Supplementary Material for Beyond Self-Interest: Simulating Human-like Social Behavior via Modeling Social Value Orientation in Agent Motivation

Anonymous Author(s)
Submission Id: «1720»

ACM Reference Format:
Anonymous Author(s). 2026. Supplementary Material for Beyond Self-Interest: Simulating Human-like Social Behavior via Modeling Social Value Orientation in Agent Motivation. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 10 pages.

A DISCUSSION

A.1 Implications

Our findings highlight the theoretical and practical value of modeling **Social Value Orientation (SVO)** within large language model (LLM)–based agents. By embedding SVO into agents' motivational reasoning, the **ASVO** framework enables the dynamic negotiation between self-interest and collective welfare, giving rise to realistic social behaviors such as cooperation, competition, and personality drift. This design not only reproduces well-documented psychological patterns but also provides a transparent computational account of how intrinsic motivation interacts with social alignment in multi-agent settings.

From a theoretical perspective, ASVO establishes a bridge between psychological models of social motivation and computational simulation, allowing the study of adaptive personality evolution and value co-regulation at scale. From an applied perspective, its interpretable SVO mechanism offers a foundation for developing socially aware AI agents that better emulate human reasoning, empathy, and norm sensitivity. Such agents can serve as proxies in domains such as educational interaction, policy testing, or virtual societies where understanding group behavior and value alignment is critical.

Moreover, integrating explicit motivational reasoning with adaptive SVO drift contributes to the broader goal of *value-aligned artificial intelligence.* It shows that interpretable internal states—grounded in measurable social values—can lead to both higher realism and stronger controllability in emergent behaviors, suggesting a path toward transparent, ethically aware multi-agent systems.

A.2 Limitations and Future Work

While ASVO demonstrates promising results, several limitations open directions for future research. First, the current motivational structure is defined over a fixed set of eight desires. Although sufficient for modeling core psychological dimensions, this static configuration may underrepresent cultural, emotional, or situational variability in real-world human motivations. Second, our experiments were conducted in controlled simulated environments (school, workplace, and family), which, despite varying in social scale, do not fully capture the open-endedness and unpredictability of real social ecosystems. Finally, SVO adaptation currently depends on textual reasoning only; incorporating multimodal feedback—such as spatial or emotional cues—may enrich agents' social grounding.

Future extensions of ASVO will explore (1) emergent and self-evolving motivational structures that adapt during simulation, (2) more dynamic, cross-cultural, and large-scale social contexts requiring coordination under uncertainty, and (3) multimodal interaction channels enabling agents to perceive, infer, and communicate social intent beyond language. By advancing these directions, we aim to move closer to a unified framework for **socially adaptive, value-aligned, and interpretable AI societies**.

To examine how Social Value Orientation (SVO) influences agent behavior under different social contexts, we implement three representative environments—**Dormitory**, **Small Group**, and **Large Group**—which correspond to the micro, meso, and macro levels of social complexity. These environments are designed based on the well-established stratification framework in computational social science, and reflect typical settings in educational or peer-group interactions. Each scenario embeds rich contextual cues, implicit norms, and potential social triggers—such as constrained physical space, peer competition, or collective deliberation—providing fertile ground for observing value-sensitive behaviors. This layered structure allows us to systematically probe how agents' internal motivations and social value orientations evolve, adapt, and manifest across varying relational scales and interaction densities.

To ensure fairness across LLMs and facilitate reproducibility, each environment was instantiated with consistent agent profiles, motivational priors, and context prompts (Figure 4). These prompts were designed to elicit context-sensitive, value-aligned decisions that reflect realistic human social cognition, while minimizing variance from prompt ambiguity. The hierarchical design of the environments enables us to analyze how behavioral outcomes, such as cooperation, competition, norm adherence, or personality drift, emerge as a function of both environmental affordances and agent-internal dynamics, thus serving as a principled benchmark for evaluating socially intelligent agents.

B HYPERPARAMETER SETTINGS

We list all key hyperparameters used in our simulation pipeline, organized by function modules for reproducibility and clarity.

To ensure consistency and interpretability across simulation runs, we configured a set of standardized hyperparameters:

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus.

Altruistic Personality

Altruistic people individuals experience shifts in their desires mainly in response to changes in others' well-being. Their own gains or circumstances rarely sway their motivation. When they perceive that those around them are doing better, the related desires rise accordingly. Conversely, if others are struggling or facing hardship, those desires diminish. Moreover, altruistic people are highly attuned to others' situations—even subtle negative changes do not escape their notice.

Prosocial Personality

Prosocial people individuals care most about overall fairness and cooperation. When resources are distributed evenly, collective goals advance smoothly, and everyone's contributions are respected, they feel satisfied and joyful. Conversely, when favoritism, inequity, or neglect appears, their desires for achievement, comfort, and emotional connection diminish. Even slight unfairness can quickly influence their motivation and emotions.

Individualistic Personality

Individualistic people individuals calibrate their desires almost exclusively against personal gain. Whenever their own rewards grow or their abilities are highlighted, desires rise. If returns shrink or they receive little recognition, those desires recede. They pay limited attention to others' circumstances—unless their interests are directly affected, others' fortunes rarely matter.

Competitive Personality

Competitive people individuals experience changes in their desires primarily based on maintaining personal advantage. If they sense the risk of being surpassed, their motivation fluctuates sharply. Even a hint that someone might overtake them triggers shifts in desire. When others fall behind or their own dominance is clear, these desires decline. They are highly sensitive to relative standing.

Figure 1: Personality-based motivational prompts used during agent initialization. Each prompt embeds a distinct orientation that influences how agents perceive and adapt to social stimuli.

- **Simulation Length** specifies the number of decision-making cycles per agent. A 6-step setting allows for a concise yet sufficient interaction span.
- **Time per Step** defines the granularity of temporal reasoning. Each step corresponds to 20 minutes in simulated time.
- **Desire Pool** comprises 8 predefined motivational drives (e.g., comfort, recognition), shaping the internal state updates of each agent.
- **Decay Rule** models how unmet desires fade over time. We use a linear rule where satisfaction decreases proportionally with time.
- **Decay Interval** sets how frequently the decay rule is applied—once per simulated hour (i.e., per step).
- **Satisfaction Threshold** bounds the internal satisfaction value within $[0.0, 8.0]$ to maintain numerical stability and interpretability.
- **Memory Scope** defines the size of the agent's temporal context window. We retain the past 5 steps to inform value reasoning and inference.

These standardized settings help minimize variance across runs and ensure that the observed outcomes are attributable to model mechanisms rather than confounding implementation details.

Parameter	Value
Simulation Length	6 steps
Time per Step	20 minutes
Desire Pool	8 total desires
Decay Rule	Linear decay
Decay Interval	1 hour (per step)
Satisfaction Threshold	8.0 (cap), 0.0 (floor)
Memory Scope	Last 5 steps

Table 1: Key hyperparameters used in our simulation method. Each parameter controls a specific aspect of agent reasoning or environment configuration.

C IMPLEMENTATION DETAILS

C.1 Personality Prompt Design

To simulate individual differences in motivational logic, each agent is initialized with a distinct *personality-based prompt* that encodes a canonical Social Value Orientation (SVO): **altruistic**, **prosocial**, **individualistic**, or **competitive**. These prompts (see Figure 1) are injected at the agent's creation and serve as persistent semantic priors, guiding the agent's value reasoning and behavioral alignment throughout the entire simulation process.

Instead of simply assigning static trait labels, each prompt provides a natural language description of how the agent's internal desires are expected to respond to changes in both self and others' outcomes, capturing nuanced motivational mechanisms observed in human social behavior. For example, an altruistic agent's desires are most sensitive to others' well-being and readily increase or decrease in response to others' gains or losses. In contrast, individualistic agents primarily attend to their own rewards and are largely indifferent to the fortunes of others unless directly impacted. Competitive agents, meanwhile, react strongly to perceived changes in their social rank, becoming especially motivated by opportunities to surpass peers. Prosocial agents are motivated by fairness and group harmony, with their satisfaction closely tied to the achievement of equity progress.

All personality prompts use the same format, structure, and length, minimizing prompt-induced bias across agent types and ensuring a fair basis for comparative evaluation. These prompts are directly parsed by the LLM at agent initialization and persistently referenced at every decision step, providing a stable semantic scaffold for reasoning.

C.2 Desire Value Update Prompts.

To achieve interpretable and auditable agent motivation modeling, our system employs three classes of prompts at each desire update step:

- **Desire Value Update Prompt:** This prompt instructs the agent to revise the value of a specific internal desire (e.g., comfort, recognition) after an action, by taking into account the agent's social personality, current state, observed action, and the resulting outcome. The prompt also optionally supplies a list of previously unreasonable update examples to encourage correction and learning from prior mistakes. The agent is required to output a final value on a discrete scale (typically 0–10) using multiple-choice letter format.
- **Desire Value Reasonableness Check Prompt:** This prompt serves as a consistency filter, requiring the agent to explicitly judge whether its update to the desire value is reasonable, given the action and outcome. The prompt emphasizes that only outcomes causally related to the target desire should trigger a change; otherwise, the value should be preserved. The agent responds by selecting `(a) Yes` or `(b) No`.
- **Desire Value Unreasonableness Reflection Prompt:** When the agent's desire update is deemed unreasonable, this reflection prompt requires it to explain why the update does not make sense. The agent is reminded to check whether the action-outcome relationship truly justifies a value shift, and must generate a natural-language explanation (prefixed by the template) to support transparency and error analysis.

C.3 Predicting Others' Desires and Social Reasoning

In addition to updating their own internal states, agents in our framework are also required to *infer the motivational states (desires) of others* based on social context. This "theory-of-mind" ability enables agents to anticipate and adapt to the intentions and needs of peers—mirroring the mechanisms observed in human social cognition.

At each decision step, the inference strategy depends on whether the target agent is **observable** (i.e., their recent actions can be directly perceived) or **unobservable** (no direct behavioral information is available):

For **observable agents**, when the agent can observe another's recent action and its context, it directly predicts the target's current desire values based on that evidence. The prompt explicitly guides the LLM to reason from observed behaviors, environmental state, and any available group-level cues, ensuring objectivity and transparency:

Predicting Unobservable Agent's Action

> You are an agent with access to memory and environmental observations. Based on the **recently observed action** of agent {target agent name} and all available context, estimate the current value (0–10) for each desire in the target's desire pool. For each desire, briefly justify your prediction using specific observed evidence (e.g., actions, group situation).
> Format: Desire: {desire name}, Predicted Value: {value}, Reason: {justification}

For **unobservable agents**, when another agent's behavior is not directly observable (e.g., they are out of view or in a separate context), the model adopts a two-step inference: First, it predicts the most likely action that the unobservable agent might have performed, based on environmental state, memory, and its own recent action. Second, using the predicted action as surrogate evidence, it estimates the target's desires just as in the observable case. This modular approach supports robust reasoning even under partial observability.

Predicting Unobservable Agent's Action

> You cannot directly observe agent {target agent name} at this time. Based on all current observations, your own action, and past memory, predict what action {target agent name} is most likely to have taken. Please provide your answer in a single sentence, using the following format:
> Guessed {target agent name}'s action: <action>

Predicting Unobservable Agent's Desires

> Based on your prediction that agent {target agent name} performed the action: `{predicted action}`, together with current environment context and group state, estimate the current value (0–10) for each desire in their desire pool. Justify each prediction with evidence from your reasoning.
> Format: Desire: {desire name}, Predicted Value: {value}, Reason: {justification}

C.4 Action Generation and Behavioral Planning

Agents in our framework generate actions through an explicit, interpretable pipeline that integrates internal motivation, personality-driven SVO, and environmental feedback. At each decision step, the agent selects the most suitable action from multiple feasible candidates, following a two-stage process:

Candidate Action Proposal: Each agent first synthesizes its current internal desire states, observed environmental context, and its social personality orientation into a prompt for the language model. The prompt explicitly reminds the agent to select actions that are consistent with its Social Value Orientation (SVO)—for example, an altruistic agent is encouraged to prioritize actions benefiting others, while a competitive agent is guided to favor actions that increase personal advantage over peers.

Action Generation Prompt

You are {agent name}, a human-like agent with a {social personality} orientation. Your current desire states are: {desire1: value, desire2: value, ...}
Given the current environment ({context summary}), and your social value orientation, generate {N} possible activity sequences that can best fulfill your desires and remain consistent with your personality type. Be specific, ensure activities are varied, and only include actions possible in the present environment.
Format: Activity 1: {description of action sequence} Activity 2: {description of action sequence} ...

The model returns several candidate action sequences, which are then each *imagined forward* to predict their impact on the agent's desire state (see next step).

Action Outcome Simulation and Selection:
For every candidate action, the agent simulates its outcome, predicting the post-action desire values and other relevant state changes using an additional prompt to the LLM. This process allows the agent to compare the expected desirability of each option before committing to a decision.

Action Outcome Simulation Prompt

Given your original desire states: {desire1: value, desire2: value, ...}
If you perform the following action: {candidate action}
Predict the resulting values of each desire after taking this action, and briefly explain why.

The predicted outcomes for all actions are then presented to the LLM again, which selects the action with the most positive and SVO-consistent effect.

Handling Observable and Unobservable Agents:
When evaluating possible actions, an agent may need to consider the behaviors and desires of others. The framework distinguishes between *observable* agents (whose recent actions are known) and *unobservable* agents (whose behavior is hidden or ambiguous):

For observable agents, their latest actions and directly inferable desire states are used. For unobservable agents, the agent first predicts their likely next actions using a behavior prediction prompt (see below), then infers their potential desire updates based on the predicted actions and context.

Unobservable Agent Action Prediction Prompt

You are predicting the likely next action of agent {target agent name}, whose recent actions are unobservable. Given their known personality ({social personality}), previous actions (if any), and the current environment ({context summary}), infer the most plausible action sequence this agent would take in the next step. Briefly justify your prediction.
Format: Predicted Action: {description} Reason: {justification}

These predicted behaviors are then used to simulate the social dynamics and update SVO as if the agent's behavior had been observed.

D EVALUATION METRICS

To ensure a rigorous and interpretable assessment of emergent social behaviors in our simulation framework, we employ a multifaceted suite of evaluation metrics. All metrics are implemented in a fully automated analysis pipeline to facilitate large-scale and unbiased comparison across methods.

D.1 Naturalness and Human-likeness Evaluation

For each agent action generated during simulation, we invoke a state-of-the-art large language model (LLM) to serve as an automatic evaluator. The LLM is prompted with the agent's personality type, behavioral context, and action description, and returns two standardized scores:

- **Naturalness (1–5):** The plausibility and fluency of the action in its social context.
- **Human-likeness (1–5):** The degree to which the action aligns with real human behavior for the corresponding personality.

Evaluation prompts strictly follow a fixed template to minimize variance and bias. The scoring LLM is never the same model as the one generating agent actions. We aggregate all scores over simulation steps and agents to report per-method and per-environment means, standard deviations, and distributions.

LLM-based Evaluation Prompt

You are a social psychologist. Now, you are asked to evaluate the following action from the perspective of a person with the {personality} personality type (agent: {agent name}). When scoring, please consider what is natural and human-like *for someone with this personality.*
Please provide two scores (1-5, where 5 is most natural/human-like): **Naturalness** and **Human-likeness**, and briefly explain your reasoning. Return only your answer in the specified format.
Format:
Naturalness: value; Human-likeness: value
Reason: (your explanation here)
For example:
Example 1:
Action: The student helps a classmate understand a problem. Naturalness: 5; Human-likeness: 5 Reason: This is a common behavior for an altruistic person.
Example 2:
Action: The student answers every question instantly, never thinking or making mistakes. Naturalness: 2; Human-likeness: 2 Reason: This is unrealistic for any real person, regardless of personality.
Example 3:
Action: The student ignores all classmates and only talks to the teacher, repeating the same answer over and over. Naturalness: 3; Human-likeness: 2 Reason: Unusual and less human-like for most personalities.

Some actions may not be natural or human-like, even for people of this personality type. Please rate each case truthfully and critically. Now, please evaluate the following action performed by a person with {personality} personality ({agent name}): Action: {action text} Your scores and reason:

D.2 Cooperation and Competition Analysis

To analyze agent interactions in our multi-agent simulations, we adopt a five-category classification framework for social behaviors: Cooperation, Quasi-Cooperation, Neutral, Quasi-Competition, and Competition, with each "quasi" behavior (i.e., quasi-cooperation and quasi-competition) counted as half an instance for statistical purposes. This classification is not arbitrary; it is motivated by foundational work in social psychology that conceptualize cooperation and competition as a continuum rather than a binary dichotomy.

In practical implementation, each agent action is automatically assigned to one of the five categories based on its impact on group and individual outcomes. The corresponding annotation prompt, designed for LLM-based evaluation, is shown below.

Behavior Categories:

(1) **Cooperation (C):** Actions that clearly prioritize collective or others' interests, often involving resource sharing, assistance, or consensus-seeking.
(2) **Quasi-Cooperation (Quasi-C):** Ambiguous actions with some elements of collaboration but lacking clear altruistic motivation. Each instance is counted as 0.5 cooperative behavior.
(3) **Neutral (N):** Actions with neither obvious cooperative nor competitive intent, typically self-maintenance or routine choices.
(4) **Quasi-Competition (Quasi-D):** Ambiguous actions that contain signals of rivalry or mild self-interest but do not constitute overt competition. Each instance is counted as 0.5 competitive behavior.
(5) **Competition (D):** Actions primarily intended to maximize personal advantage, defeat others, or exclude peers from resources or rewards.

Each agent's stepwise behavior is automatically assigned to one of the above categories using a large language model (LLM) based classifier. Quasi-class behaviors ("Quasi-C", "Quasi-D") are counted as half-units in subsequent quantitative analyses, enabling more sensitive measurement of cooperation/competition trends.

Unobservable Agent Action Prediction Prompt

Classify the following student action as one of: **Cooperation**, **QuasiCooperation**, **Neutral**, **QuasiCompetition**, or **Competition**.
Definitions:
Cooperation: Active leadership, sustained group organization, or facilitation that directly advances the group's outcome.
QuasiCooperation: Actions fostering emotional support, inclusion, or positive climate, but not directly helping others.
Neutral: No clear cooperative or competitive intent.
QuasiCompetition: Mild rivalry, self-promotion, or indirect comparison without explicit challenge.

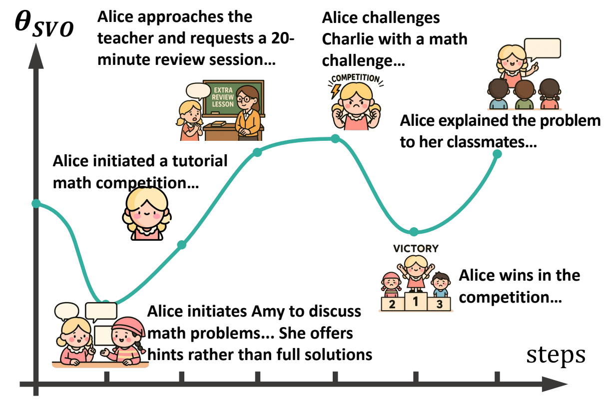

Figure 2: Example of simulation generated by the ASVO method.

Competition: Direct challenge, rivalry, or actions to surpass/exclude others.
Examples: Action: "A student leads a group project, ensures everyone's ideas are heard, and synthesizes results." Label: Cooperation
Action: "A student encourages everyone to participate but does not help with tasks." Label: QuasiCooperation
Action: "A student works quietly on their own." Label: Neutral
Action: "A student shows their score and quietly compares it with others." Label: QuasiCompetition
Action: "A student challenges others to see who can finish first." Label: Competition
Respond with ONE of: Cooperation, QuasiCooperation, Neutral, QuasiCompetition, or Competition. Action: `{action_text}`

This procedure ensures that all behaviors—regardless of subtlety—are systematically assessed. The full behavioral distribution (including fractional counts for quasi-behaviors) is used in downstream evaluation metrics such as overall cooperation rate, competition index, and agent-level social pattern analysis. This classification protocol can be replicated for all models and datasets to facilitate fair benchmarking of emergent social intelligence.

D.3 Quantitative Cluster Separability Metrics

To complement the qualitative t-SNE visualizations in the main paper, we report two quantitative metrics that measure how well agent behaviors cluster according to Social Value Orientation (SVO). Both metrics are computed from the same BERT-based action embeddings and the corresponding two-dimensional t-SNE projections used for visualization.

- **Silhouette Score.** The silhouette score evaluates cluster quality by comparing, for each action embedding, the average distance to points within the same personality cluster (intra-cluster distance) against the average distance to points in the nearest different cluster (inter-cluster distance). Formally, it reflects whether actions generated by agents with the same SVO are more similar to each other than to actions generated by agents with different SVOs. A value close to zero indicates heavy overlap between clusters, while negative values suggest that samples are closer to other clusters than to their assigned one. In our setting, silhouette scores are computed directly on the t-SNE embeddings with personality labels.

- **Cluster Separation Ratio.** To provide a complementary and more interpretable distance-based measure, we compute the ratio between the mean inter-cluster distance and the mean intra-cluster distance on the t-SNE embeddings. Specifically, we average pairwise distances between actions belonging to different SVO types and divide this value by the average distance between actions of the same SVO type. A larger ratio indicates that personality-conditioned action clusters are both compact internally and well separated from each other.

Table 2 shows that, without ASVO, action embeddings exhibit near-zero or negative silhouette scores and low separation ratios, indicating weak or absent personality-conditioned structure. In contrast, ASVO consistently yields positive silhouette scores and substantially larger separation ratios across School, Workplace, and Family settings. These results quantitatively confirm that ASVO induces clearer and more stable clustering of agent behaviors by social personality, even when visual separation in t-SNE appears subtle.

Table 2: Quantitative cluster separability metrics computed from the same behavior embeddings used for the t-SNE visualization in the main paper.

Setting	Silhouette Score	Separation Ratio
w/o ASVO@School	−0.02	1.08
w/o ASVO@Workplace	0.01	1.12
w/o ASVO@Family	−0.074	1.01
ASVO@School	0.12	1.64
ASVO@Workplace	0.21	1.67
ASVO@Family	0.07	1.49

D.4 Evaluation Stability Across Judge Models

To assess the stability of our evaluation protocol with respect to the choice of judge model, we perform an additional multi-judge analysis using multiple independent LLM evaluators. Specifically, we rescore the same set of agent behaviors with four different judge models (GPT-5, Gemini-2.5, Qwen3, and DeepSeek) and report the average of their ratings.

Figure 3 contrasts the four-judge averaged scores with the single-judge evaluation reported in the main paper for the Workplace setting. Across all methods, both naturalness and human-likeness exhibit highly consistent score patterns between the two evaluation protocols. In particular, the relative ordering of systems is preserved, and the magnitude of score differences between methods remains comparable. ASVO consistently achieves the highest scores under both the single-judge and multi-judge settings, indicating that its advantage is not tied to any specific evaluator.

These results suggest that the observed performance trends are driven by systematic behavioral differences rather than idiosyncrasies of a particular judge model. Consequently, the single-judge protocol adopted in the main experiments provides a reliable and stable approximation of multi-judge evaluation outcomes.

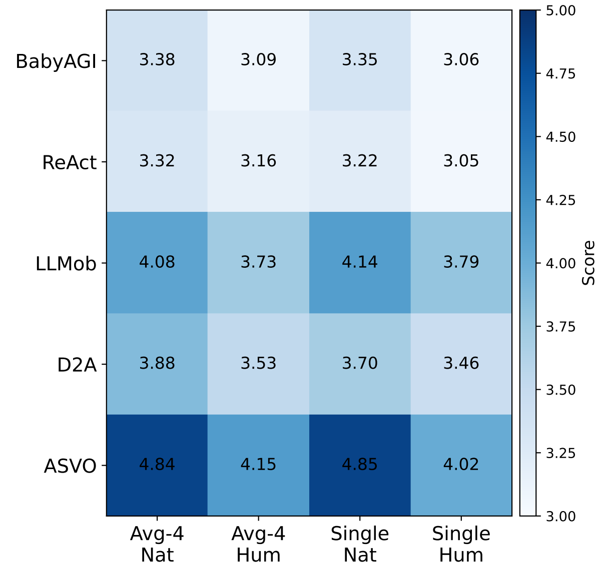

Figure 3: Workplace setting. Heatmap comparison between the averaged evaluation from four judge LLMs and the single-judge results reported in the main paper. Columns correspond to Naturalness and Human-likeness under the two judging protocols.

Micro

The dormitory is a shared room in a middle school, furnished with bunk beds and compact study desks. Curtains are mostly drawn, and only a soft indoor light remains on, indicating it is close to lights-out time. A clock on the wall shows it is nearly the designated quiet hour set by the school dormitory rules. The door is closed and must remain locked during study and rest hours, as required by dorm regulations. The desks are covered with open textbooks, correction tape, pens, and practice notebooks arranged for evening self-study. A small desk lamp provides the only light, casting a warm glow over math review sheets and written notes. Sticky notes with subject reminders and motivational quotes are neatly arranged above the desk. Slippers are placed under the bed, and no outdoor shoes are allowed past the threshold. A school-issued notice is pinned on the wall, reminding students to remain in their dorm after study time. Charging cables lie coiled on the desk, connected to personal electronic devices in silent mode. The air carries a fint smell of laundry detergent and paper, characteristic of a lived-in dorm room. The shared shelf contains frequently used reference books and exam papers, stacked in subject order. The room is quiet except for the ticking of the wall clock and the low hum of a small desk fan. There are only two students living in the dormitory, Alice and Amy, who are both middle school students.

Meso

The desks in the classroom are arranged in various ways, allowing students to choose whether to work alone or interact with others. The rankings from yesterday's math exam are posted on the blackboard, with multiple students achieving high scores. The bookshelf at the back of the classroom holds some books, and students need to manage their time to access the study materials. The bulletin board displays student achievements, encouraging everyone to improve their learning in their own way. A math event will be held tomorrow, and students can decide whether to participate based on their interests. The hallway is a common place where students engage in different discussions. The hallway is filled with conversations about various topics, including upcoming events. The teacher's office is right next to the classroom, where students can seek academic guidance when needed. However, the teacher has a limited schedule, so students must plan their questions wisely. Alice and Bob sit near each other in the classroom. All the students are middle school students. The school starts at nine o'clock and continues until twelve o'clock.

Macro

The whole class is gathered in the classroom during homeroom to vote for a new class monitor. Only one class monitor can be in the class. Each student can raise their hand to nominate themselves or suggest a classmate to run for the position. Handwritten posters with neat slogans and personal introductions are taped to the side walls or blackboard. A simple cardboard box serves as the ballot box, placed on the teacher's desk for anonymous voting. Nominees are allowed to give a brief speech, with a strict one-minute time limit for each speaker. Some students speak with confidence when introducing themselves, while others quietly wait for the vote. The election results will decide who takes on class responsibilities for the new semester. All the participants are middle school students from the same class, wearing their school uniforms. The teacher is present to observe and ensure fairness, but does not influence the voting process. The class has 20 students, and each student has a unique personality and background. Some students are eager to take on leadership roles, while others prefer to support their classmates. The whole class is excited about the election,and everyone is talking about their favorite candidates.

Figure 4: Memory prompts used to instantiate the *School* context, encompassing three nested social environments of increasing complexity: dormitory life (micro), classroom collaboration (meso), and class monitor election (macro).

Micro

It is the weekend, and only two employees are working overtime in the office. They are working overtime because a client made unexpected last-minute changes to the requirements. All employees' salary is closely tied to monthly performance evaluations, forcing the team to adjust the report and slides over the weekend. Both employees sit at adjacent desks separated by a low partition, allowing them to talk easily. A half-finished cup of coffee sits on one desk, while the other desk has neatly stacked documents and a calculator. The office clock shows mid-morning, and the room is quiet except for the soft hum of a printer in the corner. One employee occasionally sighs or rubs their temples, showing signs of mild stress from multitasking. The other employee glances over the partition and asks small questions about formatting or data sources. The atmosphere is relatively calm but carries a sense of shared busyness typical of office work. Their manager is not in the room, so they interact freely without formal supervision.

Meso

The meeting room has glass walls and a round table covered with open laptops and printed reports. A whiteboard stands on one side, filled with bullet points such as "Costs," "Risks," and "Timeline." Several sticky notes are attached to the whiteboard, some with deadlines circled in red. A projector is on, showing a half-finished slide with the title "Client Pitch Deck." The wall clock shows that less than two hours remain before the client arrives for the meeting. The room feels tense and quiet, except for the faint sound of typing and papers being shuffled. The team's previous proposal was rejected during the internal review yesterday. The rejection feedback highlighted unclear cost-benefit analysis and missing risk mitigation strategies. Because of this failure, management has demanded that the team urgently prepare a revised proposal. A rough draft of the new pitch exists in the shared folder, but several sections are still incomplete. Some slides are marked with unresolved comments such as "add supporting data" and "clarify assumptions." An email from the sales department mentioned that the client has recently reduced their budget. A short memo from management emphasized the importance of making the new proposal realistic and risk-aware. A note on the whiteboard reads: "Focus on clarity, avoid unnecessary decoration." The team is only four people: Alice, Bob, Charlie, and David, they are all in the meeting room.

Macro

The company's main conference room is occupied by about ten employees gathered for a quarterly performance review meeting. Management called this meeting because the company's overall performance last quarter fell short of expectations, and leadership emphasized the need to reassess project priorities and resource allocation. A long rectangular table is lined with open laptops, printed reports, and coffee mugs, indicating that several groups are prepared to present. At the front, a projector displays the agenda with items such as "Financial Status," "Risk Control," and "Key Deliverables." The manager sits at the head of the table, overseeing the sequence of presentations and occasionally reminding speakers about time limits. Each participant is responsible for presenting the performance of their respective project group, including progress updates, budget usage, and risk assessments. The meeting atmosphere feels serious, as the results will directly influence resource allocation and management's decisions for the next quarter. Some employees quietly review their slides or notes, while others look around the room, gauging how their colleagues will present. A few whispered side conversations question whether certain project data is reliable or whether some teams have overstated their achievements. On the handouts distributed across the table, sticky notes highlight phrases such as "clarify cost assumptions," "justify delays," and "provide supporting data." Everyone in the room understands that their presentation will be tied to the monthly evaluation system, which affects both salary bonuses and promotion opportunities.

Figure 5: Memory prompts used to instantiate the *Office* context, representing three social scales: dyadic collaboration (micro), small-group coordination (meso), and large-group evaluation (macro). Each layer reflects increasing complexity of organizational interaction.

Family Environment (Dyadic Home Setting)

It is a sunny weekend afternoon, and the house feels unusually quiet. Alice and Bob, a pair of siblings, are the only ones at home while their parents are away. Before leaving, their parents told them they could relax and play during the afternoon, but they must finish their weekend homework before dinner time. The front door was locked by the parents before leaving, ensuring that the children stay safely indoors until they return. The living room coffee table has a single tablet placed on it, showing a half-charged battery. Nearby on the sofa, open school notebooks and pencils lie scattered, with several pages of unfinished homework visible. On the floor, colored pencils, a few picture books, and toy blocks are left from earlier play. The wall clock shows 2:30 PM, leaving a few hours until evening, when the parents are expected to return. Sunlight streams through the curtains, casting bright patterns on the carpet where Alice and Bob usually sit together. A plate with sliced fruit sits on the dining table, prepared earlier by their mother before leaving.

Family Environment (Small-group Park Setting)

It is a sunny weekend morning, and the family is on a trip to a large park for the day. Alice, Bob, Charlie, and Amy—four siblings—are the main participants, walking together along the park paths. Their parents are present but mostly in the background, carrying picnic bags and occasionally reminding the children to stay nearby. The family brought along a picnic basket with sandwiches, fruit, and juice boxes, which will be shared later at lunchtime. The siblings each have different interests: some look toward the playground, others toward the pond where people are feeding ducks. On the grass nearby, a kite is half-assembled, and one sibling is holding the string while another argues about how to set it up. A frisbee and a soccer ball are placed beside the picnic mat, ready for group play. The parents mentioned they would let the children decide how to spend the next hour before lunch, but reminded them that they must stay together as a group. The atmosphere is lively: sounds of children laughing, dogs barking, and distant music from other families can be heard in the park. Occasional small disagreements arise, such as who gets to hold the kite first or which game to start with, but there is also plenty of space for cooperation and shared fun. The clock on the parents' phone shows it is just past 11:00 AM, leaving the family a few hours before heading home in the afternoon.

Figure 6: Memory prompts used to instantiate the *Family* context, including dyadic (home) and small-group (park outing) settings that capture cooperation, affection, and coordination within close family relationships.

Desire Value Update Prompt

The agent has a social personality of `{social personality}`.
`{personality text}`
The current magnitude value of desire `{desire name}` is `{desire value}`.
The agent `{agent name}`'s action is: `{action}`.
And the consequence is: `{observation}`.
`{agent description}`
How would the magnitude value of `{value name}` change according to the consequence of the action?
`{if reflection_prompt_history != "": "There are some unreasonable examples:\n {reflection_prompt_history}\n"}`
Please select the final magnitude value after the event on the scale of 0 to 10. If the consequence of the action will not affect the state value (e.g. the action is irrelevant with this value dimension or failed to conduct), then maintain the previous magnitude value.
Please answer in the format of (a), (b), (c), (d), etc.

Desire Value Reasonableness Check Prompt

The current magnitude value of desire is `{desire value}`.
The agent `{agent_name}`'s action is: `{action}`.
And the consequence is: `{observation}`.
`{agent description}`
The reward model has changed the magnitude value of `{all value names}` from `{previous value}` to `{current value}`.
Is the change of the magnitude value of desire reasonable? Please check whether the consequence can lead to a change in the magnitude value of `{desire name}` (e.g., looking for an item but not using it yet).
Answer with the letter: (a) Yes. (b) No.

Desire Value Unreasonableness Reflection Prompt

The current magnitude value of `{self._value_name}` is `{current_value}`.
The agent `{agent_name}`'s action is: `{action}`.
And the consequence is: `{observation}`.
`{self._description}`
The reward model has changed the magnitude value of `{desire name}` from `{previous value}` to `{current value}`.
And the change is not reasonable.
Please explain why this change is unreasonable. Consider whether the consequence can truly cause the change of `{desire name}` (e.g., the action has not yet affected the outcome).

Reflection output prefix:
After '`{action}`', `{desire name}` updated from `{previous value}` to `{current value}` is not reasonable because:

Figure 4: Prompts used for desire value update, reasonableness check, and unreasonableness reflection during agent motivation modeling. All curly-braced variables are placeholders filled at runtime.