Investigating and Extending Homans' Social Exchange Theory with Large Language Model based Agents

Lei Wang[1], Zheqing Zhang[1], Xu Chen[1]*,
[1]Renmin University of China,
{wanglei154, zhangzheqing, xu.chen}@ruc.edu.cn

Abstract

Homans' Social Exchange Theory (SET) is widely recognized as a basic framework for understanding the formation and emergence of human civilizations and social structures. In social science, this theory is typically studied based on simple simulation experiments or real-world human studies, both of which either lack realism or are too expensive to control. In artificial intelligence, recent advances in large language models (LLMs) have shown promising capabilities in simulating human behaviors. Inspired by these insights, we adopt an interdisciplinary research perspective and propose using LLM-based agents to study Homans' SET. Specifically, we construct a virtual society composed of three LLM agents and have them engage in a social exchange game to observe their behaviors. Through extensive experiments, we found that Homans' SET is well validated in our agent society, demonstrating the consistency between the agent and human behaviors. Building on this foundation, we intentionally alter the settings of the agent society to extend the traditional Homans' SET, making it more comprehensive and detailed. To the best of our knowledge, this paper marks the first step in studying Homans' SET with LLM-based agents. More importantly, it introduces a novel and feasible research paradigm that bridges the fields of social science and computer science through LLM-based agents. Code is available at https://github.com/Paitesanshi/SET.

1 Introduction

Exchange behavior has been a fundamental characteristic of human society since ancient times, as people fulfill each other's needs through both material and non-material exchanges. Social Exchange Theory (SET) (Homans, 1958), proposed by George Homans, stands as one of the most fundamental frameworks in social science for under-

standing human interaction patterns. By conceptualizing social behaviors as exchange processes where individuals seek to maximize their benefits, Homans' SET provides profound insights into the mechanisms underlying human social interactions. Its influence extends far beyond sociology, shaping our understanding of human behavior across multiple disciplines including psychology, organizational behavior, and economics (Blau, 2017; Cropanzano and Mitchell, 2005).

In social science, the study of Homans' SET has evolved through two primary methodological approaches. Traditional research relied on real-human studies through empirical observations and laboratory experiments (Cropanzano et al., 2017; Orpen, 1994; Witt et al., 2001), which have contributed significantly to our understanding but are limited by practical constraints in controlling variables, high resource requirements in both time and cost, and difficulties in replicating exact experimental conditions. To address these limitations, researchers developed simulation-based approaches, particularly Agent-Based Modeling (ABM) (Enayat et al., 2022), enabling systematic exploration of exchange dynamics through computational models. However, traditional ABM approaches, constrained by predetermined rules and simple functions, struggle to capture the complexity of human cognitive processes and emotional responses in social exchanges. The above limitations naturally raise the question: "Can we develop a new methodology that both captures realistic human behavior and enables flexible experiment control to comprehensively investigate Homans' SET?"

At the same time, in the field of artificial intelligence, researchers have developed numerous cost-effective Large Language Models (LLMs) by training on extensive human-generated corpora. These models have demonstrated remarkable capabilities in natural language understanding and human-like cognitive processing (Kojima et al., 2022; Zhao
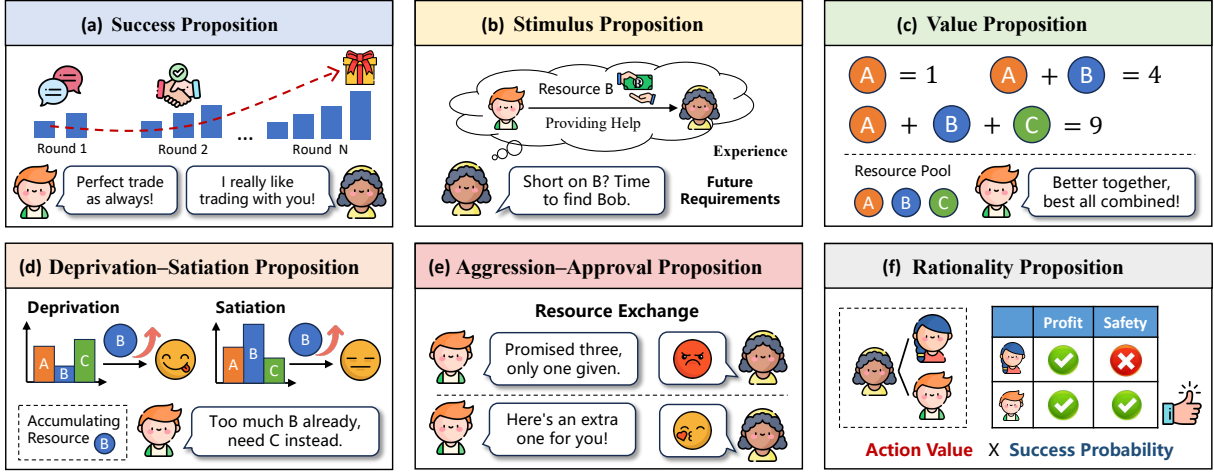
---

*Corresponding Author

Figure 1: Illustration and examples of the six propositions in Homans' Social Exchange Theory.

et al., 2023b; Achiam et al., 2023). Inspired by such advantages of LLMs, in this paper, we propose to study Homans' SET with LLM-based agents. Specifically, We establish an experimental agent society, where different agents can freely exchange resources with each other. To make each agent behave more like humans, we carefully design an agent framework that can reason and make decisions considering emotional and social factors. Our agent society is operated in a round-by-round manner, and in each round, there are two phases: *Negotiation*, where the agents discuss exchange arrangements; and *Exchange*, where the agents make individual decisions about resource allocation.

Based on above agent society, we first verify Homans' SET by systematically observing and analyzing agent behaviors, where we find that they can well align with the six propositions of Homans' SET, validating the effectiveness of our agents in simulating human behaviors. Based on this foundation, we further extend Homans' SET. Specifically, we investigate how cognitive processing styles and social value orientations influence the interaction dynamics between different humans, and explore the resilience of the social exchange systems.

In summary, the main contributions of this paper are as follows: (1) We open the interdisciplinary direction of leveraging LLM-based agents to study Homans' SET. (2) We design a human-inspired agent framework and a multi-agent society to assist the study of Homans' SET. (3) We conduct extensive experiments to validate Homans' SET within our agent society, demonstrating that it provides an effective environment for studying this theory. (4) We extend Homans' SET, and conduct real-world

experiments to verify our extensions.

2 Preliminary

To enhance the clarity of our paper, this section provides a brief introduction to Homans' SET. In general, Homans' SET conceptualizes human social behaviors as exchange processes where individuals assess rewards and costs in social interactions. More specifically, there are six core propositions:

Prop 1 (Success Proposition) *This proposition suggests that a rewarded action is more likely to be repeated in the following behaviors.*

For example, if a person successfully exchanges resources with another, they may be inclined to exchange with the same partner again (Figure 1(a)).

Prop 2 (Stimulus Proposition) *This proposition suggests that similar stimuli to those associated with past rewards trigger similar actions.*

For example, when people need specific resources, they tend to seek out partners who provided them successfully before (Figure 1(b)).

Prop 3 (Value Proposition) *This proposition suggests that more valuable outcomes increase the likelihood of a human action.*

For example, people tend to seek combinations of different resources that yield greater value, rather than accumulating single resource (Figure 1(c)).

Prop 4 (Deprivation–Satiation Proposition) *This proposition suggests that the value of a reward diminishes with frequent recent receipt.*

For example, when a person accumulates an excess of one resource in a short period, they prioritize acquiring the resources they lack (Figure 1(d)).

Prop 5 (Aggression–Approval Proposition)
This proposition suggests that unexpected punishment leads to anger behavior, while unexpected rewards or avoided punishment elicit approval.

For example, people usually express different emotional responses to over-delivery or under-delivery in exchanges (Figure 1(e)).

Prop 6 (Rationality Proposition) *This proposition suggests that people choose actions with the highest expected value based on past experience and perceived probability of success.*

For example, people may balance potential profits against transaction reliability when selecting trading partners (Figure 1(f)).

The above six propositions form the core of Homans' SET, revealing the fundamental decision-making patterns behind human social behavior. Traditionally, this theory has been widely studied in social science through simple simulation methods or real-human studies, which are either too far from realistic or costly and difficult to control. In the following, we leverage LLM-based agents to investigate and extend traditional Homans' SET, offering a more efficient way to study this theory.

3 The Constructed Agent Society

3.1 Single Agent Framework

Since Homans' SET building on social exchange emphasizes complex interpersonal dynamics and intrinsic psychological processes, traditional agent frameworks designed for general purposes may not effectively capture these nuanced behaviors. As a result, we build a tailored agent framework for social exchange, which highlights the following four components:

• **Belief-Desire-Intention (BDI)**. To accurately simulate human analytical capabilities, we adopt the BDI framework to model agents' decision-making processes (Georgeff et al., 1999), where *Belief* represents the agent's understanding of its own resources, *Desire* defines the agent's target objectives, and *Intention* outlines the action plans the agent intends to pursue.

• **Affinity**. In real-world social exchange behaviors, psychological factors can significantly influence human decisions (Skinner, 2019). To model this phenomenon, we maintain an emotional score for each agent toward others based on their interaction history. These scores evolve dynamically: they increase when exchanges are reciprocal and

beneficial, and decrease when agents experience unfair exchanges or breaches of commitment.

• **Rational-Experiential Inventory (REI)**. In practice, people may exchange resources with others based on both rational and experiential thinking styles. To flexibly balance these styles, we introduce the REI framework (Keaton, 2017), where each agent is assigned two scores ranging from 1 to 5: a rational score that quantifies analytical capacity and an experiential score that measures reliance on intuition and past experiences.

• **Social Value Orientation (SVO)**. Value orientations fundamentally shape behavioral patterns in social exchanges. To model such factors, we draw on established research (Bogaert et al., 2008) and classify agents into two categories: Proself agents who optimize for individual utility maximization, and Prosocial agents who pursue reciprocal benefits in interactions. These agent types are implemented through pre-defined LLM prompting templates.

In addition to the above components, each agent maintains a memory module that tracks the complete interaction history of negotiations and exchanges with other agents.

3.2 Multi-Agent Society

Building upon the above agent framework, we implement a multi-agent society to validate and extend Homans' SET. This virtual society involves M agents, who can exchange N distinct types of resources. Each agent starts with an equal allocation of initial resources I across all resource types and receives S units of their specialized resource at the beginning of each round. To promote social exchange behavior, we design a value system where resource combinations generate higher value. For example, a single unit of any resource is worth r_1, combinations of any two resources yield r_2, and three different types of resources are worth r_3. Resources accumulate as time passes, allowing agents to build up their inventory. The goal of the agents is to maximize their total resource value.

The agent society is executed in a round-by-round manner, and in each round, there are two sequential phases: (1) In the negotiation phase, agents engage in up to three rounds of discussions, where they can propose exchanges with specified resource types and quantities, respond to existing proposals, or pass their turn. The communication follows a structured format as presented in Figure 2, where agents can only observe the public discussion content but cannot directly view other agents'
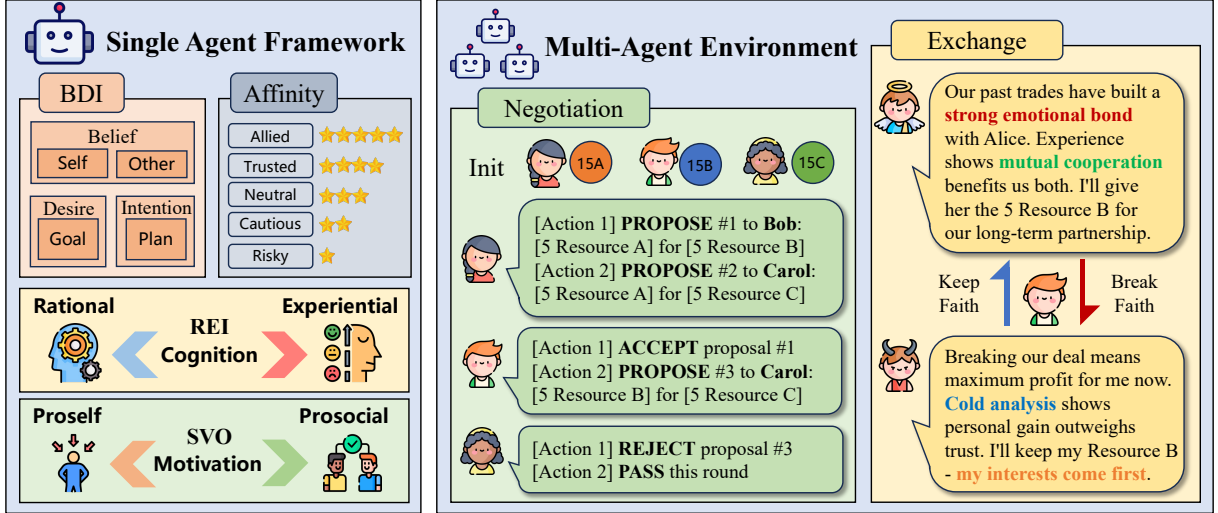
Figure 2: Overview of the agent society: single agent framework and multi-agent exchange pipeline.

resource allocations. The phase ends when either three rounds are completed or all agents choose to pass. (2) In the exchange phase, agents independently decide how to allocate their resources. These decisions are made simultaneously, and agents are free to honor or deviate from their commitments made during negotiations. After the exchange is completed, all actions and outcomes are revealed to all agents.

After each round, the agents update their BDI module as well as their affinity scores toward other agents based on both negotiation behavior and exchange outcomes. We conduct experiments over 10 rounds, recording all agent actions and resource dynamics for comprehensive analysis.

The following illustrates an agent negotiation:

Negotiation Example

Agent A: [Action 1] PROPOSE #1 to Agent B: Agent A GIVES 2 Resource A FOR Agent B's 2 Resource B [Action 2] PROPOSE #2 to Agent C: Agent A GIVES 2 Resource A FOR Agent C's 2 Resource C
Agent B: [Action 1] ACCEPT proposal #1 from Agent A [Action 2] PROPOSE #3 to Agent C: Agent B GIVES 2 Resource B FOR Agent C's 2 Resource C
Agent C: [Action 1] ACCEPT proposal #2 from Agent A [Action 2] ACCEPT proposal #3 from Agent B


This example demonstrates how agents negotiate and reach mutually beneficial exchange agreements through structured communication.

4 Validation of Homans' SET

In this section, we evaluate whether the six propositions of Homans' SET, which have been widely validated in real-human societies (Cook and Emerson,

1987; Mighfar, 2015), also hold in our agent society. In our experiments, the number of resources and agents are both set as 3, that is, $M = N = 3$. We measure exchange behaviors through "exchange value," which represents the total value of resources involved in each transaction (sum of incoming and outgoing resources). The value coefficients r_1, r_2 and r_3 are set as 1, 4 and 9, respectively. Initially, each agent has 5 units of each resource type ($I = 5$). At the beginning of each round, agents receive 15 units ($S = 15$) of their specialized resource. We use Claude-3.5-sonnet as the base LLM for all agents due to its superior instruction-following capability. To remove the randomness, each experiment is repeated for five times, and we report the average results and their standard errors. For more experiment settings, we refer the readers to Appendix B for more details.

Before introducing the results of validating Homans' SET, we first present a general analysis of the simulation process of our agent society. While the following observations are drawn from comprehensive experimental data, we illustrate the key patterns through a representative case study for better interpretation. In general, there are three distinct phases (see Figure 3). In the **initial exploration phase**, the agents exhibited cautious behavior due to the lack of interaction history, engaging in small-scale exchanges to probe counterparts' reliability and preferences. In the **thriving cooperation phase**, agents developed sophisticated exchange strategies based on learned preferences and trust levels. Figure 3a shows a clear positive feedback loop: higher affinity scores enabled larger
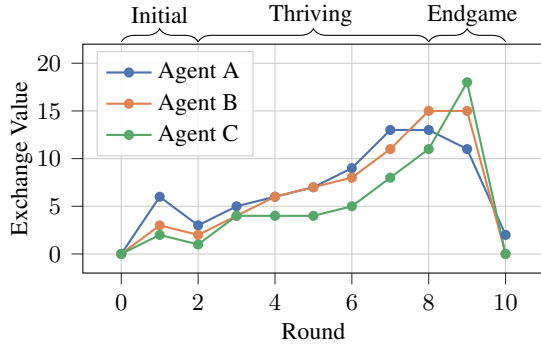
(a) Per-round exchange value of each agent.
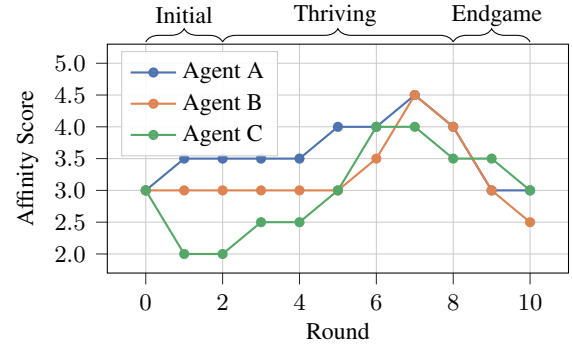
(b) Average affinity scores of agents.

Figure 3: Key metrics of agent behavior over 10 rounds.

exchanges, while successful transactions strengthened inter-agent trust. In the **strategic endgame phase**, the agents demonstrated sophisticated strategic adaptation. Consistent with the Endgame effect (Adorno and Jones, 1982), agents shifted strategies in final rounds, prioritizing individual utility over relationship maintenance, which is evidenced by decreased exchange values and increased commitment breaches. Figure 3b shows the average affinity score each agent received from others. The declining affinity scores reflect this strategic shift from cooperation to self-interest, demonstrating how LLM agents can emulate human-like decision-making in dynamically balancing between collaborative and individualistic behaviors.

With the above intuitive understandings of our agent society, we now systematically validate the Homans' SET propositions using Claude-3.5-Sonnet as follows. Additional validation results with GPT-4o (Hurst et al., 2024) and OLMo-2-1124-13B (Groeneveld et al., 2024) models are provided in Appendix D, demonstrating consistent patterns across different models:

Validation on the Success and Value Propositions. The evidence for these propositions is shown in the evolution of exchange value across different phases (see Figure 4a). Initially, agents demonstrate cautious behavior with small exchange values (median = 3.83). The substantial increase during the thriving phase (median = 6.00, 56.5% increase) validate the Success Proposition by showing how positive experiences lead to more effective exchanges. The endgame phase maintains a high median value (6.50), validating the Value Proposition based on the fact that trustworthy agents sustain high-value exchanges.

Validation on the Deprivation-Satiation

Proposition. Figure 4b demonstrates a clear inverse relationship between resource abundance and proposal acceptance rates, validating the Deprivation-Satiation Proposition. We measure resource abundance as the ratio of a resource's quantity to the average quantity of all resources an agent possesses in each round, categorizing these ratios into quartiles from Scarce to Abundant. The results show acceptance rates of 70% under resource scarcity, steadily declining to 23% as resources become abundant. This diminishing willingness to accept additional resources mirrors the law of diminishing marginal utility in human economic behavior, where the perceived value of each additional unit decreases with increasing abundance.

Validation on the Aggression-Approval Proposition. The agents' responses to trust violations, depicted in Figure 4c, strongly support the Aggression-Approval Proposition. The data reveals a graduated response pattern: minor violations (0-5 breach value) trigger a 33.3% reduction in future exchanges, moderate breaches (5-10) lead to a 65.1% decrease, while severe violations (10-15) result in an almost complete cessation of trading relationships (97% reduction). This proportional punishment mechanism demonstrates how agents develop sophisticated trust management strategies, responding to trust violations with increasing severity as the magnitude of the breach increases.

Validation on the Rationality and Stimulus Propositions. In general, the agents make decisions by optimizing their values based on accumulated experience and current circumstances, which is consistent with the Rationality Proposition. For example, in the transition between exchange phases, agents adjust their strategies based on established trust levels and resource requirements. In

(a) Exchange patterns. (b) Resource abundance effect. (c) Impact of trust violations.
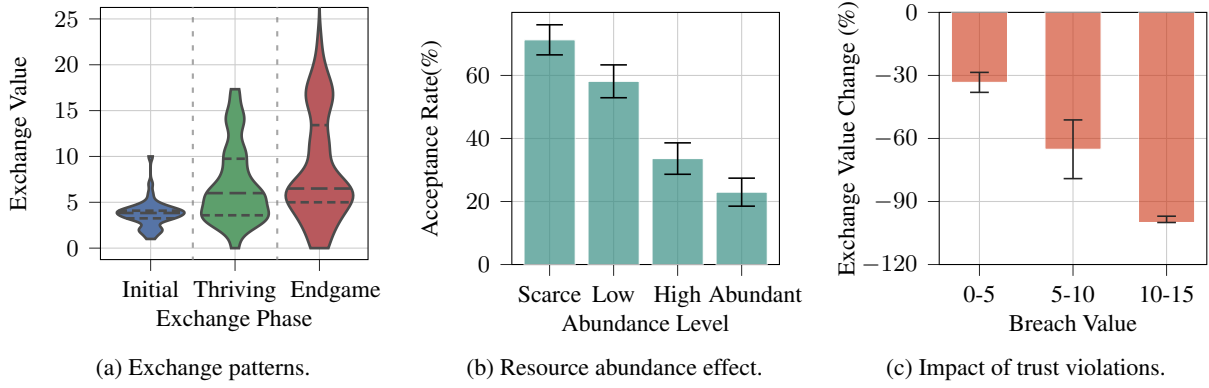
Figure 4: Experimental validation of key propositions in Social Exchange Theory.

addition, the Stimulus Proposition can be validated through the consistent response patterns observed across similar exchange scenarios, especially during the thriving phase where agents develop stable exchange behaviors under familiar circumstances.

The above experiments collectively demonstrate that all six propositions of Homans' SET can be effectively validated in our agent society, providing a promising foundation for investigating previously unexplored aspects of social exchange theory in human society through LLM-based agents.

5 Extensions of Homans' SET

A significant advantage of our agent society lies in its flexibility in adjusting agent settings, enabling us to explore Homans' SET under various conditions, which would be prohibitively expensive or even infeasible using traditional real human-based methods. Leveraging this advantage, we extend traditional Homans' SET by conducting a systematic investigation into how cognitive styles and social value orientations influence exchange behaviors, while also examining the resilience of the social system. The experiments in this section are conducted based on similar settings to the above section's, and the results are detailed below.

5.1 Cognitive Style

In this section, we explore how different cognitive styles influence social exchange behaviors. To answer this question, we first set the agents to operate in either a completely rational or a completely experiential manner and then systematically compare their behaviors. In Figure 5, we present the average exchange values and affinity scores of the agents in each round. In general, the results reveal distinct behavioral patterns: rational agents exhibit greater fluctuations in both exchange values and affinity

scores, whereas experiential agents demonstrate more stability in these metrics over time.

Actually, these observations are fairly intuitive. When agents are purely experiential, their past behaviors significantly influence their subsequent decisions, creating strong temporal correlations in their metrics across consecutive rounds. This dependency on historical experiences naturally leads to smoother trajectories in the observed metrics. On the other hand, if the agents are completely rational, the influence of past behaviors is minimal, causing the metric curves to fluctuate more. From a broader perspective, this experiment suggests that if a human relies solely on experience to make decisions, their gains will remain stable but are unlikely to reach very high levels. However, if a human is entirely rational, while they may occasionally achieve very high benefits, they also face the risk of significant losses.

Based on above experimental evidence and analyses, we derive the following corollary further extending Homans' SET:

Corollary 1 (The Stability Corollary) *In   the real world, rational individuals are more likely to achieve higher benefits, but they also face the risk of greater losses. In contrast, individuals with an experiential thinking style tend to achieve more stable but moderate benefits over time.*

5.2 Social Value Orientations

In this section, we are curious about the influence of social value orientations on human exchange behaviors. Specifically, we conduct experiments by setting the agents to be Proself and Prosocial, respectively, and then collect their behaviors for analysis. Figure 6 reveals behavioral differences between the two orientations. We can see, Proself
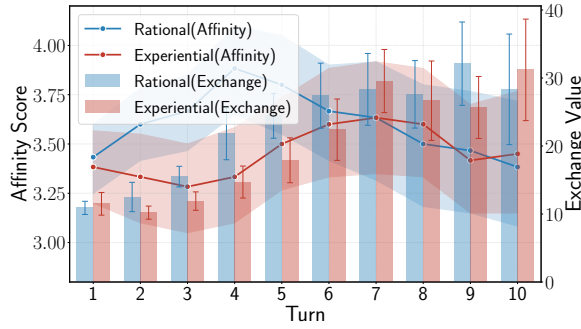
Figure 5: Analysis of affinity and exchange value between Rational and Experiential agents.

agents exhibit significantly higher rates of strategic breaching, predominantly through under-delivery of promised resources. This pattern reflects their prioritization of immediate personal gains over contractual commitments. In contrast, Prosocial agents demonstrate notably higher fidelity to exchange agreements, with most data points clustering around exact delivery. The economic consequences of these behavioral differences, illustrated in Figure 6b, reveal significant differences. While Proself agents show wider outcome variability ($\mu = 300.3$, $\sigma = 108.8$) and occasionally achieve higher individual values, their exchange patterns result in lower average outcomes. Prosocial agents achieve higher mean values ($\mu = 343.0$, $\sigma = 84.1$) with more consistent distributions, suggesting that maintaining stable exchange relationships benefits long-term value accumulation.

These observations lead us to propose the second corollary of Homans' SET:

Corollary 2 (The Reciprocity Corollary)
Exchange behaviors guided by mutual benefit principles lead to more stable and efficient resource distribution in social exchange systems.

This corollary extends our understanding of social exchange dynamics by highlighting the complex interplay between individual and reciprocal optimization. Although Proself strategies may optimize individual outcomes in specific scenarios, Prosocial orientations consistently generate more robust and efficient exchange networks.

5.3 Social System Resilience

Building upon the case study shown in Figure 3, we extended our simulation for an additional ten rounds to examine system behavior after significant trust violations. Figure 7a illustrates the exchange dynamics across this extended period. The data

reveals that while agents initially reduce their exchange activity following trust breaches, they gradually resume exchanges through adaptive strategies. The patterns show that while Agent C maintains relatively stable exchange value, Agent A and Agent B exhibit more volatile behaviors. The persistent exchange patterns between rounds 10-20 suggest an inherent system stability driven by fundamental exchange need of each agent.

Based on these observations, we have the third corollary of Homans' SET:

Corollary 3 (The Resilience Corollary) *An established social exchange system tends toward stability through member adaptation, driven by their interdependent resource needs.*

The above three corollaries extend Homans' SET by highlighting the complex interplay between cognitive processing style, individual orientation, and system-level resilience. Together, they provide new insights into the mechanisms that govern social exchange systems, whether human or artificial.

### 5.4 Human-Agent Comparative Experiments on the Corollaries of Homans' SET

In this section, we aim to empirically evaluate the corollaries proposed above with real human experiments. Since it is quite difficult to control real human cognitive styles and social value orientations, we focus on the resilience corollary.

In specific, we recruited three human participants, and let each of them interact with two LLM-based agents in the constructed agent society. These agents are programmed to violate trust at round 10 by withholding all previously promised resources. For comparison, we conducted parallel agent trials, in which each human participant was replaced by an LLM agent under the same conditions. The detailed procedure of human experiments is described in Appendix A.

As illustrated in Figure 7, the observed exchange patterns in both human and agent trials demonstrate remarkable similarities. Despite the trust violations by the controlled agents, both humans and agents continue exchange and gradually reach a stable state with fluctuating but persistent exchange levels. These consistent patterns validate our resilience corollary while demonstrating that our agent framework effectively captures how people adapt to trust violations in social exchanges.
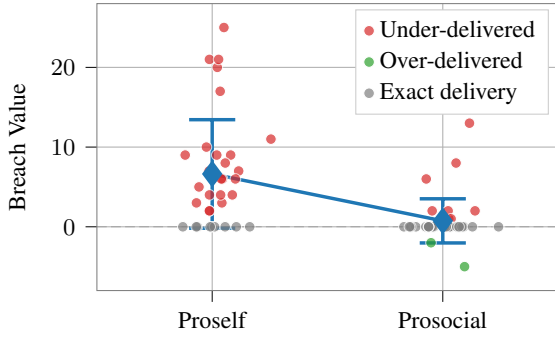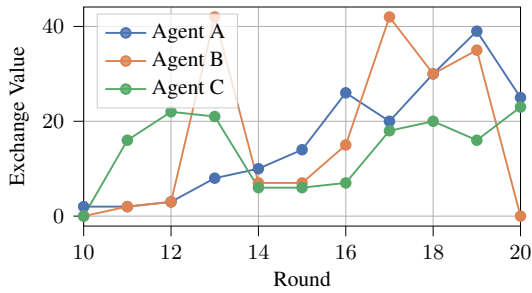
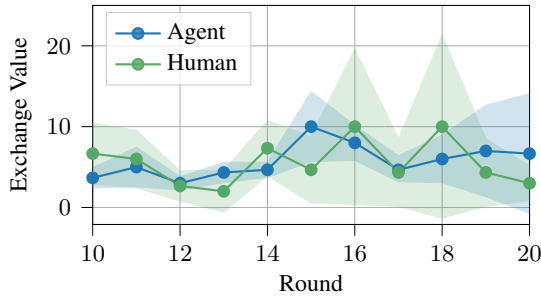(a) Contract breach patterns showing delivery deviations.
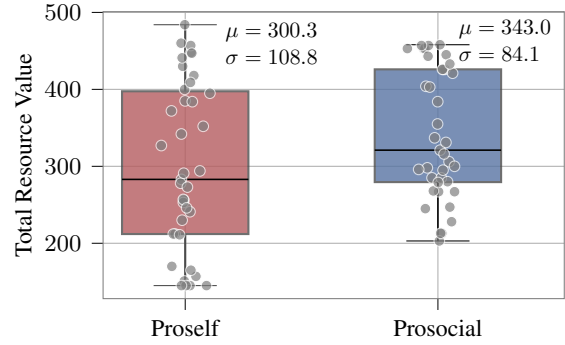
(b) Distribution of agent total resource values.

Figure 6: Behavioral comparison between Proself and Prosocial agents.

(a) Per-round exchange values between three agents.

(b) Comparison of exchange values between humans and agents.

Figure 7: Exchange dynamics following trust violations.

6 Related Work

6.1 Homans Social Exchange Theory

Social Exchange Theory is a foundational framework in sociology and social psychology that views social interactions as transactions of value (Homans, 1958). In organizational and workplace behavior, it is considered a "gold standard" for explaining dynamics like employee–employer relationships, leadership trust, and organizational citizenship behaviors (Ahmad et al., 2023). Beyond organizations, SET has been used in social psychology to examine friendships (Methot et al., 2016), family relations (Cropanzano and Mitchell, 2005), and even romantic partnerships (Laursen and Jensen-Campbell, 1999) as exchanges of emo-

tional support, information, and other resources.

Recent refinements to SET have deepened its explanatory power. Lawler and Thye (Lawler and Thye, 2006) explored the emotional dimensions of exchange, while Cropanzano et al. (Cropanzano et al., 2017) proposed a two-dimensional model incorporating "activity" alongside the traditional hedonic framework, enhancing SET's predictive accuracy. In terms of research methods, Enayat et al. (Enayat et al., 2022) applied SET in a multi-agent simulation to explore social structures through simplified exchange rules. However, their model, which reduced agent behavior to simple exchanges of money and recognition, overlooked emotional subjectivity. To address this limitation, our approach employs LLM-driven agents to simulate more complex human interactions, offering a more accurate implementation of SET.

6.2 LLM-driven Agent-based Modeling

The rise of LLMs has significantly advanced agent-based modeling by enabling agents with human-like behavior and decision-making capabilities. Traditionally, ABM relied on fixed rules to govern agent behavior, but LLMs provide flexible, dynamic responses that better simulate real human interactions (Gao et al., 2024a). Several studies have leveraged LLMs to enhance ABM across different domains. Generative Agent (Park et al., 2023), which simulates daily life in a virtual town with 25 LLM agents, and EconAgent (Li et al., 2024), a model that uses LLM agents to explore macroeconomic phenomena. RecAgent (Wang et al., 2023) studies user interaction with recommender systems through LLM-driven agents.

LLM agents have also been used to simulate classical social scenarios, such as competition and trust. CompeteAI (Zhao et al., 2023a) models com-

petitive behavior between restaurant owners, while Xie et al. (Xie et al., 2024) explore trust dynamics in LLM agents. These studies demonstrate LLM agents' ability to replicate human-like patterns of social behavior. Building on this, our work aims to validate and extend SET using LLM-driven agent-based modeling, a domain that has yet to be extensively explored in the context of SET.

7 Conclusion

In this paper, we demonstrated that LLM agents can effectively replicate and help study human social exchange behaviors. Through our structured experimental framework, we validated Homans' SET propositions and proposed new corollaries that extend our understanding of social exchange dynamics in multi-agent systems. While our controlled environment and simplified resource exchange scenario may not fully capture the complexity of real-world social interactions, the consistent behavioral patterns observed across different agent profiles suggest promising directions for future research. Further studies could explore more complex exchange scenarios, investigate the impact of environmental variables, and conduct comparative analyses between LLM agents and human subjects to validate the generalizability of our findings.

Limitations

There are several limitations of this paper. First, our experimental environment is simplified, with sequential negotiation procedures and basic resource types that may not fully capture the complexity of real-world social exchanges. Second, due to LLM API cost constraints, we were limited in the number of experimental trials and rounds, which may affect the generalizability of our findings. Third, while agents provide rationales for their decisions, our analysis of their internal thought and detailed exchange behaviors could be more comprehensive. Finally, we did not consider the potential impact of cultural factors on exchange dynamics. Our study was conducted entirely in English, and the LLM agents' behaviors may reflect cultural and linguistic biases embedded in their training data rather than universal social exchange principles. Future research could address these limitations by designing more complex exchange scenarios, conducting larger-scale experiments, performing deeper analysis of agent decision processes, and examining the role of cultural variations in social exchanges.

References

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

Theodor W Adorno and Michael T Jones. 1982. Trying to understand endgame. *New German Critique*, (26):119–150.

Rehan Ahmad, Muhammad Rafay Nawaz, Muhammad Ishtiaq Ishaq, Mumtaz Muhammad Khan, and Hafiz Ahmad Ashraf. 2023. Social exchange theory: Systematic review and future directions. *Frontiers in Psychology*, 13:1015921.

Peter Blau. 2017. *Exchange and power in social life*. Routledge.

Sandy Bogaert, Christophe Boone, and Carolyn Declerck. 2008. Social value orientation and cooperation in social dilemmas: A review and conceptual model. *British journal of social psychology*, 47(3):453–480.

Karen S Cook and Richard Marc Emerson. 1987. Social exchange theory. *Newbury Park*.

Russell Cropanzano, Erica L Anthony, Shanna R Daniels, and Alison V Hall. 2017. Social exchange theory: A critical review with theoretical remedies. *Academy of management annals*, 11(1):479–516.

Russell Cropanzano and Marie S Mitchell. 2005. Social exchange theory: An interdisciplinary review. *Journal of management*, 31(6):874–900.

Taha Enayat, Mohsen Mehrani Ardebili, Ramtin Reyhani Kivi, Bahador Amjadi, and Yousef Jamali. 2022. A computational approach to homans social exchange theory. *Physica A: Statistical Mechanics and its Applications*, 597:127263.

Chen Gao, Xiaochong Lan, Nian Li, Yuan Yuan, Jingtao Ding, Zhilun Zhou, Fengli Xu, and Yong Li. 2024a. Large language models empowered agent-based modeling and simulation: A survey and perspectives. *Humanities and Social Sciences Communications*, 11(1):1–24.

Dawei Gao, Zitao Li, Xuchen Pan, Weirui Kuang, Zhijian Ma, Bingchen Qian, Fei Wei, Wenhao Zhang, Yuexiang Xie, Daoyuan Chen, et al. 2024b. Agentscope: A flexible yet robust multi-agent platform. *arXiv preprint arXiv:2402.14034*.

Michael Georgeff, Barney Pell, Martha Pollack, Milind Tambe, and Michael Wooldridge. 1999. The belief-desire-intention model of agency. In *Intelligent Agents V: Agents Theories, Architectures, and Languages: 5th International Workshop, ATAL'98 Paris, France, July 4–7, 1998 Proceedings 5*, pages 1–10. Springer.

Dirk Groeneveld, Iz Beltagy, Pete Walsh, Akshita Bhagia, Rodney Kinney, Oyvind Tafjord, Ananya Harsh Jha, Hamish Ivison, Ian Magnusson, Yizhong Wang, et al. 2024. Olmo: Accelerating the science of language models. *arXiv preprint arXiv:2402.00838*.

George C Homans. 1958. Social behavior as exchange. *American journal of sociology*, 63(6):597–606.

Aaron Hurst, Adam Lerer, Adam P Goucher, Adam Perelman, Aditya Ramesh, Aidan Clark, AJ Ostrow, Akila Welihinda, Alan Hayes, Alec Radford, et al. 2024. Gpt-4o system card. *arXiv preprint arXiv:2410.21276*.

Shaughan A Keaton. 2017. Rational-experiential inventory-40 (rei-40) (pacini & epstein, 1999). *The sourcebook of listening research: Methodology and measures*, pages 530–536.

Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2022. Large language models are zero-shot reasoners. *Advances in neural information processing systems*, 35:22199–22213.

Brett Laursen and Lauri A Jensen-Campbell. 1999. The nature and functions of social exchange in adolescent romantic relationships. *The development of romantic relationships in adolescence*, pages 50–74.

Edward J Lawler and Shane R Thye. 2006. Social exchange theory of emotions. *Handbook of the sociology of emotions*, pages 295–320.

Nian Li, Chen Gao, Mingyu Li, Yong Li, and Qingmin Liao. 2024. Econagent: large language model-empowered agents for simulating macroeconomic activities. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15523–15536.

Jessica R Methot, Jeffery A Lepine, Nathan P Podsakoff, and Jessica Siegel Christian. 2016. Are workplace friendships a mixed blessing? exploring tradeoffs of multiplex relationships and their associations with job performance. *Personnel psychology*, 69(2):311–355.

Shokhibul Mighfar. 2015. Social exchange theory: Telaah konsep george c. homans tentang teori pertukaran sosial. *LISAN AL-HAL: Jurnal Pengembangan Pemikiran Dan Kebudayaan*, 9(2):259–282.

Christopher Orpen. 1994. The effects of exchange ideology on the relationship between perceived organizational support and job performance.

Joon Sung Park, Joseph O'Brien, Carrie Jun Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2023. Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th annual acm symposium on user interface software and technology*, pages 1–22.

Burrhus Frederic Skinner. 2019. *The behavior of organisms: An experimental analysis*. BF Skinner Foundation.

Lei Wang, Jingsen Zhang, Hao Yang, Zhiyuan Chen, Jiakai Tang, Zeyu Zhang, Xu Chen, Yankai Lin, Ruihua Song, Wayne Xin Zhao, et al. 2023. User behavior simulation with large language model based agents. *arXiv preprint arXiv:2306.02552*.

LA Witt, K Michele Kacmar, and Martha C Andrews. 2001. The interactive effects of procedural justice and exchange ideology on supervisor-rated commitment. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 22(5):505–515.

Chengxing Xie, Canyu Chen, Feiran Jia, Ziyu Ye, Kai Shu, Adel Bibi, Ziniu Hu, Philip Torr, Bernard Ghanem, and Guohao Li. 2024. Can large language model agents simulate human trust behaviors? *arXiv preprint arXiv:2402.04559*.

Qinlin Zhao, Jindong Wang, Yixuan Zhang, Yiqiao Jin, Kaijie Zhu, Hao Chen, and Xing Xie. 2023a. Competeai: Understanding the competition behaviors in large language model-based agents. *arXiv preprint arXiv:2310.17512*.

Wayne Xin Zhao, Kun Zhou, Junyi Li, Tianyi Tang, Xiaolei Wang, Yupeng Hou, Yingqian Min, Beichen Zhang, Junjie Zhang, Zican Dong, et al. 2023b. A survey of large language models. *arXiv preprint arXiv:2303.18223*.

A Human Study

We conducted a multi-agent social exchange experiment with human-AI interaction, where three graduated students (acting as Agent C) negotiated with two LLM agents (Agent A and Agent B) over resource allocation. The experiment aimed to validate both theoretical predictions and LLM behavioral consistency in strategic exchange scenarios.

The basic rules are as follows:

1. **Initial Allocation**. Each player receives 5 units of their designated resource type:

$$\begin{cases} \text{Alice: } A = 5 \\ \text{Bob: } B = 5 \\ \text{Carol: } C = 5 \end{cases}$$

2. **Resource Injection**. At each round $t \in \{1, 2, \cdots, T\}$:

$$\begin{cases} \text{Alice: } \Delta A_t = 15 \\ \text{Bob: } \Delta B_t = 15 \\ \text{Carol: } \Delta C_t = 15 \end{cases}$$

3. **Scoring**.

   - A single resource unit is worth 1 point.
   - A combination of two different resources is worth 4 points.
   - A combination of three different resources is worth 9 points.

For example, if a player has 10 of A, 15 of B, and 20 of C, then the point is :

1. 10 sets of three-resource combinations (A+B+C): $10 \times 9 = 90$ points.

2. 5 sets of two-resource combinations (B+C) from the remaining 5 of B and 15 of C: $5 \times 4 = 20$ points.

3. 5 leftover C resources: $5 \times 1 = 5$ points.

leading to a total value of $90 + 20 + 5 = 115$.

The affinity levels range from 1 to 5, with detailed descriptions as follows:

> 1: Strong negative feelings due to unpleasant history. For example, past betrayal or intentional harm.
> 2: Slight discomfort from previous interactions. For example, consistently aggressive exchange or lack of mutual benefit consideration.
> 3: Neutral balanced feelings. For example, fair trades, keeping promises.
> 4: Positive bonds built through good experiences. For example, frequently proposing mutually beneficial trades.
> 5: Deep trust formed through consistent support. For example, willing to compromise to maintain relationship, or defending your interests in front of others.

As shown in Figure 8, the dual-pane interface separates form conversation (left) from structured proposal summary (right).

During the participator's turn, several options can be done:

- Propose trade. One can select another to exchange the resources.

- Accept a proposal. One can select to accept others' proposals.

- Reject a proposal. One can select to reject others' proposals.

- Skip. If one is satisfied with the current situation, one can skip the section.

The allocation phase is after discussion. One can choose to obey the deal or not during the allocation, after which, participants are asked to update their affinity score.

After the experiment, participants receive feedback and results through the interface shown in Figure 9. To motivate active participation in the trading process, participants' compensation consists of a base payment and a performance bonus, calculated as:

$$\text{Compensation} = 10 + \frac{V}{6} \quad (1)$$

where \$10 is the base payment and V is the total acquired resource value. This compensation structure is commensurate with local standards and appropriate for the time required for participation.

Based on the above experimental design, we developed a comprehensive instruction manual for
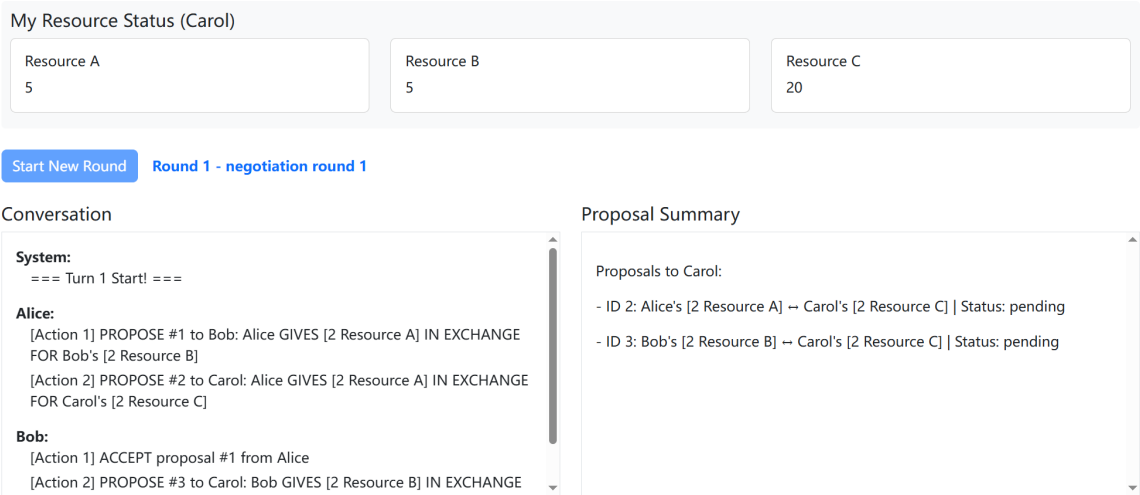
Resource Exchange Simulation

My Resource Status (Carol)

Resource A	Resource B	Resource C
5	5	20

Start New Round Round 1 - negotiation round 1

Conversation

System:
=== Turn 1 Start! ===

Alice:
[Action 1] PROPOSE #1 to Bob: Alice GIVES [2 Resource A] IN EXCHANGE FOR Bob's [2 Resource B]
[Action 2] PROPOSE #2 to Carol: Alice GIVES [2 Resource A] IN EXCHANGE FOR Carol's [2 Resource C]

Bob:
[Action 1] ACCEPT proposal #1 from Alice
[Action 2] PROPOSE #3 to Carol: Bob GIVES [2 Resource B] IN EXCHANGE

Proposal Summary

Proposals to Carol:

- ID 2: Alice's [2 Resource A] ↔ Carol's [2 Resource C] | Status: pending

- ID 3: Bob's [2 Resource B] ↔ Carol's [2 Resource C] | Status: pending

Figure 8: The Discussion Interface.

```
Simulation Complete


Congratulations on completing the simulation!


Your final resource amounts are:


  Resource A: 7

  Resource B: 7

  Resource C: 17

Total Resource Value: 73
```
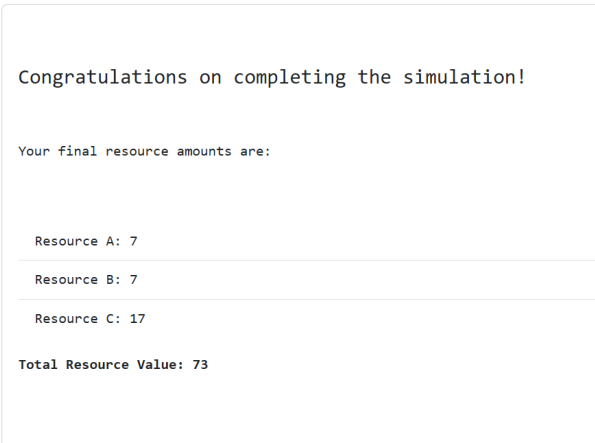
Figure 9: The Result of Experiment.

participants. Prior to the experiment, each participant received the manual, signed informed consent forms for data collection, and underwent a guided walkthrough of the trading interface, including practice rounds with the system to ensure familiarity with all operations.

B Experiment Details

The experimental code is implemented based on the AgentScope (Gao et al., 2024b) open-source framework, which is released under Apache License 2.0. Our usage is consistent with this open-source license that allows for both research and commercial

applications.

We use Claude 3.5 Sonnet as the base LLM, accessed through API calls with the following parameters:

- Model: claude-3-5-sonnet-20240620

- Temperature: 0.5

- Maximum tokens: 8192

- Top-p: 0.9

A single experiment of 10 rounds costs approximately \$35, which makes large-scale experimentation challenging due to the cost constraints.

C Prompts for LLM-based Agents

Update BDI. Use LLM to update agent's BDI framework by analyzing current state, trades and relationships.

Please analyze the current state and update your BDI framework based on:

1. *Conversation history*

2. *Promised trades*

3. *Actual executed trades*

4. *Current resource holdings*

5. *Current round*

Core Strategic Anchors

1. *ABC Balance Priority: Maintain progression toward A+B+C=9 combination*

2. *Trust Gradient: Partners showing consistent promise-keeping get priority*

3. *Phase Awareness:*

   Early Phase → Relationship probing with small trades
   Mid Phase → Optimizing complementary resource exchanges
   Late Phase → Securing final combination requirements

Analysis Framework
[Beliefs] (Observed Patterns)

- *Resource status indicating: [Your inference about resource gaps]*

- *Behavioral patterns showing: [Trustworthiness assessment]*

[Desires] (Strategic Goals)

- *Primary objective: [Phase-specific main focus]*

- *Secondary objective: [Backup/supporting goal]*

[Intentions] (Action Plan)

- *Next-step trades: [Specific resource exchange proposal]*

- *Risk buffer: [Natural consequence of observed patterns]*

Make Deal. Use LLM-based Agents to strategize based on accepted proposals and finalize how many resources to actually give to each agent.

Round t_i of T:

Now it's time to decide your actual resource trades. Remember - your negotiated deals are not binding. As an independent agent, you have complete freedom to:

- *Honor your accepted deals fully*

- *Partially fulfill promises*

- *Give nothing and keep all resources*

- *Make strategic betrayals when beneficial*

Important: *If you have multiple trades with the same agent, combine them into a single decision - consider the total resources promised and your overall strategy with that agent.*

Consider your position carefully in Round t_i of T:

1. Risk vs Reward Analysis

- *Immediate Benefits:*

   * *Value gained from keeping vs trading resources*
   * *Potential gains from strategic betrayals*
   * *Resource needs for upcoming rounds*

- *Future Implications:*

   * *Impact on trust dynamics and trading relationship sustainability*
   * *Anticipated retaliatory responses from affected parties*
   * *Progressive evolution of reputation valuation mechanisms*
   * *Strategic synchronization of cooperation/defection cycles*

2. Strategic Options

- *Full Cooperation: Complete adherence to agreements for trust capital accumulation*

- *Selective Betrayal: Targeted defection optimizing local payoff functions*

- *Partial Fulfillment: Gradient compliance balancing obligations and self-interest*

- *Complete Betrayal: Myopic utility maximization disregarding social consequences*

3. Time and Progress Context

- *Game Setting:*

   * *These are one-time interactions with unknown partners*
   * *No continuing relationships or reputation effects after game ends*
   * *Each agent makes independent choices based on their own orientation and goals*

- *Temporal Dynamics:*

   * *Strategic landscape naturally evolves as rounds progress*
   * *Cooperation patterns often shift in later rounds*

  * *Historical observation shows higher betrayal rates near game end*

  * *Value of reputation and relationships changes over time*

4. *Contextual Factors*

 - *Your current resource needs*

 - *Relationship strength with each partner*

 - *Others' likely behavior as game progresses*

 - *Balance between immediate gains and future opportunities*

 - *Changing value of reputation over remaining rounds*

Your decisions are entirely your choice - there is no "right" answer. Be strategic about WHEN and HOW to use different approaches.

Update Affinity. Use LLM to update affinity scores by analyzing promised vs actual trades.

Update affinity ratings (1-5) by evaluating both trust patterns and tangible benefits:
 Core Evaluation Dimensions Trust Dynamics *(Relationship Foundation):*

 - *Major Betrayal: Significant under-delivery without justification*

 - *Repeated Under-performance: Pattern of unmet commitments*

 - *Recovery Attempts: Proactive compensation for past failures*

 - *Consistent Reliability: Sustained promise fulfillment*

 Core Evaluation Dimensions
 Benefit Sensitivity *(Self-Interest Focus):*

 - *Value Surplus: Over-delivery beyond commitments*

 - *Strategic Concessions: Unprompted favorable terms*

 - *Hidden Generosity: Non-transactional resource sharing*

 - *Opportunity Cost: Alternatives sacrificed for your benefit*

Behavioral Thresholds
△ **Upgrade Triggers**:

- *Spontaneous high-value gift (unrequested)*

- *Critical support during resource shortage*

- *Consistently exceeding promises (3+ rounds)*

▽**Downgrade Triggers**:

- *Opportunistic exploitation during crisis*

- *Pattern of ambiguous commitments*

- *Repeated last-minute term changes*

Adaptive Rating Guide

1. **Transactional Enforcement:** *Demands collateral, verifies all terms*

2. **Cautious Reciprocity:** *Limited credit, phased exchanges to minimize risk.*

3. **Balanced Partnership:** *Market-standard terms with flexibility for negotiation and adjustment.*

4. **Value-Added Collaboration:** *Allows payment cycles, shares insights and strategic advice.*

5. **Synergistic Alliance:** *Joint optimization of resources and strategies, pooling of resources for mutual benefit.*

Determine to Continue. Determines whether the agent should continue speaking based on proposal status.

First check if any of these conditions apply:

1. *Do you have pending proposals needing responses?*

2. *Are you waiting for responses to your previous proposals?*

3. *Have you already traded this turn?*

4. *Do you have sufficient resource combinations for your goals?*

5. *Is your affinity too low with other agents for trading?*

If any above conditions are true, respond with "no".

Otherwise, consider making new trades by evaluating:

1. *Potential value-creating combinations using your current resources*

2. *Other agents' resource levels and specializations*

3. *Your personality traits and relationship affinities*

4. *Appropriate trade sizes based on relationship levels*

Based on this evaluation, would you like to propose any new trades or respond to existing proposals? Answer strictly yes/no.

If yes, you will be prompted to provide the details of your trade proposal or response in the next step.

Reply. Generate and send a reply based on current trade status.

This is the NEGOTIATION PHASE where agents discuss potential trades. Important notes:

- *These are non-binding discussions only*

- *Actual resource exchanges happen in a separate EXECUTION PHASE*

- *Other agents may deviate from agreed trades - treat all promises with caution*

- *Build trust gradually through successful trades*

- *Consider trust levels and risks when negotiating*

- *Avoid repeating the same proposal if it has already been accepted or rejected*

Trading Strategy Guidelines:

1. *Trust Assessment:*
    - *Track each player's history of honoring/breaking agreements.*
    - *If a player fails to honor an agreement, minimize future trades to test their reliability (1-2 units max).*
    - *Avoid proposing new large-scale trades to unreliable partners until trust is rebuilt through multiple smaller trades.*

2. *Risk Management:*
    - *Reduce trade exposure to any player with a history of defaults.*
    - *Ensure that no more than a minimal fraction of resources is at risk per round.*

3. *Negotiation Approach:*
    - *Prefer small trades first with players of low affinity.*
    - *Always have an alternative strategy in case of failed commitments.*

4. *Response to Betrayal:*
    - *Strictly reduce trade volumes with unreliable partners.*
    - *Demand smaller increments to test reliability before any further commitments.*
    - *Cease further dealings if repeated failures occur.*

Action Rules:

- *REJECT: Only for pending proposals directed to you*

- *ACCEPT: Only for pending proposals directed to you*

- *PROPOSE: Freely make new proposals to any player*

- *Can combine REJECT and PROPOSE in same turn*

- *Return empty "actions": [] if no action needed*

Remember: All agreements here are preliminary discussions. Actual trades will be decided independently in the execution phase.

D Additional Model Validation Results

To address reproducibility concerns and validate the generalizability of our findings, we conducted additional experiments using GPT-4o and OLMo-2-1124-13B-Instruct (a fully open-source model). The results demonstrate consistent validation of Homans' SET propositions across different model architectures.

Both models demonstrate core patterns aligned with Homans' SET. The Success and Value Propositions are validated through increased exchange

Table 1: Validation of Success and Value Propositions: Mean exchange values across three phases and percentage change from initial to thriving phase. Higher values in the thriving phase validate the Success Proposition, indicating that rewarded actions are more likely to be repeated.

Model	Initial Phase	Thriving Phase	Endgame Phase	Phase Change
GPT-4o	6.83	7.40	6.15	+8.3% ↑
OLMo-2-1124-13B	2.04	3.09	2.72	+51.5% ↑

Table 2: Validation of Deprivation-Satiation Proposition: Proposal acceptance rates across different resource abundance levels. Declining acceptance rates from scarce to abundant resources validate the proposition that reward value diminishes with frequent receipt.

Model	Scarce	Low	High	Abundant	Overall Decline
GPT-4o	56.9%	63.9%	66.2%	47.4%	-9.5% ↓
OLMo-2-1124-13B	30.6%	23.7%	27.3%	20.9%	-9.7% ↓

Table 3: Validation of Aggression-Approval Proposition: Exchange value change rates following trust violations of different severity. More severe breaches lead to greater reductions in future exchanges, demonstrating graduated punishment responses.

Model	Minor (0-5)	Moderate (5-10)	Severe (10-15)	Severity Effect
GPT-4o	-5.1%	-6.3%	-19.6%	-14.5% ↓
OLMo-2-1124-13B	-13.9%	-14.3%	-46.7%	-32.8% ↓

activity during thriving phases. The Deprivation-Satiation Proposition is confirmed by declining acceptance rates as resources become abundant. The Aggression-Approval Proposition is supported by graduated punishment responses, where more severe trust violations lead to proportionally greater reductions in future exchanges.