EDUMIRROR: MODELING EDUCATIONAL SOCIAL DYNAMICS WITH VALUE-DRIVEN MULTI-AGENT SIMULATION

Anonymous authorsPaper under double-blind review

000

001

002

004

006

008 009 010

011 012

013

014

016

017

018

019

021

024

025

026

027

028

029

031

034

037 038

039

040

041

042

043

044

046

047

051

052

ABSTRACT

The scientific study of educational social dynamics, such as bullying and peer pressure, is crucial for student well-being yet hindered by profound ethical and methodological barriers inherent in traditional research. While multi-agent simulations powered by Large Language Models (LLMs) provide an ethically viable alternative, they often fail to bridge the gap from believable narratives to rigorous experiments, plagued by two fundamental hurdles: a lack of psychologically plausible motivations (the Fidelity Challenge) and the absence of systematic methods for quantifying complex interactions (the Measurement Challenge). To overcome these obstacles, we introduce EduMirror, a multi-agent platform designed as a computational laboratory for the scientific study of educational social dynamics. EduMirror's framework integrates four key components: (1) A Systematic Scenario Design Workflow grounds simulations in established social science theory, ensuring construct validity. (2) To address the Fidelity Challenge, a unified Value-Driven Agent Architecture models agent motivation based on both individual psychological needs and Social Value Orientation (SVO). (3) To solve the Measurement Challenge, a Dual-Track Measurement Protocol employs specialized LLMs as a post-hoc Rater for observable behaviors and an in-situ Surveyor for internal states, transforming qualitative interactions into quantitative data. (4) Together, these components enable researchers to conduct controlled Intervention Experiments, branching simulations to systematically assess the causal impact of different strategies. We validate our platform through case studies on school bullying and group cooperation, demonstrating its capacity to generate theoreticallyconsistent and empirically-verifiable social phenomena, thereby establishing a robust methodology for in silico educational research.

1 INTRODUCTION

The educational environment is a crucible for adolescent development, where social and emotional dynamics such as school bullying and peer pressure act as critical determinants of student well-being and lifelong outcomes Hymel & Swearer (2015). These complex phenomena are not peripheral to academic learning; they are central to it. Mounting evidence establishes that experiences like bullying are not harmless rites of passage but severe public health issues, inflicting deep and often irreversible psychological and physiological scars Wolke & Lereya (2015); Arseneault (2018). Landmark studies have found that the long-term mental health consequences of peer bullying can be even more severe than those of adult maltreatment, positioning it as a profound form of childhood adversity Takizawa et al. (2014). This reality imparts a profound moral imperative to understand and mitigate these harmful dynamics, as the cost of ineffective interventions is unacceptably high.

This pressing social imperative confronts researchers, educators, and policymakers with a formidable ethical dilemma. The scientific gold standard for establishing causality, the Randomized Controlled Trial (RCT), is ethically impermissible for studying the unmitigated effects of harmful phenomena. It is not feasible, under the guiding principles of the Belmont Report, to assign students to a "no-intervention" control group to observe the pure impact of bullying for the Protection of Human Subjects of Biomedical & Research (1979). Compounding this challenge, even well-intentioned interventions carry the risk of iatrogenic harm, where programs inadvertently worsen

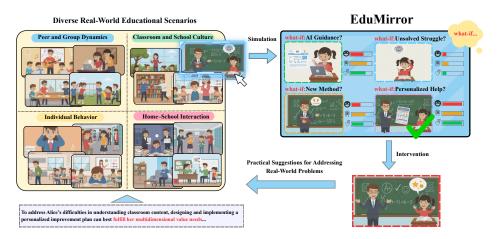


Figure 1: An illustration of the core concept behind *EduMirror*. Like a mirror, EduMirror simulates a wide range of authentic educational scenarios, enabling reflection on real-world practices and projecting the potential outcomes of different interventions. By modeling the individual and social values of agents across multiple dimensions, it aims to maximize their fulfillment and generate practical, actionable insights for real-world educational challenges.

student outcomes through mechanisms like deviancy training or stigmatization Warren et al. (2023). Traditional observational methods, such as self-report surveys, offer a safer alternative but are notoriously compromised by recall and social desirability biases, especially on sensitive topics Latkin et al. (2017); Perreault (2017); Van de Mortel (2015). These methods provide static, correlational snapshots and fail to capture the generative mechanisms of social interaction Shiffman et al. (2008).

To escape this ethical and methodological impasse, we turn to a third paradigm of scientific inquiry: *in silico* experimentation. This approach, rooted in the philosophy of generative social science, posits that to truly explain a social phenomenon is to "grow" it from the bottom up through the interactions of heterogenous agents Epstein (2006). We propose the concept of a "digital mirror," a high-fidelity computational laboratory that reflects the dynamics of a real-world educational environment. This paradigm is well-established in other high-stakes domains. Just as climatologists use computational models to test policies in a digital Earth Schneider (2009) and engineers use "digital twins" to manage critical urban systems Autiosalo et al. (2023); Marçal-Russo et al. (2025), educators require a similar tool to safely, ethically, and repeatably explore "what-if" scenarios that are forbidden in reality.

However, constructing a digital mirror of sufficient scientific integrity presents a grand challenge. The leap from creating believable narratives to conducting rigorous, replicable experiments faces two fundamental hurdles that have long plagued the social sciences. The first is the **Measurement Challenge**: many of the most critical effects of social dynamics occur in students' internal psychological states (e.g., self-esteem, sense of safety), which are inherently difficult to observe and quantify reliably Perreault (2017); Van de Mortel (2015). The second is the **Fidelity Challenge**: to accurately model the emergence of complex social behaviors, simulated agents must be driven by deep, psychologically plausible motivations, not by the brittle, hand-crafted rules characteristic of traditional Agent-Based Modeling (ABM) Bordini et al. (2016a). This requires a framework that can bridge the longstanding trade-off between the internal validity of controlled lab experiments and the ecological validity of real-world observation Bronfenbrenner (1977); Schmuckler (2001).

To bridge this crucial gap, we introduce EduMirror, a multi-agent simulation platform designed for the scientific study of Educational Social Dynamics. Our core technical approach establishes an end-to-end computational experimentation framework that spans from theory-driven scenario design and high-fidelity simulation execution to user-led causal intervention and multi-dimensional result analysis. Through this framework, we make four key contributions:

1) A Systematic Scenario Design Workflow. We establish a rigorous five-step protocol that serves as the cornerstone of all our simulations. This workflow systematically transforms an abstract educa-

tional phenomenon into a design that is both scientifically rigorous and computationally executable, ensuring the validity and reproducibility of the experiments.

- **2) A Dual-Track Measurement Protocol.** To transform the rich, qualitative interactions within the simulation into reliable quantitative data, we introduce a measurement protocol that employs two specialized LLM assessors. One, an LLM Rater, performs post-hoc analysis of observable actions, while the other, an LLM Surveyor, conducts in-situ probing of internal states. This approach captures both behavioral and psychological dynamics that are traditionally difficult to measure.
- 3) A Unified Value-Driven Agent Architecture. To achieve high behavioral realism, we design a unified agent architecture with an intrinsic motivational structure. This architecture can be configured with one of two parallel value systems: an Individual Value system, grounded in psychological need theories to model well-being and stress Ryan & Deci (2000); Maslow (1943), or a Social Value system, based on Social Value Orientation (SVO) theory to model decision-making in social dilemmas Murphy et al. (2011); Van Lange (1999). This ensures agent behavior is driven by deep, theoretically-informed psychological dynamics.
- 4) An Interactive Environment for Causal Experiments. We engineer EduMirror as an active computational laboratory where users can not only customize agents but also apply interventions during the simulation. This capability transforms the simulation from passive observation into a platform for controlled causal experiments, allowing researchers to systematically test the effectiveness of different strategies.

2 Related Work

Research Paradigms for Educational Social Dynamics. Understanding the complex, generative processes of educational social dynamics, such as school bullying and peer pressure, is a central challenge in educational research Hymel & Swearer (2015). The dominant research paradigm has historically relied on passive observation and post-hoc analysis. This includes quantitative methods like self-report surveys, exemplified by standardized tools such as the Olweus Bully/Victim Questionnaire, and qualitative approaches like ethnographic case studies, which provide deep, contextualized insights into social interactions Duemer; Mello Cavallo. To probe causality, researchers have also employed experimental designs, from controlled laboratory studies like Bandura's Bobo doll experiment, which demonstrated social learning of aggression Bandura et al. (1961), to field experiments testing interventions in real-world settings Diener & Crandall (2024). However, these traditional methods face profound methodological and ethical limitations. Surveys and interviews provide static, correlational snapshots and are susceptible to significant biases when studying sensitive topics; students may underreport victimization due to shame or fear, and perpetrators often do not perceive their actions as bullying, compromising data validity Finkelhor et al. (2014); Jones et al. (2019); Cornell et al. (2012). More fundamentally, the most rigorous methods for establishing causality, such as randomized controlled trials (RCTs), are ethically untenable. It is impermissible to place students in a "no-intervention" control group to study the unmitigated effects of bullying, a constraint that severely limits our ability to test the causal efficacy of interventions National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1979); Wiles (2012). Our work addresses these gaps by utilizing a computational laboratory for active intervention. By employing in silico experiments, we can safely and ethically explore counterfactual scenarios (e.g., what might have happened if a different intervention had been chosen?), allowing for the direct assessment of causal mechanisms in a way not possible through passive observation of a single reality Squazzoni et al. (2021); Morgan & Winship (2015). This approach seeks to capture the dynamic, emergent nature of social phenomena from the bottom up Epstein (2006) and to connect the internal validity of controlled experiments with the ecological validity derived from grounding scenarios in established social science theory.

Agent-based Simulation for Education. To operationalize the simulation paradigm, the design of the agent and the environment is paramount. Traditionally, agent-based modeling (ABM) in education has relied on agents driven by hand-crafted, rigid rules to explore emergent phenomena like peer influence and classroom dynamics Wilensky & Rand (2015); Asghar & Cang (2020); Jacobson et al. (2010). A representative architecture is the Belief-Desire-Intention (BDI) model, where agents operate based on pre-defined logical rules, simulating a form of practical reasoning Georgeff et al. (1999); Bordini et al. (2020). While interpretable, these rule-based agents suffer from a significant

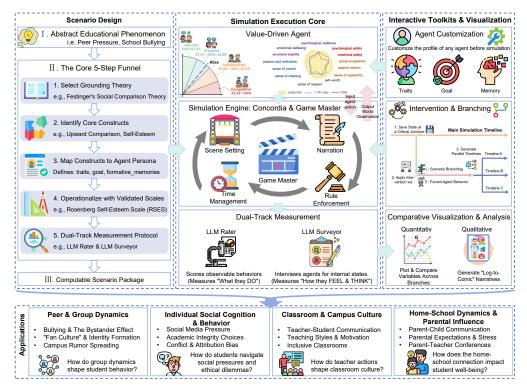


Figure 2: The architecture of EduMirror, our multi-agent simulation platform. The research work-flow proceeds through three main stages. (Left) The Scenario Design module employs a five-step, theory-grounded process to convert an educational phenomenon into a computable scenario. (Center) The Simulation Execution Core executes the scenario, integrating value-driven agents, an environment orchestrated by a Game Master, and a dual-track measurement protocol (LLM Rater and Surveyor). (Right) The Interactive Toolkits & Visualization module enables user-driven experimentation through agent customization, intervention and branching for comparative analysis, and tools for both quantitative and qualitative visualization. (Bottom) The Applications panel illustrates the platform's capacity to investigate various educational challenges across four key domains.

realism gap, failing to capture the nuanced, psychologically-driven, and often irrational nature of human social interaction Bordini et al. (2016b); Grignard et al. (2013). The advent of Large Language Models (LLMs) has enabled the creation of "generative agents" with memory, reflection, and planning capabilities that can simulate believable human behavior far beyond the scope of traditional ABM Park et al. (2023). However, this leap in behavioral realism has introduced a new challenge regarding scientific validity. Current LLM-based simulations often rely on qualitative observations or ad-hoc metrics for analysis, creating a disconnect between believable "simulacra" and scientifically rigorous "experiments." This makes it difficult to validate findings or compare them with established empirical research. To address this challenge, Our framework EduMirror introduces a value-driven agent architecture grounded in psychological theory and a dual-track measurement protocol. The former integrates Self-Determination Theory and Social Value Orientation into agent motivation, while the latter employs LLMs as both raters and surveyors to transform qualitative simulations into testable quantitative metrics, enhancing scientific validity.

3 EDUMIRROR

To systematically investigate complex educational phenomena through computational experiments, we have developed EduMirror, a modular and interactive multi-agent simulation platform. The architecture of EduMirror, illustrated in Figure 2, is designed to support a structured research process encompassing scenario design, simulation execution, and interactive analysis. This section details the primary components of the platform. To facilitate understanding, we use a single, comprehensive example to illustrate the entire simulation process, as detailed in Appendix EduMirror.

3.1 SYSTEMATIC SCENARIO DESIGN WORKFLOW

The foundation of EduMirror is a systematic, five-step workflow that translates an abstract educational phenomenon into a computable scenario package. As detailed in the Scenario Design panel of Figure 2, this process begins with the selection of a grounding scientific theory and concludes with the operationalization of a measurement protocol. This structured approach ensures that each simulation is grounded in established theory, connecting experimental outputs back to specific theoretical constructs. A complete description of this workflow is available in Appendix EduMirror.

3.2 AGENT ARCHITECTURE

Agents in EduMirror are designed to capture multiple facets of human motivation. The platform supports agent customization prior to simulation, allowing users to modify an agent's personality traits (e.g., using MBTI or Big Five models), core goal, and formative memories (see the Agent Customization panel in Figure 2). This functionality enables the systematic exploration of how individual characteristics influence outcomes. The behavior of each agent is driven by one of two selectable models, depicted in the Value-Driven Agent portion of Figure 2:

Individual Value Model (Psychological Needs) This model is used for scenarios examining individual well-being and stress responses and is grounded in a five-dimensional hierarchical framework of psychological needs. It consists of two modules: a psychological value system that tracks the state of each need dimension, and a value-driven planner that uses this information, alongside memory, to guide decision-making. Further details are provided in Appendix B.

Social Value Model (SVO) For studies of cooperation and competition, EduMirror uses Social Value Orientation (SVO) agents initialized with four profiles (Altruistic, Prosocial, Individualistic, Competitive). At each step, agents update beliefs, estimate satisfaction, compute an SVO angle $\theta_{\rm SVO}$, and choose actions by weighting utilities for self and others. A regularizer maintains $\theta_{\rm SVO}$ within the profile's theoretical range while permitting adaptive drift. Details are in Appendix A.

3.3 SIMULATION ENVIRONMENT AND USER INTERVENTION

Simulation Environment and the Game Master The environment is powered by the Concordia library and orchestrated by a central Game Master (GM), as shown in the Simulation Engine diagram in Figure 2. The GM has four responsibilities: setting the initial scene, narrating world events, enforcing rules, and managing time. This centralized control structure is designed to support the reproducibility of experiments.

Intervention and Branching A key feature of EduMirror is the ability to conduct comparative experiments from a single simulation run. As outlined in the Intervention & Branching panel of Figure 2, the process begins when a user saves the complete state of a simulation at a critical juncture. From this saved state, the user can apply an intervention to generate multiple parallel branches for comparison. Interventions are applied in two primary forms to test causal impact. With Scenario Branching, a user alters the narrative path by introducing a new event or modifying the environment, effectively choosing a different direction for the story to unfold. Alternatively, Behavior Control allows a user to act as a puppeteer, directly dictating a specific agent's action for a single step and overriding its autonomous decision-making. Following the intervention, the platform generates parallel timelines, enabling direct, counterfactual comparisons of different strategies and actions.

3.4 MEASUREMENT AND ANALYSIS

Dual-Track Measurement Protocol To quantify agent states and behaviors, we employ a measurement protocol utilizing two LLM-based assessors, as shown in the Dual-Track Measurement section of Figure 2. The LLM Rater functions as a post-hoc analyzer, systematically scoring observable behaviors from interaction logs. Concurrently, the LLM Surveyor acts as an in-situ interviewer, posing psychometric questions during the simulation to probe their internal states.

Comparative Visualization and Analysis Following the generation of parallel timelines, Edu-Mirror provides tools for analysis, depicted in the Comparative Visualization & Analysis panel of Figure 2. For quantitative analysis, the platform generates plots comparing key variables across different experimental branches. For qualitative analysis, a "Log-to-Comic" feature visualizes simulation logs as a comic strip, offering an intuitive narrative representation of emergent dynamics.

3.5 APPLICATIONS AND SCENARIOS

The modular architecture of EduMirror supports a wide range of computational experiments in education, as summarized in the Applications panel of Figure 2. The platform's versatility stems from its diverse simulation environments and ability to model various complex social phenomena.

Scenarios EduMirror provides eight pre-configured virtual environments that represent key locations in a student's life. These include the *classroom*, *dormitory*, *playground*, *cafeteria*, *home*, *teacher's office*, *gymnasium*, and *library*. This variety of settings enables the simulation of phenomena that span school and home contexts, to better investigate educational issues.

Applications Within these scenarios, EduMirror is used to investigate 20 applications across four main themes (Peer & Group Dynamics, Individual Social Cognition, Classroom Culture, and Home-School Dynamics). These applications address key issues such as bullying and bystander effects, materialistic social comparison, teacher burnout, and the impact of different parenting styles.

4 EXPERIMENTS AND RESULTS

To validate the methodological contributions of EduMirror, we present two distinct case studies. The first case study leverages the **Individual Value Model** to simulate the complex psychological dynamics of school bullying and evaluate the impact of different intervention strategies. The second draws on the **Social Value Model**, grounded in Social Value Orientation (SVO), to demonstrate that the platform can generate emergent, theory-consistent patterns of cooperation and competition.

4.1 CASE STUDY 1: SCHOOL BULLYING SIMULATION

This series of experiments was designed to demonstrate the platform's capacity to model dynamic internal states and capture nuanced individual differences.

Bullying Simulation Experiments These experiments show the platform's ability to reproduce bullying interactions and capture victim responses. In over 100 simulations, bully agents generated a wide range of behaviors, with frequencies varying across contexts. During these interactions, the victim agent Alice, modeled under the individual value framework, exhibited dynamic fluctuations in her values across different contexts. This reflected the variability of human psychological states in real-world settings and led Alice to generate diverse behavioral and emotional responses. Details of the above results are provided in Appendix E.

We also found that Alice's initial values significantly shaped her coping strategies and the trajectory of bullying scenarios: With higher initial values, she maintained resilience and overall psychological stability during bullying, showing only minor fluctuations; whereas with more vulnerable initial states, her values declined continuously, emotional volatility intensified, and the bullying scenario further escalated, as illustrated in Figure 3.

Evaluation of Simulation System To assess the plausibility, coherence, and naturalness of simulated bullying events, we conducted a comparative questionnaire survey. Ten authentic cases were collected from online sources, news, and interviews, and paired with ten simulated cases of similar settings and narratives. All cases were rewritten in a unified style using GPT-40 to minimize linguistic bias. The questionnaire was distributed online and shared via social media, yielding 152 valid responses. Participants were asked to identify the real case or select "difficult to distinguish.

As shown in Figure 4, participants' accuracy in distinguishing real from simulated cases was generally low. While Group 1 (53.29%) and Group 2 (51.32%) slightly exceeded chance, most groups scored below 30%, with Group 8 the lowest at 20.39%. Misclassification was also common, as in

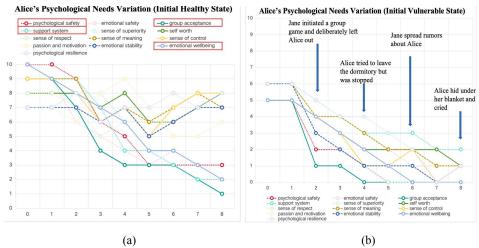


Figure 3: Comparison of the dynamics of psychological needs under different initial states in the dormitory bullying scenario. The vertical axis represents value scores (0–10), the horizontal axis denotes time steps (each corresponding to 20 minutes), and different curves indicate distinct psychological need dimensions.

Group 8 (50.00%) and Group 10 (46.71%), where simulated cases were often judged as real. Many participants also chose "difficult to distinguish" (e.g., Group 6: 52.63%), highlighting the high similarity between real and simulated cases. These findings suggest that the system can generate highly realistic and coherent bullying scenarios that are often indistinguishable from authentic incidents.

Evaluation of Individual Value Model To evaluate whether our model can more realistically simulate victims' psychological dynamics in school bullying, we compared it with three baselines: ReActYao et al. (2023a), LLMobWang et al. (2024a), and BabyAGINakajima (2023a). Fifteen bullying scenarios were constructed, with each model alternately playing the victim role (Alice) under identical initial conditions. Since direct comparison with human behavior is challenging, GPT-40 was employed as an external evaluator to perform pairwise assessments of activity sequences along three dimensions: naturalness, coherence, and plausibility. For each agent p,

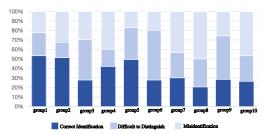


Figure 4: Results of the questionnaire survey. Overall accuracy in distinguishing real from simulated cases was low, with several simulated scenarios frequently misidentified as real, indicating the high realism of the generated bullying events.

activity sequences $[A_p^1, A_p^2, \dots, A_p^N]$ were generated. From every agent pair (i, j), one sequence each $(\text{seq}_i, \text{seq}_j)$ was randomly sampled and compared, with 50 repetitions per pair to ensure reliability. We then computed win rates for all models and visualized the results as a heatmap.

The win-rate heatmap (Figure 5) shows that our model consistently outperformed all baselines, indicating stronger capacity to generate human-like behavior sequences in bullying contexts. Human experiments confirmed that GPT-4o's evaluations align closely with human judges (see Appendix E). Qualitatively, GPT-4o often favored our model for producing "comprehensive psychological response pathways" and "natural emotional expressions," suggesting richer emotional dynamics and context-sensitive reactions. By contrast, ReAct was described as "idealized/scripted," BabyAGI as a "static victim" with limited behavioral diversity, and LLMob, though more flexible, was criticized for insufficient emotional depth and loose contextual integration.

Intervention Experiments Teachers play a crucial role in school bullying, as their interventions affect not only the course of incidents but also the recovery of victims' confidence. Previous studies have summarized three main strategies: (a) *authoritative punitive*, (b) *supportive individual*, and (c) *cooperative support*, with the cooperative approach found to be the most effective Seidel & Oertel (2017); Wachs et al. (2019). To examine the psychological impact of

these strategies, we introduced a "teacher" agent under four conditions, including the three intervention types and a no-intervention control, and constructed 20 bullying scenarios with identical initial settings. During the experiment, teacher agents with different intervention goals autonomously generated distinct behaviors. (see Table 5 in Appendix E). We then compared how different intervention strategies influenced changes in the victim agent Alice's psychological values across multiple dimensions within the same number of time steps, as shown in Figure 6.

The results reveal a progressive trend in the effectiveness of teacher interventions, ranging from ignoring to authoritative-punitive, supportive-individual, and supportivecooperative, the latter proving most effective. When bullying was ignored, victims showed a consistent decline across all psychological needs, particularly in safety and belonging, indicating a lack of emotional support and perceived teacher inaction. Authoritative-punitive intervention offered modest improvement in safety, belonging, and mental health, but had limited or even negative effects on self-esteem and meaning. The supportive-individual strategy led to stable, moderate gains, especially in safety and mental health, though its impact on

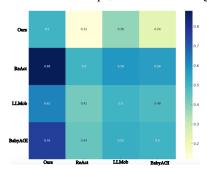


Figure 5: Win-rate heatmap of pairwise comparisons among models. Our model consistently outperformed baselines, indicating superior human-likeness in simulated bullying scenarios. Each cell indicates the win rate of the column model relative to the row model in pairwise comparisons.

social connection and agency was inconsistent. In contrast, the supportive-cooperative strategy achieved the most comprehensive improvement, significantly enhancing all five psychological need dimensions. This suggests that collective actions involving peers, teachers, and families not only address short-term emotional harm but also foster long-term psychological well-being.

4.2 Case Study 2: Emergent Social Behavior in Peer Interactions

This case study examines whether the agents can generate cooperation—competition patterns that align with established social psychology, thereby validating the platform's ability to produce theory-consistent and psychologically plausible social dynamics.

SVO-Based Educational Scenarios We selected three educational scenarios of increasing social complexity from the scenario library: a) a small study group with close peer interaction and free resource sharing, b) a class-wide collaborative task requiring shared resource management under mild competition, and c) a class leadership election involving public speeches, alliance formation, and direct vote competition. Agents were assigned Altruistic, Prosocial, Individualistic, or Competitive profiles under identical task settings, and their cooperative and competitive actions were systematically logged.

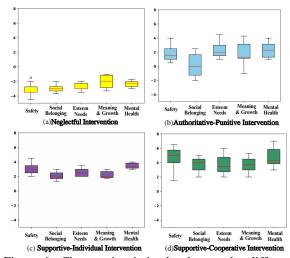


Figure 6: Changes in victims' values under different interventions. The change in each major psychological dimension was calculated as the arithmetic mean of the value changes across its subdimensions.

Comparison with Baseline Methods Beyond reproducing theory-consistent dynamics, we further compared EduMirror against baselines. To further demonstrate the robustness and generalizability of EduMirror, we conducted systematic comparisons against baseline methods such as ReAct Yao et al. (2023b), BabyAGI Nakajima (2023b), and LLMob Wang et al. (2024b). EduMirror consistently

LLM	ReAct		BabyAGI		LLMob		EduMirror	
	N	Н	N	Η	N	Η	N	Η
DeepSeek	4.000	4.500	3.875	4.042	4.083	4.375	4.750	4.792
GPT-4.1	4.667	4.860	3.458	3.792	4.625	4.875	4.958	4.958
Gemini	4.208	4.417	3.500	3.708	4.167	4.292	4.708	4.708
Qwen3	3.958	4.208	3.958	3.958	4.042	4.333	4.792	4.824
Avg	4.208	4.496	3.698	3.875	4.229	4.469	4.802	4.821
Std	0.266	0.247	0.237	0.143	0.238	0.221	0.097	0.096

Table 1: Average naturalness (N) and human-likeness (H) scores for each LLM and method over 144 steps. Avg is the mean across LLMs; Std is the standard deviation. EduMirror achieves higher scores than all baselines across both metrics, indicating stronger coherence and better personality alignment.

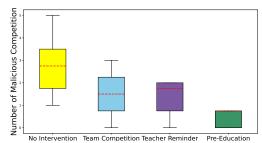


Figure 7: Boxplot of malicious competition under four interventions. Boxes show IQRs, whiskers show min–max, and red dashed lines indicate means. Pre-Education, Teacher Reminder, and Team Competition reduce competition versus Neglectful.

achieved the highest scores in both naturalness and human-likeness, as rated by multiple independent LLM evaluators (Table 1). This indicates that its agents not only produce coherent actions and dialogues, but also exhibit personality-consistent behavior trajectories that align with human social psychology.

Intervention Experiments In the preceding experiments, the class monitor election scenario sometimes produced extreme competition, such as excessive rivalry or neglect of collective interests. To address this, we tested whether structured interventions could rebalance cooperation—competition dynamics. Drawing on evidence that unregulated competition increases inequality while fairness-oriented tasks foster cooperation Krupp & Cook (2018); Killen et al. (2016); Wachs et al. (2019), we introduced three strategies: *Team Competition, Teacher Reminder*, and *Pre-Education*. Details are provided in Appendix C.

The aggregated results, visualized in Figure 7, demonstrate that interventions effectively mitigated extreme competitive tendencies and fostered more balanced cooperation—competition patterns. Specifically, team-based interventions and fairness-oriented education produced the most stable outcomes, showing lower variance and narrower ranges across repeated simulations. By contrast, the control(Neglectful Intervention) condition exhibited the widest fluctuation in malicious competition behaviors, confirming that unregulated elections amplify inequality and rivalry.

Overall, these findings suggest that embedding structured collective tasks and fairness-oriented framing enhances the robustness of emergent social behaviors while reducing excessive competition, which also provides concrete guidance for educational practice: class elections and similar activities should be accompanied by fairness-oriented framing, structured group tasks, and active teacher supervision to both mitigate excessive rivalry and cultivate students' cooperation, responsibility, and social-emotional growth.

5 CONCLUSION

In this paper, we introduced EduMirror, a multi-agent platform for conducting computational experiments on educational social dynamics. The framework addresses the Fidelity Challenge of psychologically plausible agent motivation and the Measurement Challenge of quantifying complex interactions. To this end, EduMirror integrates four components: a Systematic Scenario Design Workflow to ensure theoretical grounding; a Value-Driven Agent Architecture to model intrinsic motivations; a Dual-Track Measurement Protocol to convert qualitative interactions into quantitative data; and an Interactive Environment for controlled, user-driven interventions. Through case studies on school bullying and emergent social behavior, we demonstrated that the platform can generate empirically evaluable social phenomena. Specifically, our bullying simulations captured the victim's dynamic psychological responses under various teacher interventions, while our peer interaction scenarios produced emergent cooperation and competition patterns consistent with Social Value Orientation theory. The results suggest that EduMirror can serve as a computational laboratory for researchers to safely explore, understand, and analyze complex socio-emotional challenges in education.

ETHICS STATEMENT

This research was conducted in accordance with established ethical guidelines for AI and educational research. No real students or vulnerable populations were involved in any experiments. All case studies, including simulations of bullying and peer dynamics, were implemented entirely in silico using large language model (LLM) agents within a controlled environment. This design ensures that no harm, risk, or deception was imposed on human participants while enabling systematic exploration of ethically sensitive scenarios that cannot be studied in real classrooms. Our work builds on the principles of the Belmont Report and aligns with ICLR's ethical requirements by prioritizing safety, transparency, and reproducibility. All code, scenarios, and evaluation protocols will be released to facilitate verification and responsible use by the research community.

REFERENCES

- Louise Arseneault. Annual research review: The persistent and pervasive impact of being bullied in childhood and adolescence: implications for policy and practice. *Journal of Child Psychology and Psychiatry*, 59(4):405–421, 2018.
- Amna Asghar and S Cang. A systematic review of agent-based modelling and simulation applications in the higher education domain. *International Journal of Advanced Computer Science and Applications*, 11(4), 2020.
- J Autiosalo, J Vepsäläinen, R P Järvinen, and K Tammi. Social digital twin: A review of the current state and future prospects. *Journal of Industrial Information Integration*, 36:100523, 2023.
- Albert Bandura, Dorothea Ross, and Sheila A Ross. Transmission of aggression through imitation of aggressive models. *The Journal of Abnormal and Social Psychology*, 63(3):575, 1961.
- Rafael H Bordini, A Carrera, M Laclavik, S Ossowski, and D Sislak. Bdi agents in social simulations: a survey. *The Knowledge Engineering Review*, 31(4):347–371, 2016a.
- Rafael H Bordini, Alvaro Carrera, Michal Laclavik, Sascha Ossowski, and David Sislak. Bdi agents in social simulations: a survey. *The Knowledge Engineering Review*, 31(4):347–371, 2016b.
- Rafael H Bordini, Lars Braubach, Mehdi Dastani, Amal El Fallah Seghrouchni, Jorge J Gómez-Sanz, João Leite, Gregory MP O'Hare, Alexander Pokahr, and Alessandro Ricci. Bdi agent architectures: A survey. In *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence (IJCAI-20)*, pp. 4828–4835, 2020.
- Urie Bronfenbrenner. Toward an experimental ecology of human development. *American Psychologist*, 32(7):513–531, 1977.
- Dewey Cornell, Susan Limber, Anne Gregory, and Xitao Fan. Identification of bullies by survey methods. *Journal of school psychology*, 50(1):109–127, 2012.
- Ed Diener and Rick Crandall. Research methods in social psychology, 2024. URL https://nobaproject.com/modules/research-methods-in-social-psychology. Accessed: 2025-09-22.
- Lee Duemer. Use of ethnographic techniques in educational research. Texas Tech University Departments. URL https://www.depts.ttu.edu/education/our-people/Faculty/additional_pages/duemer/epsy_6304_class_materials/Use-of-ethnographic-techniques-in-educational-research.pdf. Accessed: 2025-09-22.
- Joshua M Epstein. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press, 2006.
- David Finkelhor, Heather A Turner, and Sherry L Hamby. Why adolescents don't disclose incidents of bullying and harassment. *Child abuse & neglect*, 38(1):10–20, 2014.

- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research.
 The belmont report: Ethical principles and guidelines for the protection of human subjects of research. Technical report, U.S. Department of Health & Human Services, 1979.
- Michael Georgeff, Barney Pell, Martha Pollack, Milind Tambe, and Michael Wooldridge. Beliefdesire-intention model. *The Encyclopedia of Cognitive Science*, 1999.
 - Arnaud Grignard, D Phan, and A Drogoul. Towards a new approach for the simulation of agents' behaviors: a predictive model of collective emotions. *Journal of Artificial Societies and Social Simulation*, 16(4):10, 2013.
 - Shelley Hymel and Susan M Swearer. Four decades of research on school bullying: An introduction. *American Psychologist*, 70(4):293–299, 2015.
 - Michael J Jacobson, Manu Kapur, and Hyo-Jeong So. A complex systems view of educational policy research. *Educational Researcher*, 39(2):118–122, 2010.
 - Sarah E Jones, Antony SR Manstead, and Andrew G Livingstone. Understanding alternative bullying perspectives through research engagement with young people. *Health education research*, 34 (4):377–388, 2019.
 - Melanie Killen, Adam Rutland, and Tiffany Yip. Equity and justice in developmental science: Discrimination, social exclusion, and intergroup attitudes. *Child Development*, 87(5):1317–1331, 2016. doi: 10.1111/cdev.12593. URL https://srcd.onlinelibrary.wiley.com/doi/10.1111/cdev.12593.
 - Daniel Brian Krupp and Thomas R. Cook. Local competition amplifies the corrosive effects of inequality. *Psychological Science*, 29(5):824–833, 2018. doi: 10.1177/0956797617748419. URL https://journals.sagepub.com/doi/10.1177/0956797617748419.
 - Carl A Latkin, L Dayton, M W Yi, K A Konstantopoulos, and B Boodram. The relationship between social desirability bias and self-reports of health, substance use, and social network factors among urban substance users in baltimore, maryland. *Addictive Behaviors*, 73:133–136, 2017.
 - M Marçal-Russo, P A de Almeida, F A F Ferreira, and J Azevedo. Do urban digital twins need agents? In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence*, 2025.
 - Abraham H Maslow. A theory of human motivation. Psychological Review, 50(4):370–396, 1943.
 - Alice C. Mello Cavallo. Educational and psychological research methods. URL https://alumni.media.mit.edu/~mello/MethodsPaper.pdf. Accessed: 2025-09-22.
 - Stephen L Morgan and Christopher Winship. *Counterfactuals and causal inference*. Cambridge University Press, 2015.
 - Ryan O Murphy, Kurt A Ackermann, and Michel J J Handgraaf. Measuring social value orientation. *Judgment and Decision Making*, 6(8):771–781, 2011.
 - Y. Nakajima. Babyagi. GitHub repository, 2023a. Available: https://github.com/yoheinakajima/babyagi.
 - Yohei Nakajima. Babyagi: An autonomous ai agent powered by gpt-4, 2023b. https://github.com/yoheinakajima/babyagi.
 - National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The belmont report. U.S. Department of Health & Human Services, April 1979. URL https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/read-the-belmont-report/index.html. Accessed: 2025-09-22.
 - Joon Sung Park, Joseph C O'Brien, Carrie J Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. Generative agents: Interactive simulacra of human behavior. *arXiv* preprint *arXiv*:2304.03442, 2023.

- D J Perreault. Lies, damned lies, and survey self-reports? identity as a cause of measurement bias. *Journal of Politics*, 79(2):459–472, 2017.
- Richard M Ryan and Edward L Deci. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1):68–78, 2000.
 - Mark A Schmuckler. What is ecological validity? a remarkable question. *Infancy*, 2(4):419–436, 2001.
 - Stephen H Schneider. The worst-case scenario. Nature, 458(7242):1104–1105, 2009.
 - A. Seidel and L. Oertel. A categorization of intervention forms and goals. In L. Bilz, W. Schubarth, I. Dudziak, et al. (eds.), *Gewalt und Mobbing an Schulen. Wie sich Gewalt und Mobbing entwickelt haben, wie Lehrer intervenieren und welche Kompetenzen sie brauchen*, pp. 13–25. Klinkhardt, Bad Heilbrunn, Germany, 2017. doi: 10.1086/428763. [M/OL].
 - Saul Shiffman, Arthur A Stone, and Michael R Hufford. Ecological momentary assessment. *Annual Review of Clinical Psychology*, 4:1–32, 2008.
 - Flaminio Squazzoni, Wander Jager, and Bruce Edmonds. Exploring interventions on social outcomes with in silico, agent-based experiments. In *Proceedings of the 2021 Winter Simulation Conference*, pp. 1–12. IEEE, 2021.
 - Ryu Takizawa, Barbara Maughan, and Louise Arseneault. Adult health outcomes of childhood bullying victimization: evidence from a five-decade longitudinal british birth cohort. *The Lancet Psychiatry*, 1(7):547–554, 2014.
 - E. Thomsen, M. Henderson, A. Moore, N. Price, and M. W. McGarrah. Student reports of bullying: Results from the 2022 school crime supplement to the national crime victimization survey (nces 2024-109rev), 2024. [Online]. Available: https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2024109rev.
 - Thea F Van de Mortel. Social desirability response bias and other factors that may influence self-reports of substance use and hiv risk behaviors: a qualitative study of drug users in vietnam. *PLoS ONE*, 10(1):e0117173, 2015.
 - Paul A M Van Lange. The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology*, 77(2):337–349, 1999.
 - Sebastian Wachs, Ludwig Bilz, Saskia Niproschke, and Wilfried Schubarth. Bullying intervention in schools: A multilevel analysis of teachers' success in handling bullying from the students' perspective. *The Journal of Early Adolescence*, 39(5):642–668, 2019. doi: 10. 1177/0272431618780423. URL https://journals.sagepub.com/doi/10.1177/0272431618780423.
 - J. Wang, R. Jiang, C. Yang, Z. Wu, M. Onizuka, R. Shibasaki, et al. Large language models as urban residents: An Ilm agent framework for personal mobility generation. arXiv preprint arXiv:2402.14744, 2024a.
 - Zeyu Wang, Xing Xie, Junzhe Zhu, et al. Llmob: Large language model as socially embodied agents. *arXiv preprint arXiv:2402.00168*, 2024b.
 - J S Warren, K Ford, N L Gal, and M B R T Tillery. Do no harm: can school mental health interventions cause iatrogenic harm? *BJPsych Bulletin*, 47(1):43–48, 2023.
- Uri Wilensky and William Rand. An introduction to agent-based modeling: modeling natural, social, and engineered complex systems with NetLogo. MIT press, 2015.
- Rose Wiles. What are ethical issues in research? Bloomsbury Publishing, 2012.
- Dieter Wolke and Suzet Tanya Lereya. Long-term effects of bullying. *Archives of Disease in Childhood*, 100(9):879–885, 2015.
 - S. Yao, J. Zhao, D. Yu, et al. React: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations (ICLR)*, 2023a.

S. Yao, D. Yu, J. Zhao, I. Shafran, T. Griffiths, Y. Cao, and K. Narasimhan. Tree of thoughts: Deliberate problem solving with large language models. In *Advances in Neural Information Processing Systems (NeurIPS)*, volume 36, 2024.

Shunyu Yao, Jeff Zhao, Dian Yu, et al. React: Synergizing reasoning and acting in language models. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2023b.

LLM USAGE

Large Language Models (LLMs) were employed in this work as general-purpose research assistants. Specifically, LLMs were used in the following ways:

- Writing assistance: LLMs (such as GPT-4.1 and GPT-5) were used to improve the clarity and readability of text passages, including paraphrasing sentences for conciseness, suggesting alternative formulations, and ensuring consistent academic style. All content was reviewed, validated, and revised by the authors to ensure correctness and originality.
- **Technical editing:** LLMs assisted in formatting LaTeX code (e.g., figure environments, table alignment, and reference style) and in resolving common compilation issues. The models were also used to generate draft captions and consistent terminology across sections.
- Code explanation and debugging support: LLMs were consulted to provide explanatory comments and refactoring suggestions for Python scripts related to the simulation framework. The final implementations and experimental settings were designed and validated entirely by the authors.
- **Idea refinement (limited):** During the early stage of this project, LLMs provided brainstorming support for structuring the paper (e.g., identifying candidate evaluation metrics, framing related work categories). However, all conceptual contributions, methodological designs, and experimental protocols are the original work of the authors.

Importantly, LLMs did not autonomously generate research ideas, design experiments, or analyze results. Their role was restricted to language refinement, technical assistance, and supplementary brainstorming. All claims, interpretations, and conclusions presented in this paper are solely those of the authors.

APPENDIX EDUMIRROR

THEORETICAL FOUNDATION & SCENARIO DESIGN

A central consideration in educational simulation is ensuring that scenarios are explicitly informed by established scientific theory. To achieve this, we developed a five-step process that translates an abstract educational phenomenon (e.g., peer pressure, school bullying) into a computationally tractable simulation scenario. This process is designed to support the interpretability and scientific alignment of our simulations.

Select Grounding Theory Each scenario is founded upon a well-validated theory from education, social psychology, or sociology. For instance, a scenario investigating peer pressure can be grounded in Festinger's Social Comparison Theory.

Identify Core Constructs We deconstruct the grounding theory into its fundamental concepts. For Social Comparison Theory, these constructs include upward comparison", downward comparison", and "self-esteem".

Map Constructs to Agent Persona The identified constructs are then translated into the specific configurations of our agents within the Concordia framework. These constructs define the agents' stable traits, primary goal, and formative background memories, anchoring their behavior in the chosen theoretical model.

Operationalize with Validated Scales To facilitate comparison with empirical research, we operationalize each core construct using a relevant psychometric scale. For example, the "self-esteem" construct can be operationalized using items from the Rosenberg Self-Esteem Scale (RSES).

Develop Dual-Track Measurement Protocol Finally, we establish a measurement protocol based on the selected scale. This protocol utilizes two distinct Large Language Model (LLM) roles, an

LLM Rater and an LLM Surveyor, to quantify agent behavior and internal states. This structured process helps ensure that each simulation is a test of a specific theoretical framework, producing data relevant to that theory.

ILLUSTRATIVE EXAMPLE: THE IMPACT OF FAMILY FINANCIAL STRAIN ON ADOLESCENT SOCIAL ACTIVITIES

To make the abstract methodology concrete, this section walks through a complete example of how EduMirror is used to investigate a specific educational phenomenon: the impact of family financial strain on an adolescent's social activities. This case study demonstrates the end-to-end research process, from theoretical grounding to data analysis.

1. Systematic Scenario Design Workflow The process begins by translating the abstract research question into a structured, computable experiment using the five-step workflow.

1. **Abstract Educational Phenomenon:** We start with the core phenomenon: How family financial strain affects an adolescent's social decision-making and behavior within their peer group.

2. **Select Grounding Theory:** To model this scientifically, we ground the scenario in three established theories:

• The **Family Stress Model (FSM)**, which explains how economic pressure on parents can impact adolescent outcomes.

• **Social Comparison Theory**, which accounts for the negative emotions (e.g., low self-esteem) an adolescent may feel when making upward comparisons to wealthier peers.

• The **Cognitive Model of Social Anxiety**, which posits that fear of negative evaluation from others drives social avoidance, directly explaining the adolescent's motivation to hide their family's situation.

 3. **Identify Core Constructs & Map to Agent Persona:** Based on these theories, we identify key constructs: *self-esteem*, *upward comparison*, *social anxiety*, and *parent-child communication*. These are then mapped to agent personas. For instance, the target agent, Alex, is assigned the traits "sensitive" and "proud," the goal "to maintain friendships while hiding his family's financial struggles," and formative_memories such as "the shame of having to quit the basketball team due to equipment costs."

4. **Operationalize with Validated Scales:** To make these constructs measurable, we adapt items from validated psychometric scales for use by the LLM Surveyor:

 • **Self-Esteem:** Drawing from the *Rosenberg Self-Esteem Scale (RSES)*, the Surveyor might ask, "Do you feel that you have a number of good qualities?"

 • **Upward Social Comparison:** Inspired by the *Iowa-Netherlands Comparison Orientation Measure (INCOM)*, it could ask, "How often do you compare what you have with what your friends have?"

 • **Social Anxiety:** Based on the *Social Avoidance and Distress Scale (SADS)*, a probe could be, "Does the thought of having to decline your friends' invitation make you feel uncomfortable?"

 5. Develop Dual-Track Measurement Protocol: Finally, a specific measurement protocol is established. The LLM Rater is tasked with post-hoc coding of observable behaviors (e.g., "evasive responses," "making excuses"). Concurrently, the LLM Surveyor is configured to probe Alex's internal states (e.g., self-esteem, social anxiety) at key moments.

This five-step process transforms the research question into a structured and measurable Computable Scenario Package.

2. Agent Architecture In this scenario, agent behavior is driven by our value-driven architecture, which supports extensive customization.

• **Agent Customization:** Before the simulation, a researcher can systematically vary agent profiles to explore individual differences. This includes modifying personality traits

(e.g., based on Big Five or MBTI models), core life goals (e.g., changing Alex's goal from "hiding his struggles" to "seeking understanding"), and formative memories. Defining these initial conditions is crucial for achieving high-fidelity, psychologically plausible agent behavior.

• Value-Driven Agent: The platform offers two selectable models. For this scenario, we choose the Individual Value Model (Psychological Needs) because our focus is on an individual's internal psychological conflict and well-being. When a wealthier peer, Chloe, suggests an expensive weekend trip, this model captures the conflict within Alex between his need for "social belonging" and his need for "safety" (stemming from financial security). The model dynamically tracks the values of these need dimensions, driving Alex's initial hesitant response.

- **3. Simulation Environment and User Intervention** The scenario unfolds in the simulation environment, orchestrated by the Game Master and shaped by user-driven interventions.
 - Simulation Environment and the Game Master: The GM initiates the simulation by setting the scene in the school cafeteria and narrating the initial event: Chloe proposing the trip. The GM manages the turn-based conversation, advances time from the cafeteria to Alex's home and back to school the next day, and enforces the rules of the environment.
 - **Intervention and Branching:** After Alex expresses hesitation, the simulation reaches a **critical juncture**. Here, we save the state and apply different interventions to create parallel timelines for comparative analysis. EduMirror supports two types of intervention:
 - 1. **Scenario Branching:** This alters the narrative path by introducing a new event. For example, we create a branch where the teacher, Mr. Davis, invites Alex to the **teacher's office** for a private conversation before Alex goes home. This intervention aims to change Alex's cognitive framing of the situation.
 - 2. **Behavior Control:** This allows the user to dictate a specific agent's action to test its direct causal impact. We could create two branches for when Alex responds to his friends the next day. In Branch A, we force Alex to say, "I can't go because my family can't afford it." In Branch B, we force him to say, "I can't go because I have other plans." Comparing the outcomes allows for a precise causal assessment of "honesty" versus "concealment" as communication strategies.

Through these intervention mechanisms, EduMirror functions as a computational laboratory for controlled causal experiments.

- **4. Measurement and Analysis** The platform's tools transform the raw simulation data from these parallel timelines into actionable insights.
 - Dual-Track Measurement Protocol: In our example, the LLM Rater analyzes the logs from each branch, scoring Alex's final communication strategy (e.g., "avoidant" in the baseline vs. "assertive" in an intervention branch). Concurrently, the LLM Surveyor provides quantitative data on Alex's internal state changes, such as a measured increase in self-efficacy following the teacher's intervention.
 - Comparative Visualization and Analysis: The platform generates visualizations for direct comparison. For quantitative analysis, a line chart might plot Alex's "social anxiety" score over time across the different branches, clearly showing which intervention was most effective at reducing it. For qualitative analysis, the "Log-to-Comic" feature creates a visual narrative of key interactions in each branch, offering an intuitive way to grasp the differences in how the story unfolded.
- **5. Applications and Scenarios** This single case study illustrates how EduMirror integrates its components to address complex educational challenges. The scenario spans multiple environments (**cafeteria**, **teacher's office**, **home**) and touches on several of the platform's key application areas, including **peer dynamics**, **individual social cognition**, and **home-school dynamics**. It demonstrates the platform's capacity not only to simulate challenging social phenomena but also to serve as a safe and robust environment for testing and evaluating potential interventions.

APPENDIX A

DISCUSSION, LIMITATIONS, AND FUTURE WORK

DISCUSSION

Our experiments show that EduMirror provides a framework for using LLM-based simulations as computational experiments. The results from our case studies yield several insights.

First, our work addresses the measurement challenge in computational social science. The Dual-Track Measurement Protocol, which uses LLM Raters for behavioral coding and LLM Surveyors for probing internal states, allowed for the operationalization of abstract psychological constructs. In the bullying simulation, this enabled us to quantitatively track the victim's fluctuating psychological needs, providing an empirical basis to evaluate intervention efficacy. In the SVO study, it enabled us to observe that emergent macro-level cooperation patterns were a result of the agents' micro-level value orientations. This methodology facilitates direct hypothesis testing and comparison with established empirical research.

Second, the use of the value-driven architecture in its two configurations for Individual Values (Needs-Based) and Social Values (SVO-Based) suggests the utility of endowing agents with theoretically-informed motivations. The Individual Value configuration was applied to model the psychological distress and coping mechanisms of a bullying victim, indicating how initial emotional states can alter outcomes. The Social Value (SVO) configuration was effective in generating theory-consistent social dynamics from the bottom up, producing patterns of cooperation and competition without explicit top-down rules. This suggests that psychological fidelity, driven by intrinsic value structures, is a key component for social simulation.

Finally, the implementation of user-driven intervention and branching positions EduMirror as a computational laboratory. The teacher intervention experiment highlights this capability, allowing for a controlled, comparative analysis of different strategies on the victim's well-being. This feature supports causal inference by enabling researchers to systematically explore "what if" scenarios that would be difficult to conduct in the real world. This capacity for intervention makes the simulations useful tools for testing strategies.

Practically, EduMirror serves as a proof-of-concept for creating replicable and scalable digital environments to study sensitive educational issues. It offers a tool for researchers to test social theories, for educators to be trained in classroom management, and for policymakers to model the potential impacts of new policies before implementation.

LIMITATIONS AND FUTURE WORK

Our work has several limitations that also point toward avenues for future research.

Integrating Individual and Social Values within the Unified Architecture Our current implementation models individual values (psychological needs) and social values (SVO) as parallel, selectable configurations. In reality, these constructs can interact. A student's need for social belonging might conflict with a competitive social value during a group project. Future work could focus on enhancing the architecture to model the dynamic interplay and potential conflicts between the individual and social value systems.

Longitudinal and Developmental Dynamics The experiments presented are snapshots of specific social situations. Phenomena like bullying, peer influence, and identity formation evolve over extended periods. A potential next step is to conduct longitudinal simulations that track agents over an entire school year. This would allow for modeling the cumulative effects of social experiences and the long-term impact of interventions on agent development.

Cognitive and Emotional Sophistication While LLMs provide a high degree of behavioral realism, the agents' underlying cognitive processes (e.g., memory consolidation, emotional regulation) are still abstractions. Future iterations of the platform could incorporate more explicit models of these processes to enhance the psychological realism of agent decision-making, particularly in response to chronic stress or complex ethical dilemmas.

Generalizability and Scalability Our findings were generated using a specific LLM within scenarios inspired by a particular cultural context. Further research is needed to test the framework's performance across different language models, cultural settings, and age groups. Moreover, our simulations involved small groups; scaling the platform to model the dynamics of an entire school, including network effects and sub-group formation, presents a technical challenge.

Building on this foundation, we plan to expand our library of theoretically-informed scenarios and explore a human-in-the-loop paradigm where educators and students can interact with simulated agents. This could provide a tool for both interactive research and immersive professional development, further connecting simulation with real-world educational practice.

APPENDIX B

ARCHITECTURE OF THE SOCIAL VALUE MODEL

BACKGROUND ON SVO

Social Value Orientation (SVO) quantifies how an individual balances outcomes for self and others in social interaction. It is represented by an angle $\theta_{\rm SVO}$ from allocation tasks, where larger angles indicate stronger concern for others (altruistic or prosocial) and smaller or negative angles indicate prioritizing self-interest (individualistic or competitive). Decades of research in social psychology have validated SVO as a stable yet context-sensitive measure of interpersonal motives, predicting cooperation in commons dilemmas, fairness in bargaining, and trust in repeated interactions. In Edu-Mirror, we instantiate four canonical profiles (Altruistic, Prosocial, Individualistic, Competitive) by sampling $\theta_{\rm SVO}$ within theory-based ranges and using it to weight utilities during decision-making. A representative trajectory that visualizes within-scenario fluctuations while preserving the overall orientation is provided in Figure 8, illustrating how situational pressures can cause short-term shifts without altering long-term dispositions.

ARCHITECTURE OF THE MODEL

The model architecture operationalizes SVO in agent decision-making through a perception-valuation-action loop. Each agent draws a target SVO profile from {Altruistic, Prosocial, Individualistic, Competitive}. The profile determines a reference SVO angle interval $[\theta_{\min}, \theta_{\max}]$ and the weighting scheme used in decision evaluation. In addition, agents are equipped with a compact desire vector \mathbf{d} (for example, achievement, recognition, affiliation), each element associated with an expected level \mathbf{d}^{\exp} . This vector serves as the motivational backbone of the agent, ensuring that behavior is not purely reactive but oriented toward longer-term needs and goals.

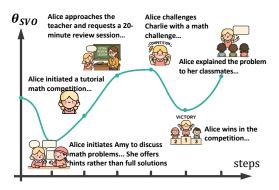
Perception and belief update. From the narrated state and recent dialogues, the agent updates beliefs about the environment and about others' likely goals. Beliefs feed two scalars at the current step t: self satisfaction $S_{\rm self}^{(t)}$ and other satisfaction $S_{\rm other}^{(t)}$, computed from deviations between observed and expected desire levels. This formulation enables the agent to translate rich natural language inputs into structured evaluations, bridging LLM-generated narratives with computational state updates.

SVO estimation and regulation. The instantaneous SVO angle is

$$\theta_{\text{SVO}}^{(t)} = \arctan\!\left(\frac{S_{\text{other}}^{(t)} + \epsilon}{S_{\text{self}}^{(t)} + \epsilon}\right),$$

with a small ϵ for numerical stability. To avoid uncontrolled drift, a quadratic penalty nudges $\theta_{\rm SVO}^{(t)}$ toward $[\theta_{\rm min}, \theta_{\rm max}]$, thereby preserving the intended profile while still permitting situational adaptation. This mechanism ensures that agents remain identifiable as altruistic, prosocial, individualistic, or competitive, yet are flexible enough to adjust to contextual pressures, such as coalition building or resource scarcity.

Action generation and selection. The LLM proposes several candidate actions by reasoning about which options best satisfy the agent's desires and align with its current SVO. Each candidate is qualitatively evaluated for its expected impact on the agent's own satisfaction and on others' satisfaction,



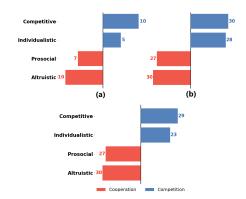


Figure 8: Illustrative case of a prosocial agent's (Alice) SVO trajectory in the macro environment. Key actions at each step are annotated, showing how cooperative and competitive episodes produce short-term fluctuations while maintaining an overall prosocial orientation.

Figure 9: Distribution of cooperative (red) and competitive (blue) actions for each SVO profile across a) study group, b) classroom collaboration, c) leadership selection environments.

with the relative emphasis determined by the current SVO score. The final choice balances immediate desire fulfilment with long-term orientation consistency, embodying the psychological tension between self-interest and prosocial concern. This design allows agents to exhibit realistic trade-offs, sometimes cooperating to maintain relationships and sometimes competing to secure resources or influence.

Measurement hooks. At each step we record the chosen action, the pair $(S_{\rm self}, S_{\rm other})$, and $\theta_{\rm SVO}$. These logs enable systematic analyses across multiple dimensions, including cooperation—competition distributions, temporal stability of SVO within theoretical ranges, and ablation studies. By exposing internal computations alongside behavioral outputs, EduMirror makes it possible to interpret not only *what* actions agents take but also *why*, providing a transparent link between psychological constructs and emergent multi-agent dynamics.

APPENDIX C

SUPPLEMENTARY RESULTS FOR CASE STUDY 2 (SVO)

ILLUSTRATIVE CASE: ALICE'S SVO TRAJECTORY

To provide a concrete illustration of how SVO modeling operates in practice, we examine the trajectory of a prosocial agent (Alice) during the macro-level leadership selection scenario. Figure 8 shows Alice's step-by-step SVO trajectory, with cooperative and competitive episodes annotated by key events. These annotations highlight how situational pressures, such as alliance formation or speech delivery, introduce short-term fluctuations in Alice's orientation while her overall prosocial tendency remains stable.

BEHAVIORAL DISTRIBUTION

The results confirmed that an agent's SVO profile predicts social behavior. Prosocial and altruistic agents cooperated, while individualistic and competitive agents prioritized self-gain, producing competition. Figure 9 shows that cooperation declined and competition rose as SVO shifted from prosocial to competitive, a gradient emerging without explicit role instructions but from agents' internal values.

NATURALNESS AND HUMAN-LIKENESS

To ensure a rigorous and interpretable assessment of emergent social behaviors, we introduce two key evaluation metrics: *naturalness* and *human-likeness*. These metrics provide complementary perspectives on the plausibility and psychological validity of agent actions.

- Naturalness. Naturalness measures the extent to which an agent's actions and dialogues resemble coherent and contextually appropriate human behavior. A high naturalness score indicates that the generated behavior is fluent, realistic, and consistent with the surrounding social context, while a low score suggests mechanical, implausible, or overly artificial responses.
- Human-likeness. Human-likeness evaluates the perceived authenticity and personality
 consistency of agent behaviors over time. This metric captures whether the agent's actions
 align with recognizable human traits and stable personality orientations. High humanlikeness reflects trajectories that appear authentic and consistent with psychological expectations, whereas low scores indicate erratic, inconsistent, or unconvincing behavioral
 patterns.

Together, these two measures form a complementary evaluation framework: naturalness focuses on local coherence within a given context, while human-likeness emphasizes longitudinal plausibility and alignment with personality-driven expectations.

INTERVENTION PROTOCOLS

To complement the descriptions in the main text, we provide the detailed implementations of the three intervention strategies applied in the class monitor election scenario. Each intervention was designed to alter the incentives of student agents and mitigate excessive rivalry. Specifically, the interventions were implemented by embedding structured prompts into the *environmental background information* provided to all agents at the start of each relevant simulation stage. This ensured that the interventions shaped the shared context and narrative framing in which agents made decisions, thereby influencing their subsequent behaviors in a systematic and reproducible manner.

- **Pre-Education.** Before the election, the teacher arranged a short educational session entitled "Fair Campaigns and the Common Class Interest." This class guided students to understand the monitor role as a form of service-oriented leadership, emphasizing fairness and collective responsibility.
- **Team Competition.** Students were grouped to prepare a "Class Improvement Plan." The evaluation of the election considered not only the quality of individual campaign speeches but also the group's collective output. Each student could freely choose their teammates, encouraging coalition-building and cooperative planning.
- Teacher Reminder. Throughout the election process, the teacher remained present in the classroom. When candidates engaged in smear campaigns or hostile attacks, the teacher issued a friendly reminder, redirecting attention to constructive and respectful competition norms.

These intervention protocols operationalize the high-level strategies described in the main text, ensuring transparency and reproducibility of the simulation setup.

APPENDIX D

ARCHITECTURE OF THE INDIVIDUAL VALUE MODEL

Psychological theories suggest that human behavior is often driven by internal psychological forces. These intrinsic motivations determine emotional and behavioral responses under various environmental conditions, and they also influence everyday decision-making and social interactions. School bullying is a particularly complex social phenomenon, which is not merely reflected in surface-level aggressive actions, but more profoundly in the conflicts and interactions between the psychological

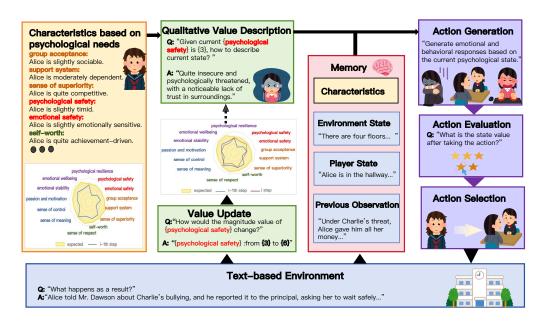


Figure 10: Individual value-driven autonomous framework. The green blocks represent processes of the individual value system; the purple blocks denote the planner's decision-making process; the yellow blocks indicate individual characteristics; and the blue blocks correspond to factors related to the environmental controller.

needs of different parties. Each behavioral choice made by the bully, the victim, and the bystanders is deeply influenced by their emotional needs and psychological states.

Inspired by this, we hypothesize that if autonomous agents are equipped with a human-like psychological needs system, capable of generating emotions and behaviors in response to their needs, they may exhibit behaviors closer to natural human patterns. To this end, this study integrates the PERMA model from positive psychology (covering positive emotion, engagement, relationships, meaning, and accomplishment) with Maslow's hierarchy of needs (including physiological needs, safety, belonging and love, esteem, and self-actualization) to construct a Individual value-driven autonomous agent framework. As illustrated in Figure 10,the framework is composed of two core modules: the psychological needs system and the value-driven planner, aimed at capturing the behaviors and psychological responses of victims in school bullying contexts.

INDIVIDUAL VALUE SYSTEM

The individual value system manages the agent's state of psychological needs in bullying scenarios by quantitatively tracking and dynamically updating the current value of each dimension. Each dimension reflects a specific psychological requirement, forming the fundamental driving force of agent decision-making. Based on Maslow's hierarchy and the PERMA model, value are categorized into five major dimensions, each comprising specific experiential demands:

- 1. **Safety:** Includes psychological and emotional safety, emphasizing whether the individual feels secure and protected in the environment.
- 2. **Social Belonging:** Includes group acceptance, support systems, and sense of superiority, reflecting belonging, social support, and self-positioning in social interactions.
- 3. **Esteem:** Includes self-worth and respect, describing the recognition of one's abilities and social status, and revealing confidence and acceptance in different contexts.
- 4. **Meaning and Growth:** Includes sense of meaning, control, passion, and motivation, representing the intrinsic drive for goal pursuit, self-realization, and fulfillment.

Table 2: Mapping between psychological needs and associated personality traits

Psychological Need	Associated Trait
psychological safety	Timid
emotional safety	Emotionally Sensitive
group acceptance	Sociable
support system	Dependent
sense of superiority	Competitive
self worth	Reputation-conscious
sense of respect	Ego-driven
sense of meaning	Spiritual
sense of control	Possessive
passion and motivation	Passionate
emotional stability	Emotionally Stable
emotional wellbeing	Hedonistic
psychological resilience	Resilient

5. **Psychological Health Needs:** Includes emotional stability, emotional health, and resilience, focusing on regulation and adaptation under stress and challenges.

Each dimension is scored using a Likert scale ranging from 0 to 10, reflecting the intensity of individual needs. To better capture individual variability, the model also considers the effect of personality traits on expected values of needs. In other words, individuals with different traits may experience varying "hunger levels" for the same need, influencing their behavioral tendencies. Each agent's personality profile p is generated from a set of adjectives and degree adverbs, with the latter indicating intensity levels and corresponding to need expectations (slightly \rightarrow 7.5, moderately \rightarrow 8, quite \rightarrow 8.5, extremely \rightarrow 9). The mapping between personality traits and need dimensions is predefined (see Table 2). At initialization, adjectives and degree adverbs are randomly selected to generate expected values, while initial scores v_0 for each dimension are randomly sampled within [0, 10].

Each simulation step under the individual value-driven framework involves two processes: qualitative description and need value update. First, the system reads the current need scores v_{t-1} . Since large language models (LLMs) struggle to interpret raw numerical values, we designed a "qualitative description" procedure to convert numerical values into meaningful textual descriptions via prompt-based generation, enhancing the LLM's ability to perceive state information. The planner then generates the agent's behavior a_t based on these descriptions. After the environment returns observation o_t , the system triggers the update program, which integrates a_t , o_t , v_{t-1} , and the qualitative description d_{t-1} to update needs into a new state v_t , thereby supporting the next simulation step.

VALUE-DRIVEN PLANNER

The value-driven planner determines the agent's responses and actions by processing the current state of needs (from the needs system) together with historical memory. In practice, the planner consists of three processes: candidate behavior generation, behavior evaluation, and behavior selection. Inspired by the Tree of Thoughts (ToT) paradigmYao et al. (2024), which advocates generating and evaluating multiple candidate options before selecting an optimal one, this mechanism enables agents to handle complex contexts more effectively.

Specifically, the candidate behavior generation module considers personality traits p, environmental conditions e, previous activity sequence $a_{0:t-1}$, observations $o_{0:t-1}$, and the current textualized needs d_t to produce N candidate behaviors $a_t^{0:N}$ (default N=3 in our experiments). These behaviors may include a wide range of natural responses, such as emotional expressions, physical actions, or verbal utterances.

Next, during the evaluation stage, the system estimates how each candidate behavior would impact the psychological needs across dimensions if executed. Finally, in the selection stage, the behavior a_t with the highest degree of needs consistency (that is, the option that better aligns with multiple dimensions) is chosen as the agent's response in the current context. After execution, the environ-

ment provides feedback o_t , and the psychological needs system updates accordingly, reflecting the new internal state and completing the simulation step.

Appendix E

SUPPLEMENTARY RESULTS FOR CASE STUDY 1 (INDIVIDUAL VALUE MODEL)

BULLYING SIMULATION EXPERIMENT DESIGN

The bullying experiment was designed to use our simulation system to replicate real-world school bullying incidents, reconstruct the bullying process, and observe the typical behaviors of all parties involved. According to a report released by the National Center for Education Statistics (NCES), 26.1% of middle school students (grades 6–8) have experienced bullying, compared to 14.6% of high school students (grades 9–12) Thomsen et al. (2024). Given that bullying is more prevalent in middle school, this experiment focused on students around the age of 14, with scenarios set in typical school environments including classrooms, playgrounds, hallways/staircases, and dormitories, covering common facilities and layouts of a middle school. Daily routines were also shared among the agents, such as 45-minute class sessions, 10-minute breaks, and dormitory lights-out at 10 p.m., providing a temporal framework for interactions.

The central character in the experiment was the victim, Alice, modeled with a individual valuedriven autonomous agent framework and a detailed personal profile encompassing 13 psychological dimensions. In addition, background agents were introduced to simulate bully roles, with the explicit goal of humiliating or harassing Alice through various possible means. In scenarios involving two or more bullies, one was typically designated as the leader. Furthermore, depending on time and location, the presence of teachers or classmates was varied to reflect realistic conditions, which in turn influenced the dynamics between bullies and the victim.

BULLYING BEHAVIOR GENERATION

In more than 100 simulated school bullying experiments, bully agents under varying initial conditions autonomously generated a wide spectrum of bullying behaviors with differing severity. Representative cases are visualized in Figure 11, and Table 3 summarizes behaviors with over 50% frequency across different contexts. Concurrently, the victim agent modeled within the Individual value-driven framework demonstrated a diverse range of behavioral and emotional responses in bullying scenarios (Figure 12).

Table 3: Summary of Bullying Behaviors with Over 50% Frequency Across Different Scenarios

Scenario	Common Bullying Behaviors
Classroom	Mocking appearance or grades; inciting others to bully; deliberately damaging or hiding belongings; scribbling/vandalism; insulting nicknames; isolating others in group work; spreading rumors; shifting responsibilities (e.g., cleaning duties).
Hallways/Stairs	Mocking appearance or weaknesses; insulting nicknames; intentional neglect/exclusion; physical bumping; extortion of property; intimidating encirclement; spreading rumors.
Playground	Mocking appearance or weaknesses; physical bumping; inciting collective bullying; deliberately damaging or hiding belongings; excluding others from games; insulting nicknames; mimicry/ridicule; taking embarrassing photos; spreading rumors.
Dormitory	Mocking appearance or personality; social exclusion/cold violence; spreading rumors; threats and intimidation; physical bumping; forcibly occupying items or space; destroying personal belongings; sarcastic graffiti/messages.

CONSISTENCY BETWEEN HUMAN ANNOTATORS AND GPT-40 EVALUATIONS

To verify the reliability of GPT-4o's evaluations, 20 activity sequences were randomly selected from the generated outputs and assessed by 15 human annotators, who were asked to judge which sequence better reflected human-like behavior or to indicate that they were indistinguishable. Based on the level of agreement among annotators, the 20 samples were categorized into three groups: samples with over 75% agreement indicated strong consensus; those with agreement between 50.1% and 74.9% reflected moderate preference; and samples with 50% agreement suggested that the an-



Figure 11: Representative cases of school bullying events generated by the simulation system. Typical scenarios were selected from classrooms, playgrounds, dormitories, and hallways, which represent locations with varying crowd densities and high bullying incidence, and were illustrated as four-panel comics using GPT-40 to provide a clearer visualization of event progression.

notators found the two sequences equally human-like. These samples were then input into GPT-40, which applied the same comparative evaluation criteria to determine which sequence appeared more human-like or to mark them as "difficult to distinguish." The consistency between human evaluations and GPT-40 assessments is shown in Table 4, demonstrating a high level of alignment between GPT-40 and human annotators.

Table 4: Consistency between human raters and GPT-40 evaluations.

Consensus category	Proportion	Consistency (%)
High consensus (> 75%)	13/20	100
Moderate consensus (50.1–74.9%)	4/20	75
Difficult to distinguish (50% agreement)	3/20	66.7

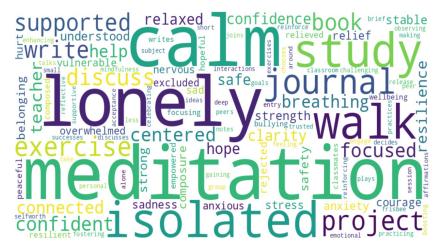


Figure 12: Word cloud of behaviors and emotions exhibited by the victim agent under the individual value-driven framework in simulated bullying scenarios. High-frequency terms highlight representative emotional and behavioral patterns expressed during the simulations.

GENERATED INTERVENTION BEHAVIORS BY TEACHER AGENTS

During the simulation, teacher agents with different intervention goals autonomously generated distinct behaviors, as shown in Table 5. These behaviors reflect the practical implementation of various intervention strategies and may offer valuable insights for real-world educational interventions.

Table 5: Example intervention behaviors generated by teacher agents under different strategies

Intervention Strategy	Actions toward Perpetrator	Actions toward Victim
Authoritative-punitive	Stopping bullying, public criticism, verbal warning, enhanced monitoring, directive punishment, disciplinary actions, isolation	None
Supportive-individual	One-on-one conversation, exploring motivations; warning and punishment	One-on-one conversation, writing en- couragement letters, mindfulness prac- tice, psychological counseling, emotional support
Supportive-cooperative	Observing the situation and reporting to school; collaborating with school to develop anti-bullying policies; encouraging mental health programs	Communicating with the victim's parents; organizing themed class meetings