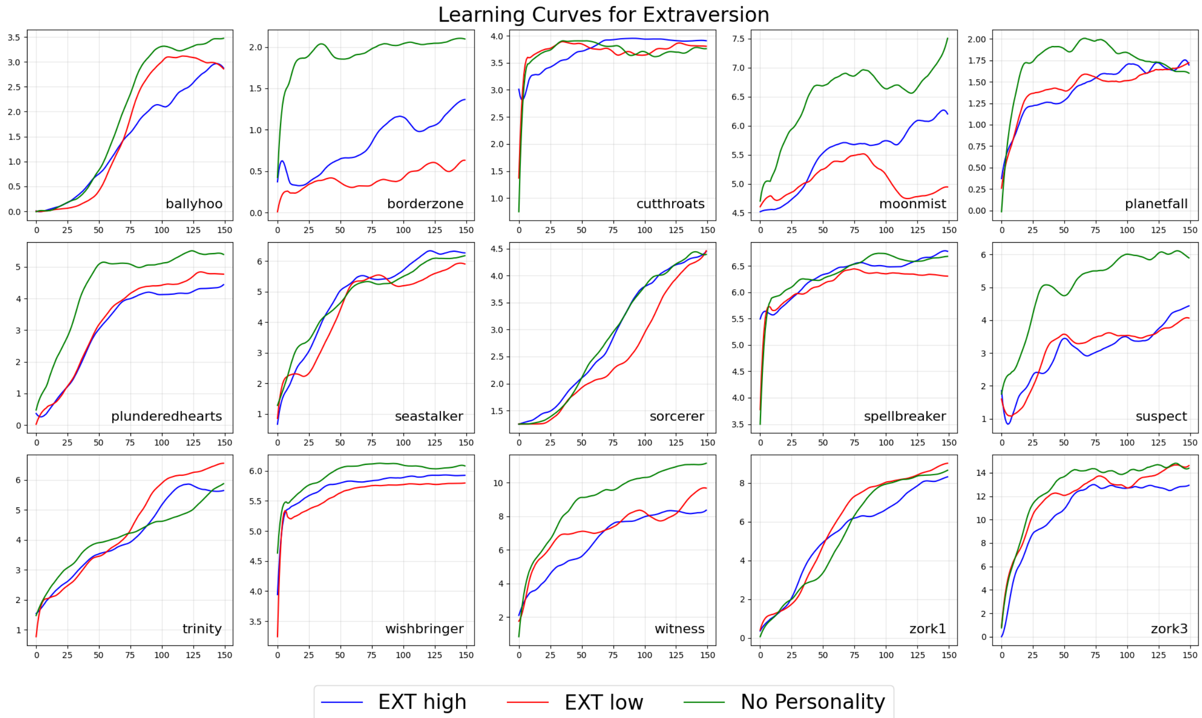


Figure 7: Learning curve each 15 games in Jiminy Cricket benchmark.

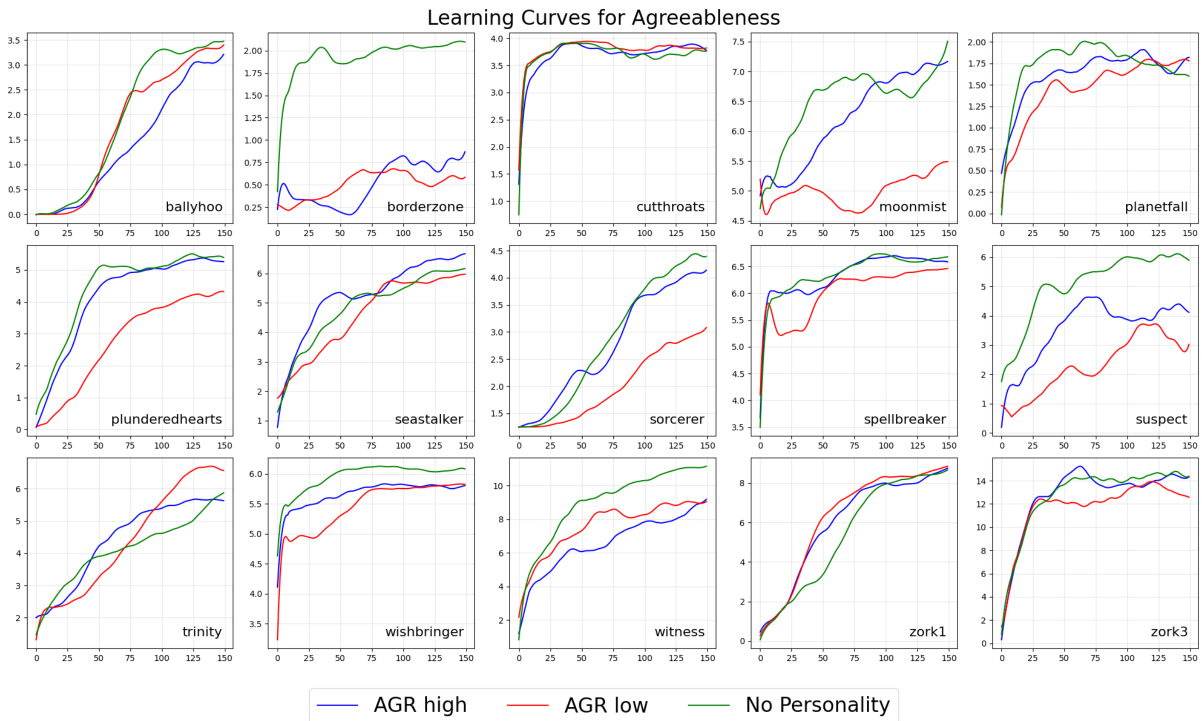


Figure 8: Learning curve each 15 games in Jiminy Cricket benchmark.

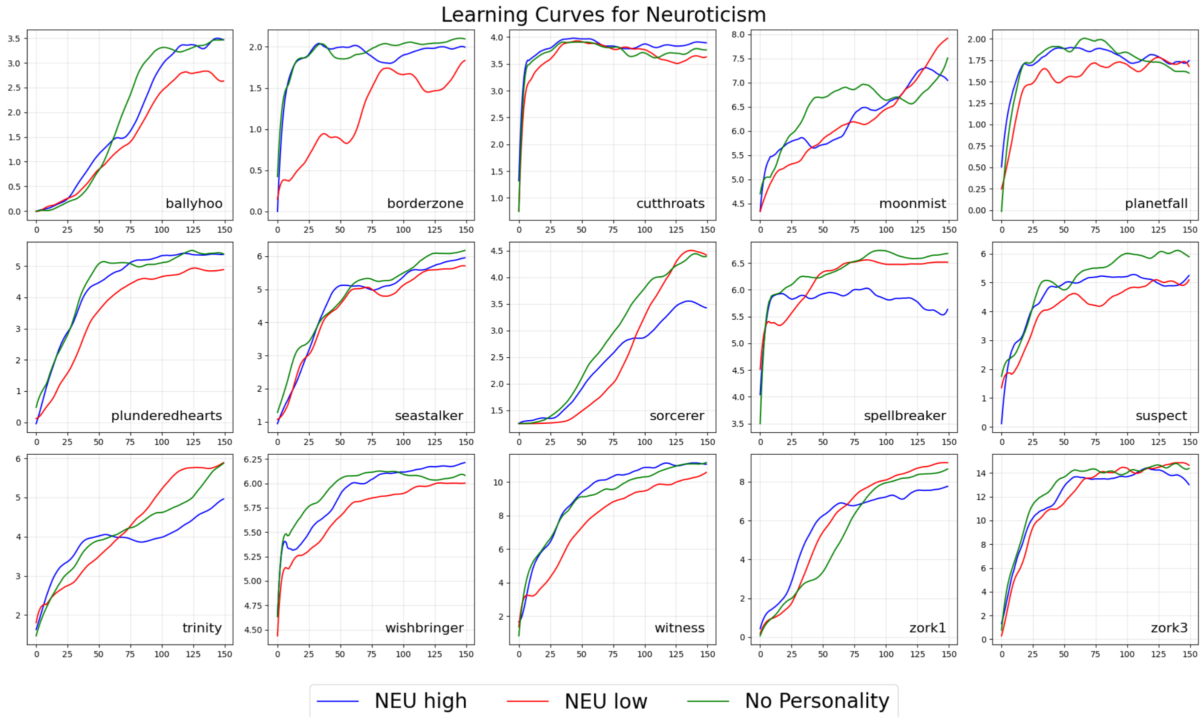


Figure 9: Learning curve each 15 games in Jiminy Cricket benchmark.

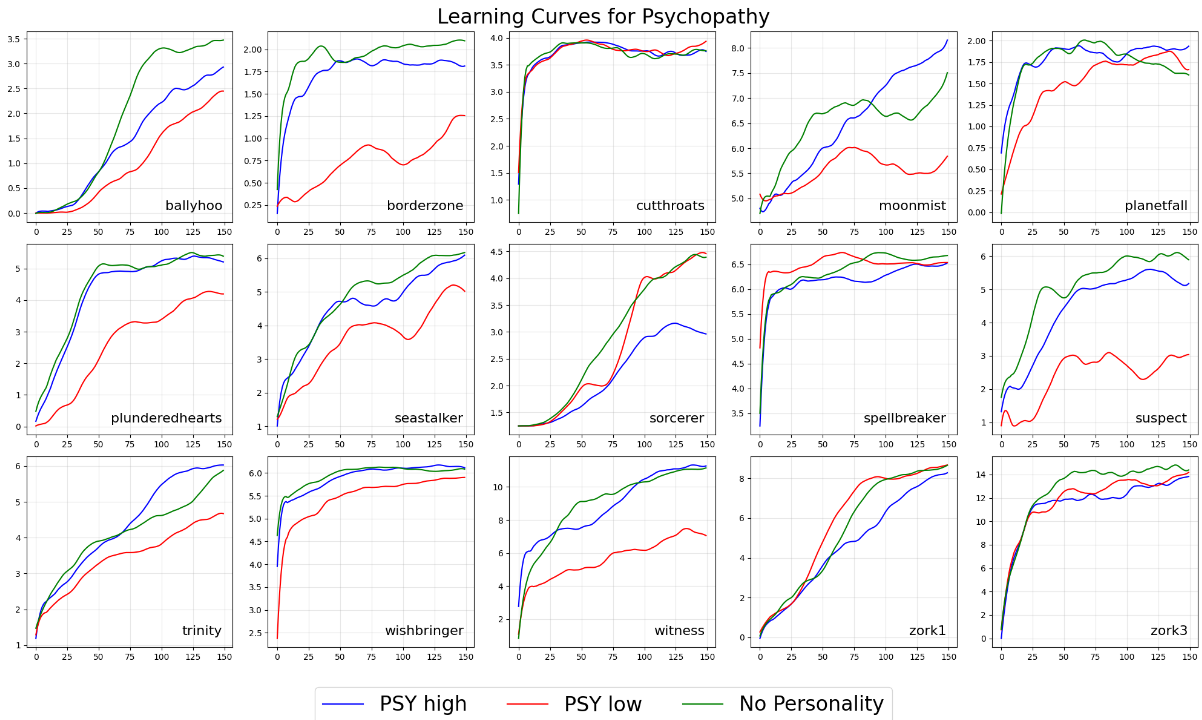


Figure 10: Learning curve each 15 games in Jiminy Cricket benchmark.

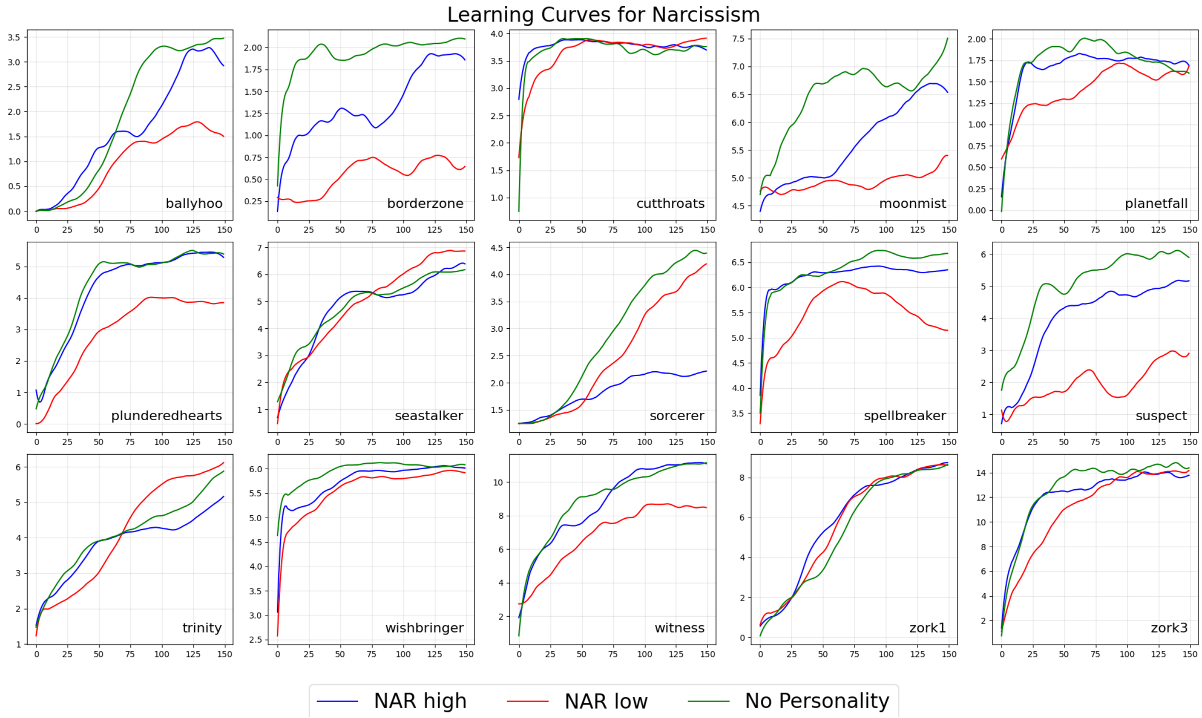


Figure 11: Learning curve each 15 games in Jiminy Cricket benchmark.

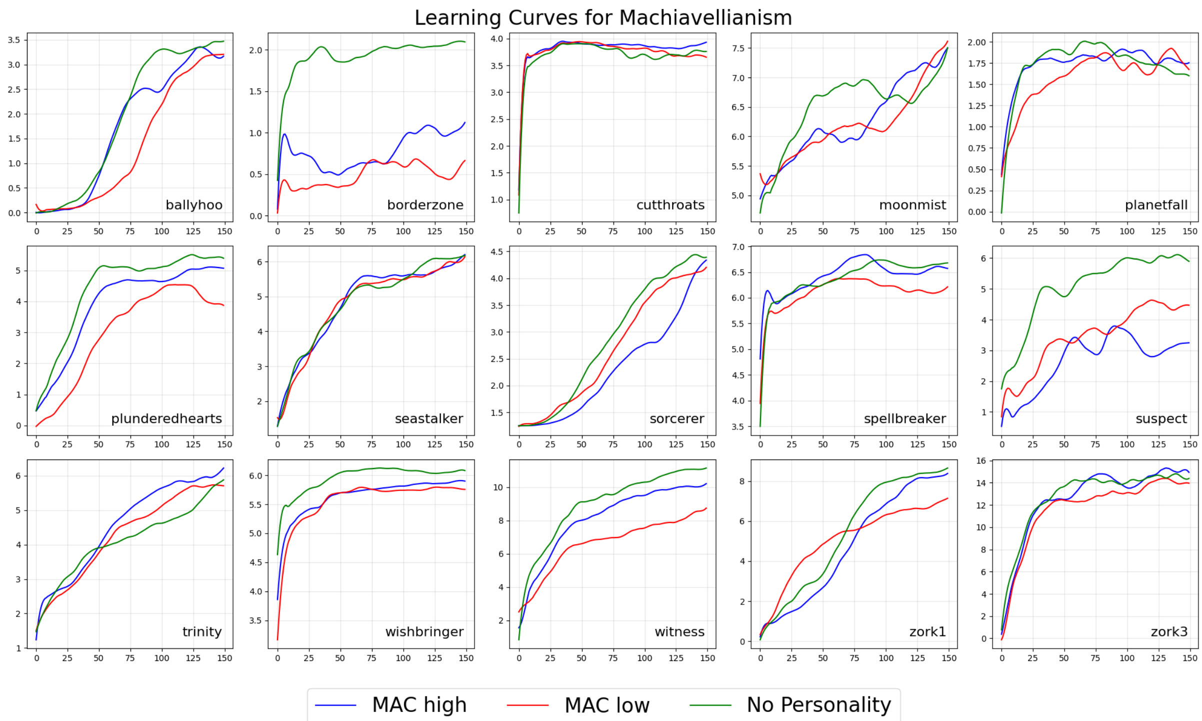


Figure 12: Learning curve each 15 games in Jiminy Cricket benchmark.

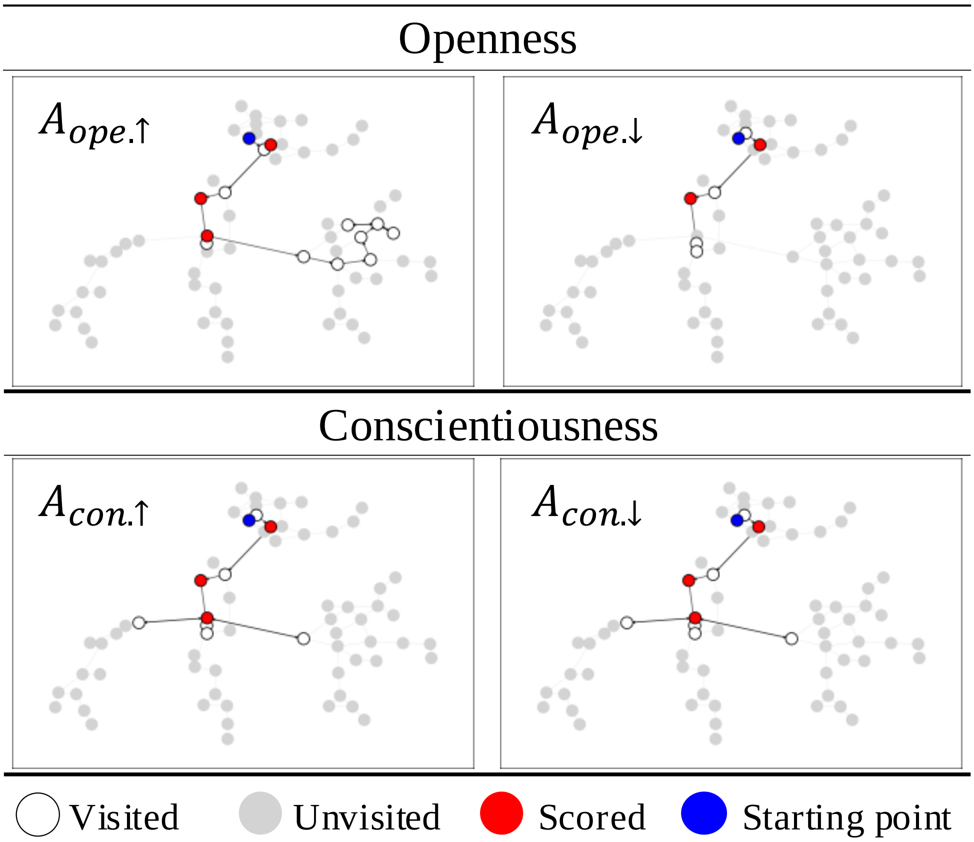


Figure 13: Trajectory of $A_{\text{Ope.}\uparrow}$, $A_{\text{Ope.}\downarrow}$, $A_{\text{Con.}\uparrow}$ and $A_{\text{Con.}\downarrow}$ in Zork1.

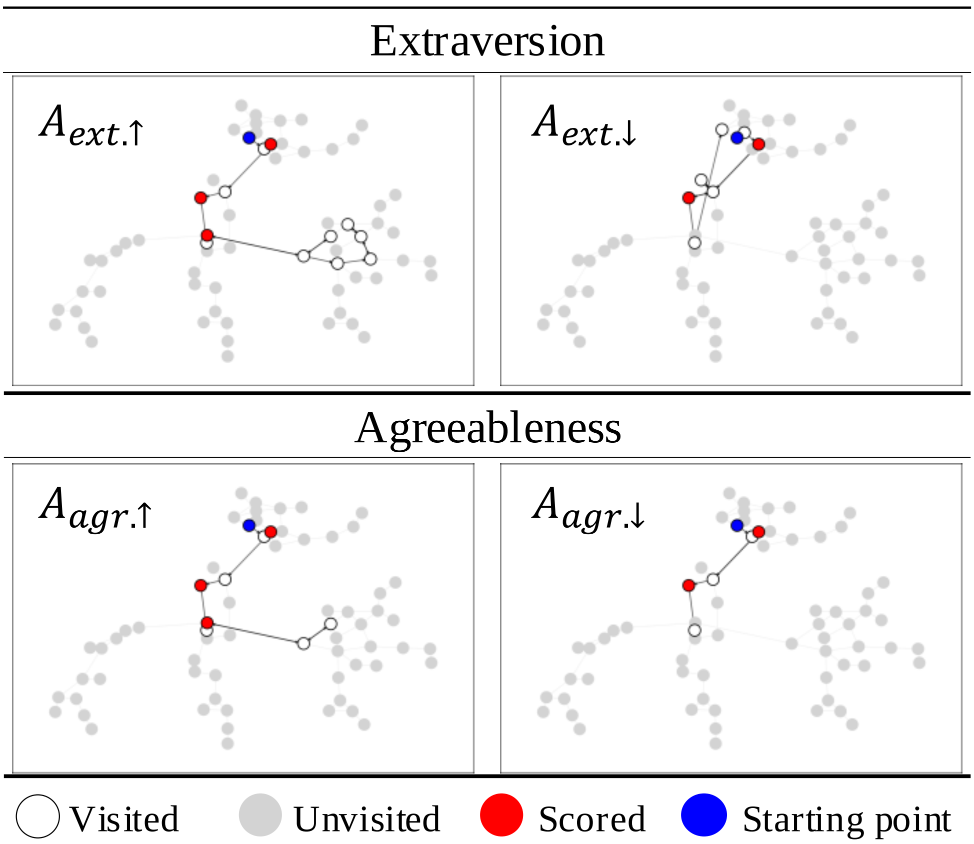


Figure 14: Trajectory of $A_{\text{Ext.}\uparrow}$, $A_{\text{Ext.}\downarrow}$, $A_{\text{Agr.}\uparrow}$ and $A_{\text{Agr.}\downarrow}$ in Zork1.

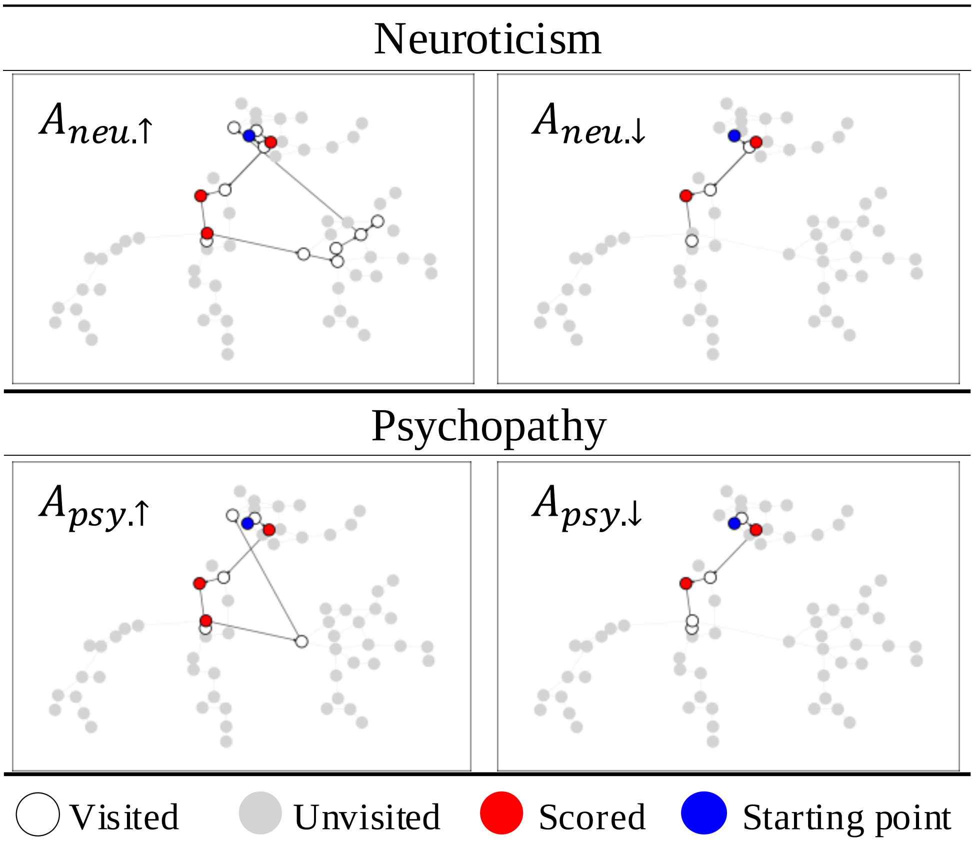


Figure 15: Trajectory of $A_{\text{Neu.}\uparrow}$, $A_{\text{Neu.}\downarrow}$, $A_{\text{Psy.}\uparrow}$ and $A_{\text{Psy.}\downarrow}$ in Zork1.

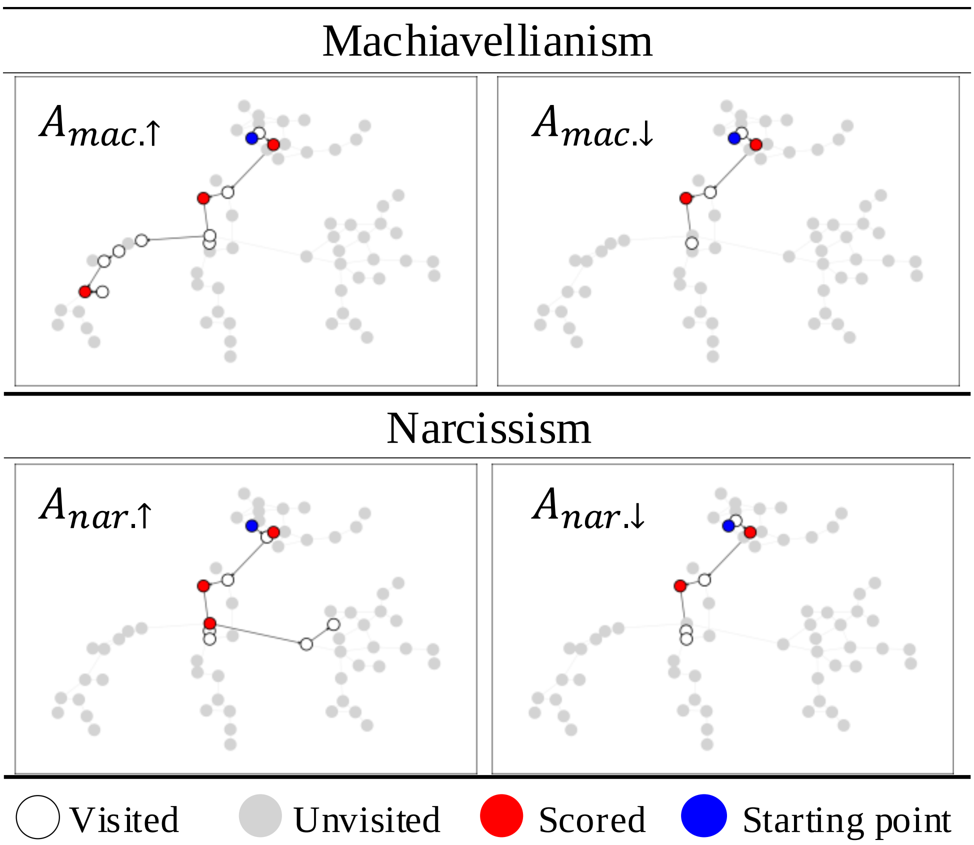


Figure 16: Trajectory of $A_{\text{Mac.}\uparrow}$, $A_{\text{Mac.}\downarrow}$, $A_{\text{Nar.}\uparrow}$ and $A_{\text{Nar.}\downarrow}$ in Zork1.

| | # Room | Visit Cnt(M) | Visit Cnt/ # Room(K) | Visit Ratio |
|---|---|---|---|---|
| *Com.* | 13 | 2.9M | 227.7K | 82.6% |
| *Unc.* | 53 | 2.5M | 47.8K | 17.4% |

Table 14: Statistical Analysis of places *Com.* and *Unc.*

### E.1 Big Five and Dark Triad

The Big Five personality traits, also known as the Five-Factor Model, is one of the most widely accepted frameworks for understanding personality. It categorizes personality into five broad dimensions: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. Big five shows high reliability and validity across cultures and times.

Dark Triad focuses on socially aversive traits: Machiavellianism, Narcissism, and Psychopathy. Machiavellianism is a trait to manipulate or deceit other people with strategic thinking for their own benefit. Psychopathy is charaterized by impulsivity, a lack of remorse or guilt, antisocial behavior, and a lack of empathy. Finally, Narcissism is a trait of grandiosity, pride, egotism, and a lack of empathy, and high narcissism have inflated sense of their own importance.

### E.2 BFI and SD-3

To measure these personality traits, psychologists have developed various assessment tools. The Big Five Inventory (BFI) is one of the most commonly used instruments for assessing the Big Five personality traits. It consists of a series of statements that respondents rate based on how accurately they reflect their own behavior and preferences. For assessing the Dark Triad traits, the Short Dark Triad (SD-3) questionnaire is widely used. The SD-3 is a brief yet effective measure, designed to assess Machiavellianism, Narcissism, and Psychopathy with just a few items per trait. Full set of BFI and SD-3 are in listed in Table 19 and 20.

## F Personality Data

### F.1 Paraphrased Personality Description

In Table 21 and Table 22, we list the full set of paraphrased personality descriptions ($n = 80$) used in the data making pipeline. We generate it with GPT-4, and '(R)' means a sentence reveals low level of the given personality trait.

### F.2 Situational Seeds

In Table 23, we list subset of the situational seeds ($n = 300$) used in the data making pipeline. We generate it with GPT-4, and uploaded 10% of the full set.

### F.3 Word Distribution

In Figure 19, we measure diverse side of personality dataset. Firstly, we draw a pie chart with two circles about most frequently used verb and noun to show a property of our dataset. Second, we do lexical analysis with the tool of LIWC, a well-known framework to statistically analysis the word distribution of given corpus.

| Game/Agent | NAIL | CALM | CMPS | CMPS+ | GALAD | $A_{\text{Ope.}\uparrow}$ |
|---|---|---|---|---|---|---|
| BAL | 0.3 | 2.5 | 1.2 | 2.2 | 1.6 | 3.5 |
| BOR | 1.4 | 3.6 | 3.3 | 3.7 | 3.5 | 3.3 |
| CUT | 4.2 | 3.9 | 3.8 | 3.6 | 3.8 | 3.9 |
| DEA | 0.8 | 1.6 | 1.6 | 1.7 | 1.8 | 1.3 |
| ENC | 0.0 | 1.8 | 1.7 | 3.6 | 3.2 | 3.0 |
| HIT | 0.0 | 7.9 | 7.2 | 10.5 | 10.0 | 0.1 |
| HOL | 0.3 | 1.7 | 1.8 | 1.6 | 1.8 | 0.0 |
| INF | 0.1 | 0.4 | 0.4 | 0.4 | 0.4 | 0.1 |
| LUR | 0.0 | 0.4 | 0.8 | 0.3 | 0.3 | 0.0 |
| MOO | 7.1 | 9.3 | 9.3 | 8.2 | 10.9 | 7.6 |
| PLA | 0.5 | 1.6 | 1.3 | 1.6 | 2.2 | 1.9 |
| PLU | 1.0 | 2.7 | 2.8 | 2.8 | 3.2 | 5.5 |
| SEA | 1.0 | 3.4 | 4.4 | 3.9 | 4.4 | 7.3 |
| SOR | 0.5 | 2.6 | 2.6 | 2.6 | 1.8 | 5.2 |
| SPE | 0.6 | 3.4 | 3.4 | 3.4 | 3.3 | 6.6 |
| STAR | -1.7 | -0.1 | -0.1 | -0.1 | 1.3 | 0.0 |
| STAT | 0.7 | 0.3 | 0.2 | 0.3 | 0.4 | 0.0 |
| SUS | 3.5 | 5.1 | 4.3 | 4.8 | 4.4 | 5.9 |
| SUSP | -1.7 | -0.7 | -0.8 | -0.4 | -0.7 | 0.0 |
| TRI | 0.1 | 1.6 | 1.6 | 1.5 | 1.6 | 6.9 |
| WIS | 0.3 | 5.0 | 5.1 | 5.0 | 5.2 | 6.0 |
| WIT | 2.8 | 9.2 | 8.6 | 9.2 | 9.9 | 11.4 |
| Z1 | -2.4 | 5.3 | 5.1 | 5.3 | 5.2 | 8.9 |
| Z2 | -2.5 | 2.5 | 4.0 | 2.5 | 2.4 | -0.2 |
| Z3 | 5.2 | 12.2 | 11.1 | 12.2 | 12.0 | 15.0 |
| **Average** | 1.7 | 4.8 | 4.6 | 4.6 | 4.9 | 4.1 |

Table 15: Comparison with previous text-game adventure agents. We report $A_{\text{Ope.}\uparrow}$ as a representative example of the **PANDA** framework

| Metric | | $A_{\text{N.P}}$ | $A_{\text{Ope.}\uparrow}$ | $A_{\text{Ope.}\downarrow}$ | $A_{\text{Con.}\uparrow}$ | $A_{\text{Con.}\downarrow}$ | $A_{\text{Ext.}\uparrow}$ | $A_{\text{Ext.}\downarrow}$ | $A_{\text{Agr.}\uparrow}$ | $A_{\text{Agr.}\downarrow}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| *Trajectory Length (↓)* | - | 45.85±3.2 | 57 ±2.6 | 39.9 ±2.1 | 50.1 ±5.2 | 51.3 ±8.2 | 60.9 ±20.1 | 55.5 ±3.5 | 49.2 ±3.3 | 47.6±3.8 |
| *Visit Count (↑)* | *Com.* | 8.66±0.3 | 9.0 ±0.4 | 8.0 ±0.1 | 8.9 ±0.3 | 8.5 ±0.3 | 7.8 ±0.9 | 8.4 ±0.4 | 8.6 ±0.3 | 8.4±0.7 |
| | *Unc.* | 0.83±0.2 | 1.20±0.4 | 0.30±0.1 | 0.89±0.1 | 0.88±0.3 | 0.67±0.4 | 1.21±0.1 | 0.82±0.2 | 1.01±0.5 |
| *Avg. Step (↓)* | *Com.* | 12.64±2.1 | 11.9 ±1.0 | 11.6 ±0.4 | 14.5 ±0.6 | 13.2 ±0.9 | 10.3 ±1.7 | 12.8 ±0.2 | 12.3 ±0.6 | 11.7±1.5 |
| | *Unc.* | 8.62±3.5 | 6.4 ±4.4 | 12 ±3.2 | 17.5 ±2.5 | 13.5 ±4.1 | 6.1 ±3.1 | 16.7 ±4.2 | 9.4 ±2.2 | 12.3±6.0 |

Table 16: Full Results on Table 5 for $A_{\text{N.P}}$, $A_{\text{Ope.}\uparrow}$, $A_{\text{Ope.}\downarrow}$, $A_{\text{Con.}\uparrow}$, $A_{\text{Con.}\downarrow}$, $A_{\text{Ext.}\uparrow}$, $A_{\text{Ext.}\downarrow}$, $A_{\text{Agr.}\uparrow}$, and $A_{\text{Agr.}\downarrow}$.

| Metric | | $A_{\text{N.P}}$ | $A_{\text{Neu.}\uparrow}$ | $A_{\text{Neu.}\downarrow}$ | $A_{\text{Psy.}\uparrow}$ | $A_{\text{Psy.}\downarrow}$ | $A_{\text{Mac.}\uparrow}$ | $A_{\text{Mac.}\downarrow}$ | $A_{\text{Nar.}\uparrow}$ | $A_{\text{Nar.}\downarrow}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| *Trajectory Length (↓)* | - | 45.85±3.2 | 48.9 ±2.9 | 54.7 ±10.0 | 50.4 ±0.9 | 44.9 ±0.3 | 46.1 ±7.5 | 53.7 ±8.8 | 48.7 ±5.5 | 46.2±3.6 |
| *Visit Count (↑)* | *Com.* | 8.66±0.3 | 8.9 ±0.2 | 8 ±0.7 | 8.3 ±0.6 | 8.1 ±0.2 | 8.6 ±0.6 | 8.2 ±0.2 | 8.1 ±0.5 | 8.1±0.5 |
| | *Unc.* | 0.83±0.2 | 1.2 ±0.4 | 0.6 ±0.2 | 0.8 ±0.6 | 0.3 ±0.2 | 1 ±0.7 | 1.1 ±0.9 | 0.6 ±0.5 | 0.5±0.1 |
| *Avg. Step (↓)* | *Com.* | 12.64±2.1 | 13.8 ±1.4 | 10 ±1.0 | 13.3 ±0.8 | 12.3 ±0.8 | 12 ±1.3 | 13.7 ±3.3 | 12.4 ±1.5 | 12.1±0.5 |
| | *Unc.* | 8.62±3.5 | 16.9 ±1.8 | 9 ±5.7 | 8.5 ±4.0 | 7.8 ±2.9 | 9.8 ±4.7 | 12.1 ±2.6 | 8.9 ±6.5 | 8.6±2.5 |

Table 17: Full Results on Table 5 for $A_{\text{N.P}}$, $A_{\text{Neu.}\uparrow}$, $A_{\text{Neu.}\downarrow}$, $A_{\text{Psy.}\uparrow}$, $A_{\text{Psy.}\downarrow}$, $A_{\text{Mac.}\uparrow}$, $A_{\text{Mac.}\downarrow}$, $A_{\text{Nar.}\uparrow}$, and $A_{\text{Nar.}\downarrow}$.

| **GPT-4 annotation in § 5.3.** |
|---|
| VARIABLES: PERSONALITY, ACTION |
| PROMPT<br>For given action, Determine whether the action exhibits high [PERSONALITY] or low [PERSONALITY] or is neutral with respect to [PERSONALITY].<br>You can choose from the following options, you should choose only one option, without any description.<br>Action: [ACTION] |
| **Acquiring 10 Personality Description in § 3.1.** |
| VARIABLES: PERSONALITY, DESCRIPTION |
| PROMPT<br>Please paraphrase the following sentences describing the trait of [PERSONALITY].<br>Generate 10 semantically distinct paraphrases:<br>- 5 paraphrases that emphasize high levels of the trait, and 5 paraphrases that emphasize low levels of the trait.<br>- Each paraphrase should reflect different aspects and nuances of the trait without overlapping.<br>Descriptions: [DESCRIPTION] |
| **Acquiring 300 Diverse Situation in § 3.1.** |
| VARIABLES: - |
| PROMPT<br>Generate 300 most common everyday places.<br>- Categorize them into 30 sub-categories, with 10 places in each category.<br>- List only the places without descriptions. |
| **Augmenting 5 detailed sentences in § 3.1.** |
| VARIABLES: PERSONALITY DESCRIPTION, PLACE |
| PROMPT<br>Based on the following everyday place and personality description, generate 5 possible behaviors that this person might exhibit.<br>- Each behavior should be distinct and semantically different from the others.<br>- The behaviors should be plausible and realistic in the context of the given place and personality description.<br>Place: [PLACE]<br>Personality Description: [PERSONALITY DESCRIPTION] |

Table 18: Prompts that were used in our work.

| Step | Extraversion (High) | Extraversion (Low) |
|---|---|---|
| Place | Living Room | Living Room |
| Description | I enjoy engaging in conversation with others frequently. | I often prefer to listen rather than speak in social settings. |
| Data | 1. I start a lively discussion about the latest TV shows.<br>2. I often host dinner parties to engage with different people. | 1. I quietly observe conversations rather than actively participating.<br>2. I keep my comments brief when asked for my opinion. |

| Step | Openness (High) | Openness (Low) |
|---|---|---|
| Place | Living Room | Living Room |
| Description | I often devise unique solutions and original creations. | I gravitate towards tasks that are consistent and unvarying. |
| Data | 1. I create a multi-functional furniture piece for the living room.<br>2. I design a unique piece of artwork for the living room wall. | 1. I always sit in the same spot on the couch.<br>2. I stick to the same routine of cleaning the living room every Saturday morning. |

| Step | Conscientiousness (High) | Conscientiousness (Low) |
|---|---|---|
| Place | Living Room | Living Room |
| Description | I work in a methodical and systematic manner. | I tend to overlook details and make mistakes due to a lack of attention. |
| Data | 1. I follow a set cleaning routine for the living room each week.<br>2. I systematically sort and arrange the DVD collection in alphabetical order. | 1. I often misplace the remote control due to not paying attention.<br>2. I forget to water the plants regularly. |

| Step | Agreeableness (High) | Agreeableness (Low) |
|---|---|---|
| Place | Living Room | Living Room |
| Description | My interactions are marked by thoughtfulness and a gentle approach toward everyone. | I tend to be critical and often pinpoint others' shortcomings. |
| Data | 1. I ask everyone's opinion before choosing a movie to watch.<br>2. I speak softly not to disturb those who are reading or studying. | 1. I criticize the arrangement of the furniture.<br>2. I point out the dust on the bookshelf. |

| Step | Neuroticism (High) | Neuroticism (Low) |
|---|---|---|
| Place | Living Room | Living Room |
| Description | I am prone to excessive worrying. | I generally maintain a relaxed demeanor, even under pressure. |
| Data | 1. I fret about guests spilling drinks on the carpet.<br>2. I worry about the kids damaging the furniture when they play. | 1. I calmly discuss disagreements without raising my voice.<br>2. I comfortably entertain guests, not worrying about minor details. |

Figure 17: Examples of our dataset and its originating seed sample, for *Ope., Con, Ext., Agr., and Neu.*

| **Step** | **Machiavellianism (High)** | **Machiavellianism (Low)** |
|---|---|---|
| Place | Living Room | Living Room |
| Description | Sees others as pawns in their scheme, believing in the ease of manipulating most people. | Engages directly in conflicts instead of avoiding them for potential future gain. |
| Data | 1. I convince others to move the furniture according to my preference.<br>2. I manipulate others into agreeing with my TV program choices. | 1. I discuss the disagreement with my roommate openly instead of ignoring it.<br>2. I speak up when I disagree with a friend's viewpoint. |

| **Step** | **Narcissism (High)** | **Narcissism (Low)** |
|---|---|---|
| Place | Living Room | Living Room |
| Description | I hold a belief in my uniqueness, reinforced by frequent affirmations from others. | I tend to feel uncomfortable and uneasy when receiving praise or accolades from others. |
| Data | 1. I decorate the living room to reflect my unique style.<br>2. I always have the most unique and interesting stories to share. | 1. I deflect compliments by praising others.<br>2. I downplay my achievements when they are brought up. |

| **Step** | **Psychopathy (High)** | **Psychopathy (Low)** |
|---|---|---|
| Place | Living Room | Living Room |
| Description | I believe retribution should be immediate and severe. | I do not seek to cause others to regret their actions towards me. |
| Data | 1. I immediately remove a roommate's belongings from the living room if they upset me.<br>2. I disconnect the WiFi as punishment if someone streams too much in the living room. | 1. I do not retaliate when my sibling uses my favorite chair without asking.<br>2. I do not hold grudges when my friend spills drink on my carpet. |

Figure 18: Examples of our dataset and its originating seed sample, for *Mac., Nar., and Psy.*

| Openness |
|---|
| I am original and come up with new ideas. |
| I am curious about many different things. |
| I am ingenious and a deep thinker. |
| I have an active imagination. |
| I am inventive. |
| I value artistic and aesthetic experiences. |
| I prefer work that is routine. (R) |
| I like to reflect and play with ideas. |
| I have few artistic interests. (R) |
| I am sophisticated in art, music, or literature. |
| **Conscientiousness** |
| I do a thorough job. |
| I can be somewhat careless. (R) |
| I am a reliable worker. |
| I tend to be disorganized. (R) |
| I tend to be lazy. (R) |
| I persevere until the task is finished. |
| I do things efficiently. |
| I make plans and follow through with them. |
| I am easily distracted. (R) |
| **Extraversion** |
| I am talkative. |
| I am reserved. (R) |
| I am full of energy. |
| I generate a lot of enthusiasm. |
| I tend to be quiet. (R) |
| I have an assertive personality. |
| I am sometimes shy and inhibited. (R) |
| I am outgoing and sociable. |
| **Agreeableness** |
| I tend to find fault with others. (R) |
| I am helpful and unselfish with others. |
| I start quarrels with others. (R) |
| I have a forgiving nature. |
| I am generally trusting. |
| I can be cold and aloof. (R) |
| I am considerate and kind to almost everyone. |
| I am sometimes rude to others. (R) |
| I like to cooperate with others. |

Table 19: questionnaire items in BFI (John and Srivastava, 1999). (R) indicates ’Reversed’, which means a low tendency toward that personality trait.

| Psychopathy |
| --- |
| I like to get revenge on authorities.<br>I avoid dangerous situations. (R)<br>Payback needs to be quick and nasty.<br>People often say I'm out of control.<br>It's true that I can be mean to others.<br>People who mess with me always regret it.<br>I have never gotten into trouble with the law. (R)<br>I enjoy having sex with people I hardly know.<br>I'll say anything to get what I want. |
| **Narcissism** |
| People see me as a natural leader.<br>I hate being the center of attention. (R)<br>Many group activities tend to be dull without me.<br>I know that I am special because everyone keeps telling me so.<br>I like to get acquainted with important people.<br>I feel embarrassed if someone compliments me. (R)<br>I have been compared to famous people.<br>I am an average person. (R)<br>I insist on getting the respect I deserve. |
| **Machiavellianism** |
| It's not wise to tell your secrets.<br>I like to use clever manipulation to get my way.<br>Whatever it takes, you must get the important people on your side.<br>Avoid direct conflict with others because they may be useful in the future.<br>It's wise to keep track of information that you can use against people later.<br>You should wait for the right time to get back at people.<br>There are things you should hide from other people to preserve your reputation.<br>Make sure your plans benefit yourself, not others.<br>Most people can be manipulated. |

Table 20: questionnaire items in SD-3 (Jones and Paulhus, 2014).

| Personality Type | Description |
|---|---|
| Machiavellianism | Tends to keep personal information and strategies concealed to maintain leverage. |
| | Employs strategic and often covert manipulation to achieve desired outcomes. |
| | Prioritizes winning the favor of influential individuals for personal gain. |
| | Sees others as pawns in their scheme, believing in the ease of manipulating most people. |
| | Believes in self-serving tactics, ensuring personal advantage in plans and interactions. |
| | Openly shares personal secrets, disregarding potential strategic advantages. (R) |
| | Prefers straightforward and honest interactions over cunning manipulation. (R) |
| | Chooses not to focus on courting favor with influential people, valuing equality in relationships. (R) |
| | Engages directly in conflicts instead of avoiding them for potential future gain. (R) |
| | Does not collect damaging information on others, believing in transparency and fairness. (R) |
| Psychopathy | I have a tendency to retaliate against figures of authority. |
| | I believe retribution should be immediate and severe. |
| | I am often perceived as lacking self-restraint. |
| | I have a propensity for being intentionally unkind. |
| | I engage in sexual activities with individuals I am not well-acquainted with. |
| | I steer clear of situations that could be harmful. (R) |
| | I have a history of abiding by the law. (R) |
| | I do not seek to cause others to regret their actions towards me. (R) |
| | I rarely, if ever, exhibit mean-spirited behavior towards others. (R) |
| | I refrain from manipulative speech to achieve my objectives. (R) |
| Narcissism | I am often viewed as someone with inherent leadership qualities. |
| | My presence is generally perceived as essential for making group events engaging. |
| | I hold a belief in my uniqueness, reinforced by frequent affirmations from others. |
| | I actively seek to connect with individuals of high status or significance. |
| | I demand recognition and the proper deference from others due to my perceived worth. |
| | I have a preference for avoiding the spotlight and not being the focal point in social situations. (R) |
| | I tend to feel uncomfortable and uneasy when receiving praise or accolades from others. (R) |
| | I consider myself to be on par with the average person, without any exceptional traits setting me apart. (R) |
| | Group activities can be just as enjoyable for me, regardless of my involvement or contribution. (R) |
| | The idea of comparing myself to celebrities or notable figures doesn't resonate with me; I see no similarity. (R) |
| Openness | I possess a knack for creativity and generating novel concepts. |
| | My interests span a broad range of topics, and I'm eager to explore them. |
| | I'm known for my clever problem-solving abilities and thoughtful insights. |
| | My mind frequently ventures into realms of fancy and hypothetical scenarios. |
| | I often devise unique solutions and original creations. |
| | I gravitate towards tasks that are consistent and unvarying. (R) |
| | My hobbies and interests are relatively specialized and limited in variety. (R) |
| | I typically don't engage in extensive contemplation or daydreaming. (R) |
| | Artistic and cultural pursuits do not significantly resonate with me. (R) |
| | I don't consider myself particularly well-versed or cultured in the arts and humanities. (R) |

Table 21: Paraphrased personality description for Machiavellianism, Psychopathy, Narcissism, and Openness.

| Personality Type | Description |
|---|---|
| Conscientiousness | I'm diligent and meticulous in my work. |
| | I'm dependable and consistently complete my work to a high standard. |
| | I'm persistent and see tasks through to completion without giving up. |
| | I work in a methodical and systematic manner. |
| | I'm proactive in organizing my activities and stick to the plans I set. |
| | I tend to overlook details and make mistakes due to a lack of attention. (R) |
| | I struggle with maintaining order and often have a cluttered workspace. (R) |
| | I have a propensity for procrastination and not fully applying myself to tasks. (R) |
| | I don't always follow through on tasks and can leave things unfinished. (R) |
| | I find it hard to stay focused and am frequently sidetracked by interruptions. (R) |
| Neuroticism | I frequently feel despondent and downhearted. |
| | I often experience tension and unease. |
| | I am prone to excessive worrying. |
| | My mood swings can be quite pronounced. |
| | I tend to succumb to nervousness with little provocation. |
| | I generally maintain a relaxed demeanor, even under pressure. (R) |
| | I am able to confront stress without becoming upset. (R) |
| | My emotional disposition is predominantly stable. (R) |
| | I stay composed and unflustered during stressful events. (R) |
| | I rarely experience undue nerves or anxiety in challenging situations. (R) |
| Extraversion | I enjoy engaging in conversation with others frequently. |
| | I have a lively and vibrant energy. |
| | My presence often inspires excitement and eagerness in others. |
| | I confidently express my thoughts and opinions. |
| | I thrive in the company of others and enjoy meeting new people. |
| | I often prefer to listen rather than speak in social settings. (R) |
| | I tend to keep to myself and enjoy solitude. (R) |
| | In groups, I usually speak less and maintain a calm demeanor. (R) |
| | I approach social interactions more cautiously or with hesitation. (R) |
| | I enjoy having a smaller circle of close friends rather than a wide social network. (R) |
| Agreeableness | I often go out of my way to assist others and put their needs before my own. |
| | I hold a compassionate attitude, easily pardoning others' mistakes or transgressions. |
| | I am characterized by a default position of believing in people's good intentions. |
| | My interactions are marked by thoughtfulness and a gentle approach toward everyone. |
| | I have a strong inclination toward collaborative efforts and seek harmony in group settings. |
| | I tend to be critical and often pinpoint others' shortcomings. (R) |
| | I have a propensity for initiating disputes and engaging in confrontations. (R) |
| | My demeanor can often be perceived as detached or lacking in warmth. (R) |
| | There are times when I disregard social niceties and come off as abrasive. (R) |
| | I have a tendency to prioritize my interests, which might lead to less altruistic behavior. (R) |

Table 22: Paraphrased personality description for Conscientiousness, Neuroticism, Extraversion, and Agreeableness.

| Home and Family | Workplaces | Educational Settings |
|---|---|---|
| • Living room | • Office | • Classroom |
| • Kitchen | • Conference room | • School playground |
| • Dining table | • Break room | • University campus |
| • Backyard | • Co-working space | • Library |
| • Family reunion | • Factory floor | • Laboratory |
| • Birthday party | • Construction site | • Tutoring center |
| • Wedding | • Retail store | • School cafeteria |
| • Funeral | • Warehouse | • Student lounge |
| • Family vacation | • Doctor's office | • Dormitory |
| • Bedroom | • Hospital ward | • School bus |

Table 23: Samples of situation seeds used in making personality dataset.

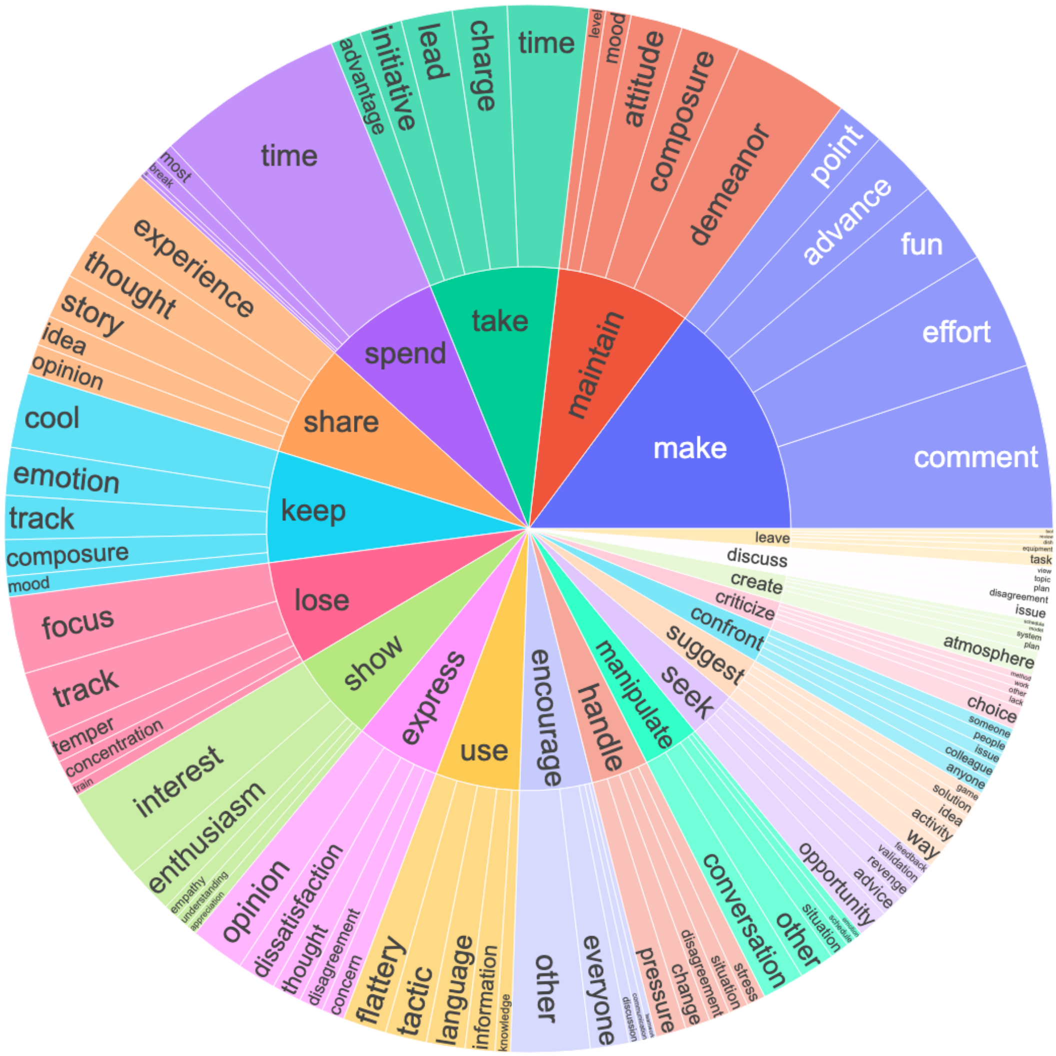

Figure 19: Word distribution in personality data. We draw the top 20 most common root verbs (inner circle) and their top 5 direct noun objects (outer circle) in the generated instructions.


Figure 20: Visualization the entire game world - Zork1


Figure 21: Visualization the entire game world - Zork2

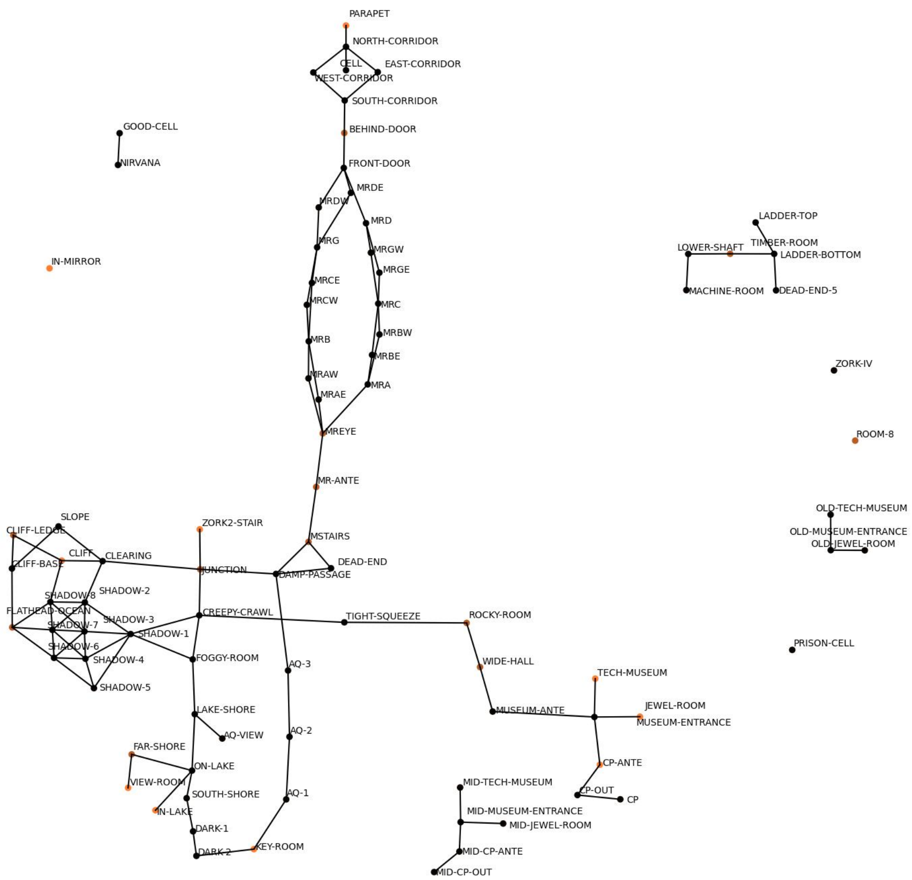


Figure 22: Visualization the entire game world - Zork3