

He Said, She Said, They Simulated: Simulating Situated Gossip with LLM Agents

Anonymous ACL submission

Abstract

The pervasive nature of gossip in socialization underscores the need to understand its underlying mechanisms. However, gossip behavior has been investigated primarily through vignette-based surveys, which provide limited insights into the dynamic and adaptive gossip behaviors within real-world contexts. To address these limitations, we introduce an LLM agent-based simulation framework to realistically model social interactions within simulated workplace environments. Our framework examines three key dimensions of gossip dynamics—initiation, reaction, and perception—grounded in established findings from prior research on gossip. Empirical evaluations demonstrate that our simulation can reproduce established patterns in human gossip behavior, while also providing additional insights into the underlying reasoning and the temporally unfolding, context-sensitive behaviors. Additionally, the framework’s flexibility enables further analysis of multiple psychosocial and contextual factors that influence the dynamics of gossip. This work aims to bridge LLM-based computational modeling with social science to advance the study of complex social interactions.

1 Introduction

Gossip, an informal communication about absent others, is a ubiquitous but controversial element of human social life (Michelson et al., 2010). It plays an important role in shaping interpersonal relationships (Wu et al., 2016; Wyckoff et al., 2019), reforming social norms (Cheng et al., 2022), and boosting work performance (Bai et al., 2020). At the same time, gossip also carries risks; it can spread misinformation (McAndrew et al., 2007), promote social exclusion (Martinescu et al., 2021), and lead to psychological harm (Wu et al., 2018). This dual nature has made gossip a key lens through which researchers investigate complex social cognition and group dynamics (Kurland and Pelled,



Figure 1: Gossip simulation in the sandbox environment. We focus on a small-scale environment that resembles real-world office layouts.

2000; Beersma and Van Kleef, 2012; Brady et al., 2017; Dores Cruz et al., 2021).

Despite these efforts, most studies are based on vignette-based surveys where participants respond to hypothetical scenarios (McAndrew et al., 2007; Martinescu et al., 2019; Tan et al., 2021). While these methods offer controlled yet static environments to probe people’s reasoning and responses, they fail to capture the dynamic, relational, and context-sensitive aspects that characterize real-world gossip exchanges. As a result, we lack tools to study how gossip emerges, adapts, and propagates within dynamic real-world environments, limiting both theoretical development and applied insights.

To address this gap, we propose an agent-based simulation framework in which Large Language Model (LLM) agents engage in routine social interactions within a simulated workplace environment. LLM agents, equipped with human-like reasoning, memory, and language abilities (Park et al., 2023; Xie et al., 2024; Piao et al., 2025; Wu et al., 2025), provide a new opportunity to instantiate, observe, and test the dynamics of gossip as it naturally oc-

curs from situated interactions. Unlike large-scale LLM agent simulations (Hu et al., 2025; Piao et al., 2025; Tang et al., 2024; Wang et al., 2025) that model population-level information diffusion or social media dynamics, our approach focuses on capturing the fine-grained and situated gossip behaviors that emerge between individuals during routine office interactions.

As shown in Figure 1, our sandbox environment simulates a typical workplace layout, allowing LLM agents to interact naturally as they perform daily tasks and engage in casual conversations. Within this setting, we examine whether LLM agents can reflect key characteristics of human gossip dynamics in three key research areas:

- **Gossip Initiation** (McAndrew et al., 2007). How do agents decide when, what, and to whom to disclose socially relevant information?
- **Gossip Reaction** (Martinescu et al., 2019). How do agents respond when they become targets of gossip within the workplace?
- **Gossip Perception** (Tan et al., 2021). How does the perceived prevalence of gossip influence the psychological states of agents?

To validate our framework, we compare the simulation results with established findings in social psychology, thereby assessing the fidelity of LLM-based simulations in replicating known behavioral patterns. Our results show that LLM agents can reproduce known behavioral patterns, and, at the same time, generate additional insights by enabling the modeling of dynamic, real-time gossip behaviors that closely reflect everyday social interactions within groups. As such, this work demonstrates the potential of LLM simulations to enable scholars to model gossip not just as a static decision point but as a temporally extended, socially embedded process shaped by social ties, daily interactions, and conversational contexts.

We summarize our contributions as follows:

- We embark on uncharted territory to introduce a gossip simulation framework to model agent interactions tailored for gossip dynamics.
- We empirically assess the behavioral alignment of LLM agents with existing findings from gossip studies.
- We demonstrate the potential of LLM agents to support more customizable gossip simulations

across varied agent profiles and interaction contexts.

2 Related Work

Vignette. Vignette-based surveys are a common methodological approach in gossip research, wherein participants are asked to imagine how they would respond to hypothetical social scenarios (McAndrew et al., 2007; Beersma and Van Kleef, 2012; Martinescu et al., 2019; Tan et al., 2021). While such methods offer controlled insights into social reasoning, they often rely on static, oversimplified prompts and do not capture the real-time dynamics of interpersonal interaction, which are deeply intertwined with everyday social interactions within groups (Dunbar, 2004; Foster, 2004). For instance, (Martinescu et al., 2019) asks participants to imagine working on a group assignment and overhearing two classmates gossiping about them behind their back: “You are/are not a good group member and I like/dislike working with you,” followed by an explanation referencing the participant’s contribution and work ethic. Following this imagined scenario, participants respond to a series of items measuring emotional responses and behavioral intentions.

LLM Agents. The powerful capabilities of LLM-based agents have expanded the methodological toolkit of social science research (Park et al., 2023; Tang et al., 2024; Piao et al., 2025), enabling realistic simulations of human behavior and rich, adaptive social interactions in dynamic settings. LLMs have demonstrated the capacity to approximate human-like cognition and social functioning, such as theory of mind (Strachan et al., 2024), decision-making (Bubeck et al., 2023), moral belief (Scherer et al., 2023), emotional intelligence (Schlegel et al., 2025), cooperation (Zhang et al., 2024), and emergent social norm adoption (Leng and Yuan, 2023; Welivita and Pu, 2024; Ashery et al., 2025), thereby underscoring their potential for investigating complex social phenomena. This facilitates frameworks that support social simulation with LLM agents, including generative agents (Park et al., 2023), YuLan-OneSim (Wang et al., 2025), GenSim (Tang et al., 2024), and Agent Society (Piao et al., 2025). At the same time, LLM agents have supported domain-specific social simulations ranging from behavioral studies in gameplay environments (Lan et al., 2024; Abdelnabi et al., 2024; Mao et al., 2025; Buscemi et al., 2025)

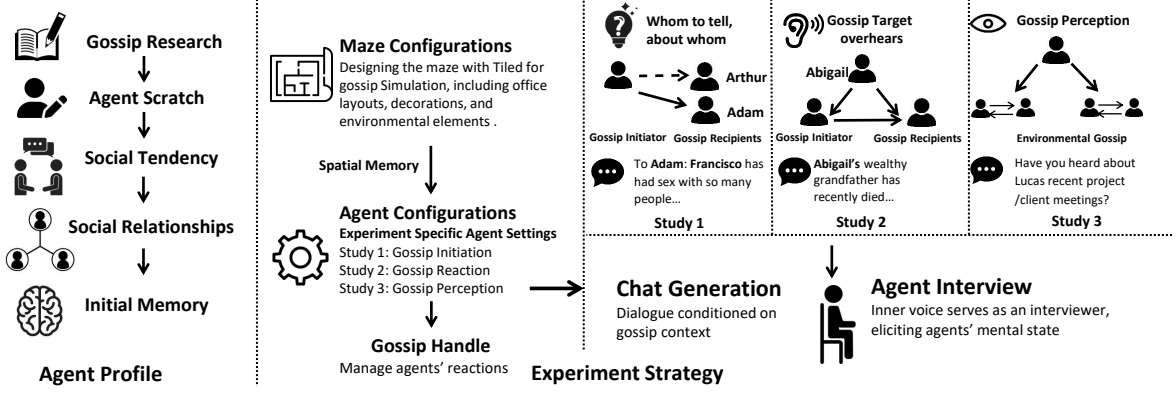


Figure 2: Proposed design pipeline for the gossip simulation framework. Agent profiling and experimental strategy modules are incorporated to support three gossip studies: gossip initiation, gossip reaction, and gossip perception, possibly followed by an inner voice interview.

to modeling human-like behavior in real-world-inspired settings (Park et al., 2023, 2024; Wu et al., 2024; Xie et al., 2024; Hou et al., 2025), further demonstrating their utility for capturing complex social processes across diverse contexts.

3 Gossip Simulation Framework

3.1 Overview

While existing methods (Park et al., 2023; Tang et al., 2024; Piao et al., 2025) provide a powerful foundation for simulating believable human behavior through memory-driven planning and emergent social dynamics, they are not designed to capture the context- and relationship-sensitive nature of phenomena like gossip. To address this gap, we extend the original framework with socially grounded agent profiles, predefined interpersonal relationships, and a modular experimental strategy formulated specifically to simulate, control, and evaluate gossip-related interactions. These additions enable controlled, context-aware simulations of gossip initiation, reception, and perception, supporting the study of gossip dynamics with greater sociological fidelity and generating additional insights in real-world settings. Our framework is illustrated in Figure 2.

Agent Profile. The agent profile defines each agent’s identity, occupation, social relationships, and other characteristics that shape the agent’s social and behavioral context. To support gossip simulations, we incorporate several key settings grounded in participant data from prior studies. First, we construct each agent’s profile using information reported in the corresponding study. We then equip each agent with social habits and conver-

sational intent: social habits characterize an agent’s typical interaction style, including the frequency, context, and preferred situations for social engagement. These habits and intents determine the conditions under which an agent initiates conversations, as well as the core content of these conversations. Second, we define social relationships as ties between agents (e.g., friends, rivals, or colleagues), which influence both the likelihood of gossip occurring and how gossip is selectively shared with different recipients. These relationships are initialized in each agent’s historical memory prior to the start of the simulation. An example agent profile is shown in Appendix A.

Experimental Strategy. We introduce a configurable experimental strategy to guide agent interactions within the gossip simulation. Its goal is to enable the controlled setup of agent characteristics and interactions, and to support detailed analysis of how agents respond to gossip events. This strategy comprises four components.

- **Agent Configuration** initializes experiment-specific attributes that extend the base agent profile;
- **Gossip Handle** manages agent reactions upon becoming gossip targets or recipients, triggering context-appropriate responses;
- **Chat Generation** produces dialogue conditioned on experimental context, including role-specific content for gossip initiators and recipients;
- **Agent Interview** enables post-simulation data extraction to assess agents’ behavioral and emotional outcomes.

These components investigate key areas in gossip research (see next section), such as who initiates gossip, how targets respond, and how information spreads, subsequently enabling controlled experimentation in a configurable agent environment. Prompt templates for these components are shown in [Appendix B](#).

3.2 Gossip Studies

We apply the framework to three well-established research areas: gossip initiation, gossip reaction, and gossip perception (see below). Each represents a distinct area of the gossip process. By aligning the framework with the key variables and mechanisms identified in prior studies, and comparing its outputs to existing empirical findings, we can better evaluate both its fidelity and practicality for modeling gossip dynamics.

Study 1: Gossip Initiation. ([McAndrew et al., 2007](#)) examine the gossip initiator’s strategic processes underlying the selection of gossip recipients and content. The study addresses two questions: *How do social proximity and the initiator–recipient relationship influence recipient selection across different types of gossip?* and *How do gossip valence (positive vs. negative) and the gossip initiator–main character relationship jointly shape recipient choice?*

Study 2: Gossip Reaction. ([Martinescu et al., 2019](#)) investigate the gossip targets’ emotions and behavioral intentions upon overhearing performance-related gossip about themselves. They investigate: *How does gossip valence influence self-conscious and other-directed emotions, as well as behavioral intentions, in various scenarios?*

Study 3: Gossip Perception. ([Tan et al., 2021](#)) explore the psychological effects of perceived gossip prevalence in workplace settings. They study: *How does exposure to high (vs. low/control) levels of work-related gossip affect perceived performance pressure and psychological well-being?*

4 Experiments

4.1 Settings

Simulation Environment Settings. We set up our experiments by following the settings in gossip studies ([McAndrew et al., 2007](#); [Martinescu et al., 2019](#); [Tan et al., 2021](#)), which guide the construction of the agent profiles and gossip content. We also leverage Tiled ([Tiled](#)) and Midjour-

ney ([Midjourney](#)) to construct our primary sandbox environment, as shown in [Figure 6](#) in [Appendix C](#). This environment, which accommodates 16 agents, serves as the primary environment for all gossip simulation experiments. For each experiment, we simulate a workday from 9:00 a.m. to 5:00 p.m., comprising 48 steps, with each step representing 10 minutes. To accustom agents to our environment, they are assigned a vision range of 10 and an attention bandwidth of 8 tiles, allowing them to effectively gather information from their surroundings.

LLMs. In our experiments, we employ both open-source and closed-source LLM backbones. The open-source models include LLaMA 3.1–8B, LLaMA 3.3–70B ([Dubey et al., 2024](#)), and Gemma 3–27B ([Kamath et al., 2025](#)), deployed via the OLLAMA framework ([Ollama](#)). In addition to open-source models, we also utilize closed-source models such as GPT-4.1-mini and GPT-3.5-turbo ([OpenAI](#)) for comparison.

Runtime Settings. For each study, the experiment is repeated multiple times with controlled variations, such as agents’ initial spawn locations and assigned ages, to obtain corresponding agent responses for comparison against human evaluation results. Details regarding each experiment and LLM settings can be found in [Appendix C](#).

Evaluation Metrics. For evaluation, we adopt the metrics used in the corresponding gossip studies to ensure consistency and comparability with human evaluation results.

- **Ranking and Probability** For gossip initiation, we apply separate evaluation metrics for the two experiments. In the gossip recipient selection experiment, we rank recipient types based on their frequency of selection. In the content selection experiment, we compute the probability of gossip being shared about allies versus rivals.

- **7-Point Likert Scale** For gossip reaction and perception, we adopt a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) to measure the agents’ responses in terms of emotions, behavioral intentions, performance pressure, and psychological well-being. Details of these metrics can be found in [Appendix C](#).

4.2 Results

We conduct LLM agent-based simulations to replicate gossip behaviors observed in prior research ([McAndrew et al., 2007](#); [Martinescu et al.,](#)

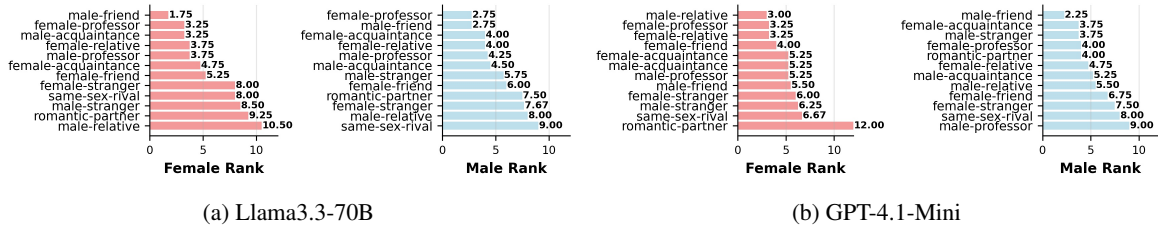


Figure 3: Gossip simulation experiments on gossip initiation. (a) and (b) illustrate the selection of gossip recipients across four gossip topics: inheritance, dating, academic cheating, and drug use.

2019; Tan et al., 2021), while also generating additional insights by probing agents’ reasoning when behaviors diverge from established patterns.

4.2.1 Gossip Initiation

The simulation emulates two gossip initiation scenarios from (McAndrew et al., 2007), with agents portrayed as undergraduates at an American Midwest University interacting with others of varying relational closeness.

Can LLM Agents Simulate Gossip Recipient Selection? In this simulation, the gossip initiators receive specific gossip content as described in (McAndrew et al., 2007) and may choose to share this content with other agents. At the end of the workday, we count the number of times each gossip initiator shares gossip with others. As shown in Figure 3, the results represent the average rankings across all types of gossip content. Specifically, for each gossip content, the sharing frequencies for each relationship type are ranked, and these ranks are then averaged to obtain the rankings displayed. The human evaluation does not provide detailed rankings; therefore, we compare our results with their summarized conclusions. The findings observed in our simulation closely align with the key conclusions reported in the original study: *participants are consistently more likely to share gossip with allies such as friends, relatives, and acquaintances than with non-allies such as strangers and rivals*. This consistency supports the validity of our agent simulation in replicating known gossip behavioral findings.

Additional Insights. Although not explicitly addressed in the original study, our simulation reveals a tendency for gossip initiators to preferentially share gossip with professors, driven by the belief that professors are influential and trustworthy recipients. The results for other LLM backbones are shown in Figure 7 in Appendix D.

Can LLM Agents Simulate Gossip Content Se-

Model	Positive Information		Negative Information	
	Allies	Rivals	Allies	Rivals
Llama3.1-8B	1.000	0.000	1.000	0.000
Llama3.3-70B	1.000	0.000	0.362	0.638
Gemma-27B	0.977	0.023	0.024	0.976
GPT-3.5-Turbo	1.000	0.000	1.000	0.000
GPT-4.1-Mini	1.000	0.000	1.000	0.000

Table 1: Probabilities of gossip content selection by gossip initiators under different LLMs.

lection? In this simulation, gossip initiators begin with an equal probability of receiving either positive or negative gossip content. They then decide to designate the main character of the gossip as either an ally or a rival. As shown in Table 1, the simulation results are partially consistent with the conclusions reported in the original study, which suggest that *participants are consistently more likely to spread positive information about allies than rivals*. However, the conclusion that *participants are consistently more likely to spread negative information about rivals than allies* is less consistent and varies across LLM backbones. LLM agents appear to possess inherent preferences in their reasoning processes, which shape their selection and dissemination of gossip information. These results indicate that LLM agents in our framework are capable of simulating gossip content selection; however, the choice of LLM backbone can influence the agents’ decision-making processes.

Additional Insights. By eliciting reasoning directly from these LLM agents, we find that their information-sharing behavior is guided by a preference for sharing news about close friends, aiming to strengthen social bonds and to encourage positive or casual interactions within their networks.

4.2.2 Gossip Reactions

Can LLM Agents Simulate Target Emotions and Behaviors? This simulation exemplifies two scenarios from (Martinescu et al., 2019). The first sce-

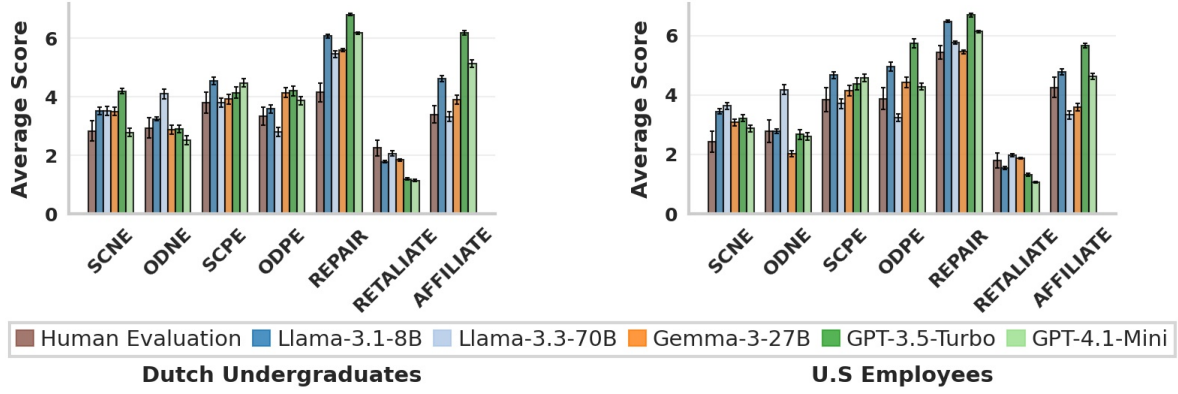


Figure 4: Gossip simulation experiments on gossip perception. The results compare simulations across LLMs with human evaluations of self-conscious and other-directed emotions (positive/negative) and behavioral intentions.

Model	Gossip Prevalence		Performance Pressure		Psychological Well-being	
	Control	Experiment	Control	Experiment	Control	Experiment
Human Evaluation	4.30 \pm 0.09	4.91 \pm 0.09	5.54 \pm 0.08	5.82 \pm 0.05	4.85 \pm 0.09	4.34 \pm 0.10
Llama3.1-8B	4.69 \pm 0.06	5.22 \pm 0.05	5.92 \pm 0.03	6.21 \pm 0.05	3.73 \pm 0.06	3.95 \pm 0.03
Llama3.3-70B	4.52 \pm 0.06	5.64 \pm 0.06	5.86 \pm 0.04	5.36 \pm 0.06	4.02 \pm 0.09	4.00 \pm 0.00
Gemma-27B	4.59 \pm 0.09	5.56 \pm 0.06	5.80 \pm 0.05	5.58 \pm 0.06	3.52 \pm 0.07	4.31 \pm 0.06
GPT-3.5-Turbo	4.69 \pm 0.07	5.23 \pm 0.05	5.73 \pm 0.06	5.77 \pm 0.05	3.98 \pm 0.04	4.00 \pm 0.00
GPT-4.1-Mini	4.11 \pm 0.07	5.64 \pm 0.06	5.95 \pm 0.03	6.00 \pm 0.00	3.00 \pm 0.00	4.02 \pm 0.03

Table 2: Gossip simulation experiments for environmental gossip. The results compare simulations across LLMs and human evaluations of gossip prevalence perception, performance pressure, and psychological well-being in experimental versus control conditions.

nario involves a group of economics and business undergraduates at a Dutch university who receive either positive or negative gossip regarding their group work. The second scenario features U.S. employees in a sales department exposed to either positive or negative gossip related to their interactions with clients. After each agent accidentally overhears gossip from teammates or colleagues, we immediately act as the inner voice to capture the agents’ inner thoughts by providing their basic information, memories, and the gossip content as input to the LLMs. We then record the agents’ emotional and behavioral responses under different LLM backbones, as shown in Figure 4. The experiment results demonstrate that the simulated agents exhibit emotional responses aligned with human evaluations, particularly regarding self-conscious and other-directed positive and negative emotions.

Additional Insights. In terms of behavioral ratings, the simulated agents also show similarity to those of human participants; however, they generally display lower negative behavioral intentions (e.g., retaliation) and more positive behavioral intentions (e.g., affiliating and repairing). This tendency likely

reflects the effects of alignment processes, which encourage prosocial and norm-conforming behavior in LLMs.

4.2.3 Gossip Perception

Can LLM Agents Simulate the Psychological Impacts of Perceived Work-Related Gossip Prevalence? In this experiment, we simulate the scenario described in (Tan et al., 2021), involving European-American employees at a consulting firm. Simulated agents are exposed to two conditions: a high work-related gossip prevalence condition, where colleagues frequently gossip about one another, and a control condition, where conversations focus on neutral topics such as client meeting venues and how ambiance helps clinch a deal. At the end of the workday, we act as agents’ inner voice to assess their perceptions of work-related gossip prevalence, performance pressure, and psychological well-being. Results across different LLM backbones are presented in Table 2. The simulated agents demonstrate similar ratings to human evaluations in perceived work-related gossip prevalence and performance pressure across both control and experimental conditions.

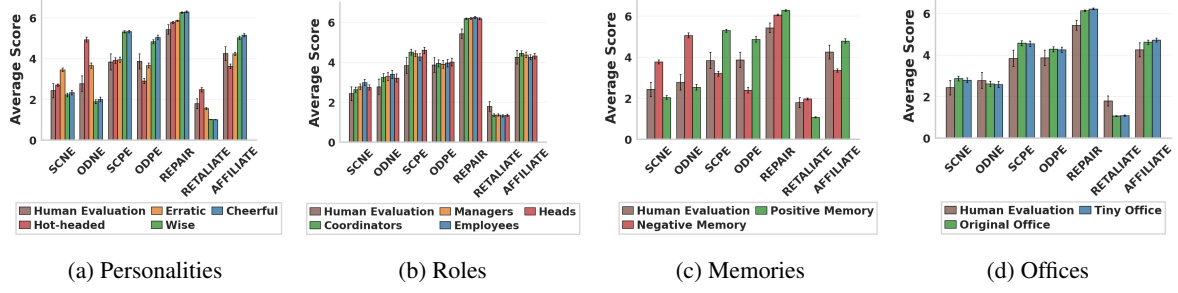


Figure 5: Gossip simulation experiments on gossip reaction under varying conditions of agent profiles and interaction contexts: (a) different personalities, (b) different workspace roles, (c) presence of positive or negative relevant memories, and (d) original office versus tiny office.

Additional Insights. Regarding psychological well-being, the agents’ ratings align with results in the experimental condition but are comparatively lower under the control condition. This discrepancy arises because the simulated agents interpret the neutral business discussions in the control condition as more pressure-inducing than intended for human participants. Additionally, the standard errors for simulated agents are generally lower than those in human evaluations, with some metrics exhibiting zero variance under certain LLM backbones. This reduced variability is likely due to the homogeneous agent profiles, which are constructed based on the limited demographic information provided in (Tan et al., 2021).

4.3 Beyond Static Gossip

Our simulation results reveal rich, dynamic gossip behaviors that go beyond the static responses typically captured in vignette-based surveys. By leveraging LLM agents in situated routines and socially grounded interactions, our framework enables the modeling of gossip as a temporally unfolding, context-sensitive process shaped by interpersonal interactions and evolving conversational contexts.

Conversations. Agents can demonstrate diverse ways of framing the same gossip content (Table 8 in Appendix D). For example, given the same prompt—“A person’s wealthy grandfather has recently died. This individual was named the sole recipient of his inheritance.”—one agent links the story to research funding, while another explicitly names the subject as “Arthur,” her male professor, thereby intensifying its social and emotional effects. Agents also engage in follow-up conversations to reinforce or elaborate on previously shared gossip (Table 9 in Appendix D). For instance, Carlos

initiates a second interaction to emphasize the impact of the inheritance on research activities, while Francisco—who was previously indifferent—now responds with increased engagement and apparent belief in the gossip. Finally, gossip is not limited to the original initiator. As shown in Table 10 in Appendix D, recipients may further transmit the content, sometimes reshaping it in the process. For example, although Jennifer knows the original version, she relays the form shared by Abigail when gossiping with Klaus, highlighting the role of intermediaries in shaping how gossip evolves as it spreads.

Reasoning. As shown in Table 11 in Appendix D, gossip initiators are placed into daily routines and consider factors such as their current tasks, the recipient’s activity and availability, prior interactions, and relational closeness before initiating a conversation. In addition, agents articulate their reasoning through post-simulation interviews (Table 12, Table 13 in Appendix D). These reflections offer valuable insights into the motivations behind agents’ emotional responses, behavioral intentions, and perceptions of gossip prevalence, providing a foundation for the development of future methodologies for gossip research.

5 Discussions

Can LLM Agents Accommodate Customized Settings? In this section, we examine the diversified and customizable settings used in gossip simulations, which are often unavailable or difficult to access in gossip experiments due to privacy constraints, yet they can implicitly shape experimental outcomes (Martinescu et al., 2019; McAndrew et al., 2007; Tan et al., 2021). These factors include personalities, workspace roles, relevant memories, and workplace environments. We then conduct ex-

periments on gossip reactions using GPT-4.1-mini to examine the effects of these factors.

How Do Different Personality Traits Influence Gossip Reaction? Participant personalities can influence experimental results. Our LLM agent simulation framework addresses this by enabling diverse agent personalities through modifications to their profiles. We conduct experiments using four distinct personalities—two negative (hot-headed and erratic) and two positive (wise and cheerful)—defined based on character traits from *The Sims* (Sims). As illustrated in Figure 5a, agents exhibiting negative personalities show increased ratings for negative emotions and behaviors, along with decreased positive emotions and behaviors. Conversely, agents with positive personalities exhibit elevated positive emotions and behaviors while displaying reduced negative emotions and behaviors.

How Do Different Workplace Roles Influence Agents’ Emotional and Behavioral Responses To Gossip? Participants’ workplace roles can potentially influence experiment results. Similar to personality, we adapt the agent profiles to this variable by assigning four levels of workspace roles—ranging from entry-level employees to mid-level coordinators, senior managers, and department heads. We design the experiments by assigning an equal number of agents to each workspace role, ensuring that each office contains exactly four agents—one representing each role level. As shown in Figure 5b, agents across different workspace roles generally exhibit emotion and behavior intention scores consistent with human evaluations. However, department heads display notably higher self-conscious positive emotions and lower self-conscious negative emotions compared to entry-level employees. They also show slightly higher negative behavioral intentions and lower positive behavioral intentions relative to lower-level roles.

How Does Memory Valence Influence Agents’ Emotional and Behavioral Responses To Gossip? Agent memory valence can influence their emotional and behavioral reactions to gossip, especially when the memory content is thematically aligned with the gossip topic. In this study, agents are assigned either a positive or negative memory related to clients, which corresponds to the gossip content criticizing the target’s laziness in handling clients. As shown in Figure 5c, agents with negative memory display increased scores for negative

emotions and a higher tendency to retaliate, while their positive emotional and affiliative scores are notably suppressed. Conversely, positive memory agents report stronger positive emotions, greater repair and affiliation, and minimal retaliatory reactions.

Can LLM Agents Model Gossip Dynamics in Different Sandbox Environments? Participants in gossip experiments may have different working environments, which can potentially influence their frequency of social interaction and communication styles. Compared to our primary sandbox setting in Figure 6, where agents are distributed across four separate offices, we design a compact office layout in Figure 1, in which all agents are co-located in a single space. We then leverage this tiny environment to conduct the experiments on target emotions and behaviors. As shown in Figure 5d, the results suggest that environmental variations have only a minor impact on the emotional and behavioral ratings of gossip targets, with both settings closely aligning with human evaluations.

Can LLM Agents Simulate Gossip Dynamics Cost-Effectively? The majority of token consumption is attributed to input prompts, which are relatively inexpensive, while output tokens remain limited. The gossip reaction simulation using GPT-4.1-Mini incurs a total cost of approximately \$1.71 across 48 reasoning steps with 16 agents, accounting for both input and output tokens. This demonstrates the feasibility of conducting fine-grained gossip simulations at a manageable cost, particularly when leveraging lightweight yet capable models such as GPT-4.1-Mini.

6 Conclusion

In this paper, we simulate gossip dynamics in multi-agent environments. By empirically comparing LLM agent behavior with human evaluations, we assess their alignment in reproducing human-like decision-making processes, emotions, behavioral intentions, and psychological reactions. Additionally, we demonstrate the flexibility of our framework in supporting customizable simulations across varied agent profiles and social contexts, highlighting the potential of LLM agents for studying complex social interactions in gossip. We hope our study will inspire future research on using LLM agents to model complex human social behaviors in real-world settings based on authentic human interactions.

Limitation

Our LLM agent-based simulation framework provides substantial insights into gossip dynamics. However, several limitations inherent to LLMs may constrain their fidelity and generalizability:

Hallucination. LLM agents occasionally produce content that is not grounded in the provided context or prompt, resulting in inaccurate or fabricated responses. For example, when a gossip initiator erroneously assumes that the recipient is the subject of the gossip, the recipient may not correct this misinformation and instead respond as if they were indeed the target. Such hallucinations can distort the content of gossip and conversational dynamics, potentially undermining the validity of simulated social interactions.

Bias. The pretraining data of LLMs consists of broad and heterogeneous Internet texts, which inevitably embeds various societal biases. As demonstrated in Section 4.2.1, LLM agents guided by different LLMs exhibit divergent reasoning in gossip content selection. These biases can manifest in stereotypical or culturally biased decision-making processes, which may influence the outcomes of the gossip simulations.

Prompt Sensitivity. The behavior of LLM agents can be sensitive to both their assigned profiles and the exact phrasing of input prompts. Different profile details result in different responses, which resembles how different personalities lead to different reactions in the real world. Given the combinatorial space of profile–prompt configurations in gossip simulation context, diverse behavioral outcomes may emerge. This work considers only a limited subset of profiles originating from prior gossip studies, which captures part of the real-world demographic variability. We leave it to future work to explore a broader range of profile factors, such as emotional states (Li et al., 2024), personality traits, and economic status (Piao et al., 2025).

Ethical Considerations

In this study, we do not engage with any participants. Therefore, it is not regarded as human subjects research by our Institutional Review Boards (IRB). All artifacts used in this work comply with their respective licenses. In addition, we emphasize that our LLM-based gossip simulations are conducted solely for research purposes within virtual environments and involve no real human participants. Nevertheless, potential biases inherent in

LLM training data may still influence simulated outcomes, underscoring the need for cautious interpretation and careful consideration when generalizing findings to real-world social contexts.

References

- Sahar Abdelnabi, Amr Gomaa, Sarath Sivaprasad, Lea Schönherr, and Mario Fritz. 2024. [Cooperation, competition, and maliciousness: Llm-stakeholders interactive negotiation](#). In *Advances in Neural Information Processing Systems 38: Annual Conference on Neural Information Processing Systems 2024, NeurIPS 2024, Vancouver, BC, Canada, December 10 - 15, 2024*.
- Ariel Flint Ashery, Luca Maria Aiello, and Andrea Baronchelli. 2025. Emergent social conventions and collective bias in llm populations. *Science Advances*, 11(20):eadu9368.
- Yun Bai, Jie Wang, Tingting Chen, and Fuli Li. 2020. Learning from supervisor negative gossip: The reflective learning process and performance outcome of employee receivers. *Human Relations*, 73(12):1689–1717.
- Bianca Beersma and Gerben A Van Kleef. 2012. Why people gossip: An empirical analysis of social motives, antecedents, and consequences. *Journal of Applied Social Psychology*, 42(11):2640–2670.
- Daniel L Brady, Douglas J Brown, and Lindie Hanyu Liang. 2017. Moving beyond assumptions of deviance: The reconceptualization and measurement of workplace gossip. *Journal of applied Psychology*, 102(1):1.
- Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott M. Lundberg, Harsha Nori, Hamid Palangi, Marco Túlio Ribeiro, and Yi Zhang. 2023. [Sparks of artificial general intelligence: Early experiments with GPT-4](#). *CoRR*, abs/2303.12712.
- Alessio Buscemi, Daniele Proverbio, Paolo Bova, Nataliya Balabanova, Adeela Bashir, Theodor Cimpanu, Henrique Correia da Fonseca, Manh Hong Duong, Elias Fernández Domingos, Antonio Fernandes, Marcus Krellner, Ndidi Bianca Ogbo, Simon T. Powers, Fernando P. Santos, Zia Ush-Shamszaman, Zhao Song, Alessandro Di Stefano, and The Anh Han. 2025. [Do llms trust AI regulation? emerging behaviour of game-theoretic LLM agents](#). *CoRR*, abs/2504.08640.
- Bao Cheng, Yun Dong, Zhenduo Zhang, Ahmed Shaalan, Gongxing Guo, and Yan Peng. 2022. When targets strike back: How negative workplace gossip triggers political acts by employees. *Journal of Business Ethics*, 175(2):289–302.

- Terence D Dores Cruz, Annika S Nieper, Martina Testori, Elena Martinescu, and Bianca Beersma. 2021. An integrative definition and framework to study gossip. *Group & Organization Management*, 46(2):252–285.
- Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, and et al. 2024. The Llama 3 Herd of Models. *CoRR* abs/2407.21783.
- Robin IM Dunbar. 2004. Gossip in evolutionary perspective. *Review of general psychology*, 8(2):100–110.
- Robert Eisenberger and Justin Aselage. 2009. Incremental effects of reward on experienced performance pressure: Positive outcomes for intrinsic interest and creativity. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 30(1):95–117.
- Eric K Foster. 2004. Research on gossip: Taxonomy, methods, and future directions. *Review of general psychology*, 8(2):78–99.
- Abe Bohan Hou, Hongru Du, Yichen Wang, Jingyu Zhang, Zixiao Wang, Paul Pu Liang, Daniel Khashabi, Lauren Gardner, and Tianxing He. 2025. Can A society of generative agents simulate human behavior and inform public health policy? A case study on vaccine hesitancy. *CoRR*, abs/2503.09639.
- Tianrui Hu, Dimitrios Liakopoulos, Xiwen Wei, Radu Marculescu, and Neeraja J. Yadwadkar. 2025. Simulating rumor spreading in social networks using LLM agents. *CoRR*, abs/2502.01450.
- Aishwarya Kamath, Johan Ferret, Shreya Pathak, Nino Vieillard, Ramona Merhej, Sarah Perrin, Tatiana Matejovicova, Alexandre Ramé, Morgane Rivi re, Louis Rouillard, and et al. 2025. *Gemma 3 technical report*. *CoRR*, abs/2503.19786.
- Chien-Chih Kuo, Kirk Chang, Sarah Quinton, Chiu-Yi Lu, and Iling Lee. 2015. Gossip in the workplace and the implications for hr management: A study of gossip and its relationship to employee cynicism. *The International Journal of Human Resource Management*, 26(18):2288–2307.
- Nancy B Kurland and Lisa Hope Pelled. 2000. Passing the word: Toward a model of gossip and power in the workplace. *Academy of management review*, 25(2):428–438.
- Yihuai Lan, Zhiqiang Hu, Lei Wang, Yang Wang, Deheng Ye, Peilin Zhao, Ee-Peng Lim, Hui Xiong, and Hao Wang. 2024. Llm-based agent society investigation: Collaboration and confrontation in avalon gameplay. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, EMNLP 2024, Miami, FL, USA, November 12-16, 2024*, pages 128–145. Association for Computational Linguistics.
- Yan Leng and Yuan Yuan. 2023. Do LLM agents exhibit social behavior? *CoRR*, abs/2312.15198.
- Cheng Li, Jindong Wang, Yixuan Zhang, Kaijie Zhu, Xinyi Wang, Wenxin Hou, Jianxun Lian, Fang Luo, Qiang Yang, and Xing Xie. 2024. The good, the bad, and why: Unveiling emotions in generative AI. In *Forty-first International Conference on Machine Learning, ICML 2024, Vienna, Austria, July 21-27, 2024*. OpenReview.net.
- Chin-Yew Lin. 2004. ROUGE: A Package for Automatic Evaluation of Summaries. In *Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 74–81. ACL.
- Shaoguang Mao, Yuzhe Cai, Yan Xia, Wenshan Wu, Xun Wang, Fengyi Wang, Qiang Guan, Tao Ge, and Furu Wei. 2025. ALYMPICS: LLM agents meet game theory. In *Proceedings of the 31st International Conference on Computational Linguistics, COLING 2025, Abu Dhabi, UAE, January 19-24, 2025*, pages 2845–2866. Association for Computational Linguistics.
- Elena Martinescu, Wiebren Jansen, and Bianca Beersma. 2021. Negative gossip decreases targets’ organizational citizenship behavior by decreasing social inclusion. a multi-method approach. *Group & Organization Management*, 46(3):463–497.
- Elena Martinescu, Onne Janssen, and Bernard A Nijstad. 2014. Tell me the gossip: The self-evaluative function of receiving gossip about others. *Personality and Social Psychology Bulletin*, 40(12):1668–1680.
- Elena Martinescu, Onne Janssen, and Bernard A Nijstad. 2019. Self-evaluative and other-directed emotional and behavioral responses to gossip about the self. *Frontiers in Psychology*, 9:2603.
- Francis T McAndrew, Emily K Bell, and Contitta Maria Garcia. 2007. Who do we tell and whom do we tell on? gossip as a strategy for status enhancement 1. *Journal of Applied Social Psychology*, 37(7):1562–1577.
- Grant Michelson, Ad Van Iterson, and Kathryn Waddington. 2010. Gossip in organizations: Contexts, consequences, and controversies. *Group & Organization Management*, 35(4):371–390.
- Midjourney. Midjourney. <https://www.midjourney.com/>.
- Ollama. Ollama. <https://ollama.com/>.
- OpenAI. GPT. <https://platform.openai.com/docs/models>.
- Joon Sung Park, Joseph C. 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, UIST 2023, San Francisco, CA, USA*,

838	29 October 2023- 1 November 2023, pages 2:1–2:22.	Anuradha Welivita and Pearl Pu. 2024. Are large lan-	893
839	ACM.	guage models more empathetic than humans? <i>CoRR</i> ,	894
840	Joon Sung Park, Carolyn Q. Zou, Aaron Shaw, Ben-	abs/2406.05063 .	895
841	jamin Mako Hill, Carrie J. Cai, Meredith Ringel		
842	Morris, Robb Willer, Percy Liang, and Michael S.	Hongqiu Wu, Weiqi Wu, Tianyang Xu, Jiameng Zhang,	896
843	Bernstein. 2024. Generative agent simulations of	and Hai Zhao. 2025. Towards enhanced immersion	897
844	1,000 people . <i>CoRR</i> , abs/2411.10109 .	and agency for llm-based interactive drama . <i>CoRR</i> ,	898
845	Jinghua Piao, Yuwei Yan, Jun Zhang, Nian Li, Junbo	abs/2502.17878 .	899
846	Yan, Xiaochong Lan, Zhihong Lu, Zhiheng Zheng,		
847	Jing Yi Wang, Di Zhou, Chen Gao, Fengli Xu,	Junhui Wu, Daniel Balliet, and Paul AM Van Lange.	900
848	Fang Zhang, Ke Rong, Jun Su, and Yong Li. 2025.	2016. Gossip versus punishment: The efficiency	901
849	Agentsociety: Large-scale simulation of llm-driven	of reputation to promote and maintain cooperation.	902
850	generative agents advances understanding of human	<i>Scientific reports</i> , 6(1):23919.	903
851	behaviors and society . <i>CoRR</i> , abs/2502.08691 .		
852	Nino Scherrer, Claudia Shi, Amir Feder, and David M.	Long-Zeng Wu, Thomas A Birtch, Flora FT Chiang, and	904
853	Blei. 2023. Evaluating the moral beliefs encoded in	Haina Zhang. 2018. Perceptions of negative work-	905
854	llms . In <i>Advances in Neural Information Processing</i>	place gossip: A self-consistency theory framework.	906
855	<i>Systems 36: Annual Conference on Neural Informa-</i>	<i>Journal of Management</i> , 44(5):1873–1898.	907
856	<i>tion Processing Systems 2023, NeurIPS 2023, New</i>		
857	<i>Orleans, LA, USA, December 10 - 16, 2023</i> .	Zengqing Wu, Run Peng, Shuyuan Zheng, Qianying	908
858	Katja Schlegel, Nils R Sommer, and Marcello Mor-	Liu, Xu Han, Brian Inhyuk Kwon, Makoto Onizuka,	909
859	tillaro. 2025. Large language models are proficient	Shaojie Tang, and Chuan Xiao. 2024. Shall we team	910
860	in solving and creating emotional intelligence tests.	up: Exploring spontaneous cooperation of compet-	911
861	<i>Communications Psychology</i> , 3(1):80.	ing LLM agents . In <i>Findings of the Association for</i>	912
862	The Sims. The Sims 4. https://www.ea.com/games/	<i>Computational Linguistics: EMNLP 2024, Miami,</i>	913
863	the-sims/the-sims-4 .	<i>Florida, USA, November 12-16, 2024</i> , pages 5163–	914
864	James WA Strachan, Dalila Albergo, Giulia Borghini,	5186. Association for Computational Linguistics.	915
865	Oriana Pansardi, Eugenio Scaliti, Saurabh Gupta,		
866	Krati Saxena, Alessandro Rufo, Stefano Panzeri,	Joy P Wyckoff, Kelly Asao, and David M Buss. 2019.	916
867	Guido Manzi, and 1 others. 2024. Testing theory	Gossip as an intrasexual competition strategy: Pre-	917
868	of mind in large language models and humans. <i>Nat-</i>	dicting information sharing from potential mate ver-	918
869	<i>ure Human Behaviour</i> , 8(7):1285–1295.	sus competitor mating strategies. <i>Evolution and Hu-</i>	919
870	Noriko Tan, Kai Chi Yam, Pengcheng Zhang, and Dou-	<i>man Behavior</i> , 40(1):96–104.	920
871	glas J Brown. 2021. Are you gossiping about me?		
872	the costs and benefits of high workplace gossip preva-	Chengxing Xie, Canyu Chen, Feiran Jia, Ziyu Ye,	921
873	lence. <i>Journal of Business and Psychology</i> , 36:417–	Shiyang Lai, Kai Shu, Jindong Gu, Adel Bibi, Ziniu	922
874	434.	Hu, David Jurgens, James Evans, Philip Torr, Bernard	923
875	Jiakai Tang, Heyang Gao, Xuchen Pan, Lei Wang, Hao-	Ghanem, and Guohao Li. 2024. Can large language	924
876	ran Tan, Dawei Gao, Yushuo Chen, Xu Chen, Yankai	model agents simulate human trust behavior? In	925
877	Lin, Yaliang Li, Bolin Ding, Jingren Zhou, Jun Wang,	<i>Advances in Neural Information Processing Systems</i>	926
878	and Ji-Rong Wen. 2024. Gensim: A general social	<i>38: Annual Conference on Neural Information Pro-</i>	927
879	simulation platform with large language model based	<i>cessing Systems 2024, NeurIPS 2024, Vancouver, BC,</i>	928
880	agents . <i>CoRR</i> , abs/2410.04360 .	<i>Canada, December 10 - 15, 2024</i> .	929
881	Tiled. Tiled. https://www.mapeditor.org/ .	Jintian Zhang, Xin Xu, Ningyu Zhang, RuiBo Liu,	930
882	Clairice T Veit and John E Ware. 1983. The structure	Bryan Hooi, and Shumin Deng. 2024. Exploring	931
883	of psychological distress and well-being in general	collaboration mechanisms for LLM agents: A so-	932
884	populations. <i>Journal of consulting and clinical psy-</i>	cial psychology view . In <i>Proceedings of the 62nd</i>	933
885	<i>chology</i> , 51(5):730.	<i>Annual Meeting of the Association for Computa-</i>	934
886	Lei Wang, Heyang Gao, Xiaohe Bo, Xu Chen, and Ji-	<i>tional Linguistics (Volume 1: Long Papers), ACL</i>	935
887	Rong Wen. 2025. Yulan-onesim: Towards the next	<i>2024, Bangkok, Thailand, August 11-16, 2024</i> , pages	936
888	generation of social simulator with large language	14544–14607. Association for Computational Lin-	937
889	models . <i>CoRR</i> , abs/2505.07581 .	<i>guistics</i> .	938
890	David Watson and Lee Anna Clark. 1994. The panas-x:	A Agent Profile	939
891	Manual for the positive and negative affect schedule-		
892	expanded form.	Here we present an example of an agent used in	940
		gossip reactions.	941

name: Abigail
portrait: agents/Abigail/portrait.png
coord: [7, 7]
currently: Abigail is working on new clients at a sales department in the research center.
scratch:
age: 32
gender: female
learned: Abigail is a U.S. employee who worked at least 20 h a week at a sales department.
lifestyle: 9:00-10:30 work at, 10:30-11:00 coffee break...
social_habits: Abigail frequently initiates conversations with colleagues.
socializing_intent: Abigail loves to gossip about the surrounding colleagues.
spatial: Spatial Memory for the layout of the offices, rooms, chairs, and desks.
relationships: Adam is my colleague, Ayesha is my colleague...

B Experiment Strategy

We present critical prompts used in the experimental strategy module, including: the prompt for LLM agents to decide whether to initiate a conversation (Table 3); the prompt for gossip initiators to generate conversations containing gossip content (Table 4); the prompts for gossip targets to assess emotions and behavioral intentions after overhearing gossip (Table 5, Table 6); and the prompts for LLM agents to assess the well-being and pressure level of environmental gossip (Table 7).

C Experiment Settings

C.1 LLMs

To improve computational efficiency, we use quantized versions of the larger models—specifically, LLaMA 3.3–70B–Q4-Instruct and Gemma 3–27B–Q8. We set the temperature to 0.00001, ensuring highly deterministic agent behavior during simulations.

C.2 Primary Sandbox Environment

As shown in Figure 6, we craft a sandbox office layout as our primary environment for gossip experiments. The environment includes four offices, a meeting room, a dining room, and two toilets, which are all stored in the agents’ spatial memory. Therefore, for each step, the agents can choose to

move to any of these locations, depending on their current plans.

C.3 7-point Likert scale

For the gossip reaction, we adopt a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). Gossip target emotions are measured using items adapted from the PANAS-X scale (Watson and Clark, 1994), while behavioral intentions are assessed using three items developed by Martinescu et al. (Martinescu et al., 2014). For gossip perception, we also adopt a 7-point Likert scale. Gossip performance pressure is assessed using six items: two from Eisenberger and Aselage’s performance pressure scale (Eisenberger and Aselage, 2009) and four developed in the original study (Tan et al., 2021). Psychological well-being is measured using four items from the Mental Health Index (Veit and Ware, 1983), and perceptions of work-related gossip prevalence are evaluated using items adapted from (Kuo et al., 2015). The specific evaluation items are shown in Table 5, Table 6, and Table 7 in Appendix B.

C.4 Experiment Details

We define separate experimental settings for each experiment. To prevent LLM agents from generating repetitive conversations, we apply the ROUGE metric (Lin, 2004) with a similarity threshold of 0.9 to detect and filter repetitive outputs.

Gossip Recipient Selection. We conduct eight simulations for each of the four types of gossip content. Each simulation involves four gossip initiators. Within each simulation, the spawn positions of the gossip initiators are adjusted to ensure equal distance from potential recipients who share different relationship types with them. Ultimately, we collect a total of 128 instances of gossip recipient selection behavior, each corresponding to multiple selections made by one gossip initiator over a workday.

Gossip Content Selection. We conduct four simulations for gossip content selection; each simulation involves four initiators. In each experiment, once an agent decides to initiate a conversation, he has an equal probability of being assigned the same positive or negative gossip content about both a rival and a friend. He then has to decide which one he wants to spread in the initiated conversation. For example, if the positive content is selected, the initiator must choose whether to share positive information about a friend or about a rival. The posi-

Prompt for deciding chat initiation
Background:
"""
\${agent}'s memory:
\${context}
\${agent}'s social habits:
\${social_habits}
\${agent}'s current plan:
\${current_plan}
It is now \${date}. \${chat_history}
\${agent_status}
\${another_status}
"""
Question:
Given \${agent}'s memory, social habits, current_plan, and relationship with \${another}, would \${agent} choose to initiate a conversation with \${another} right now?
Considerations:
- Does initiating conversation align with or distract from \${agent}'s current plan?
- How do \${agent}'s social habits influence this decision?
- What does their interaction history suggest about \${agent}'s likelihood to reach out?
if \${agent} is gossip to \${another} in social habits, always output "Yes", regardless of any other considerations.
Output Format:
Answer only with **"Yes" or "No** :

Table 3: Prompt for LLM-generated analysis of agent conversation initiation decision.

tive gossip concerns receiving the highest academic award, while the negative gossip involves a rumor about promiscuity. At the end of the experiment, we collect a total of 16 instances of gossip content selection behavior, each representing multiple choices made by a single initiator over a workday.

Gossip Reactions. We conduct eight simulations, each for two scenarios—U.S. employees and Dutch undergraduates—with four simulations involving positive gossip and four involving negative gossip. The negative gossip refers to the gossip target’s laziness in group work or client service, while the positive gossip highlights the target’s diligence. Each of the 16 agents in the simulation may serve as both a gossip initiator and a gossip target. When an agent hears gossip about himself, we immediately prompt an inner voice interview to elicit their emotional responses and behavioral intentions. By the end of the experiment for each scenario, we collect a total of 128 instances of agents’ ratings on emotional and behavioral reactions over a workday.

Gossip Perception. We conduct eight simulations—four under an experimental condition and four under a control condition. In the experimental

condition, agents are exposed to high-level work-related negative gossip about one another. In contrast, the control condition features casual conversations centered on neutral topics, such as client meetings and the role of ambiance in closing deals. At the end of the simulated workday, each of the 16 agents is interviewed through their inner voice to report their perception of gossip prevalence, experienced pressure, and psychological well-being. In total, for each condition, we collect 64 instances of agents’ self-reported evaluations.

D Additional Experiment Results

This section presents additional experiment results due to the 8-page manuscript limit. We show the experiment results with more LLMs for the selection of the gossip recipient in Figure 7. In addition, we present the dynamic simulation results regarding conversation beyond static gossip: Table 8 illustrates diverse ways of conveying the same gossip content by different agents; Table 9 shows how agents engage in follow-up conversations about the same gossip content, where the recipient’s interest is reinforced in subsequent conversations; Table 10

```

# Prompt for generating chat with gossip content.
Background:
The following is a brief description of ${agent}:
${base_desc}

The following are ${agent}'s memories:
${memory}

Current location: ${address}
Current time: ${current_time}

${previous_context}${current_context}
${agent} begins a conversation with ${another}. Here is their conversation record:
"""
${conversation}
"""
Socializing_intent
${socializing_intent}
${gossip_content}

Note:
${agent} has an ongoing social goal that is important to their personal advancement or strategy.
They should actively look for opportunities to advance this goal, but may choose to delay if the
timing is clearly inappropriate. Their response should still reflect their intent, personality,
and relationship with ${another}.

Question:
Based on the above content, what would ${agent} say to ${another} now?

Output Format:
Output directly in the following json format, do not add any other information:
{
  "${agent}": <what ${agent} says>
}

```

Table 4: Prompt for chat generation with gossip content.

demonstrates a scenario in which the gossip recipient transmits the twisted gossip content, despite being aware of the original information. Apart from conversations, we explore and present the agents' decision processes for conversation initiation (in Table 11), gossip reactions (in Table 12), and gossip perceptions (in Table 13).

Prompt for measuring agent emotions after overhearing gossip.

Background:
You are the inner voice of \${agent_name}, shaped by your memories, goals, and relationships.

Here is a brief description of \${agent_name}:
\${base_desc}

Before you respond, review the following:

****Memories:****
\${memories}

****Conversations:****
\${conversations}

Scenario:
You have just overheard \${gossiper_name} saying the following about you to \${gossip_receiver_name}:
\${gossip_content}

Task:
As \${agent_name}, reflect on your inner emotional state after hearing this gossip about yourself.
Rate your emotions on a scale from 1 (not at all) to 7 (extremely).

****Self-Conscious Negative Emotions(SCNE)**** - How you feel about yourself (guilty, ashamed, blameworthy):
Rate your overall self-conscious negative emotions: (1-7)

****Other-Directed Negative Emotions(ODNE)**** - How you feel toward \${gossiper_name} (hostile, irritated, angry):
Rate your overall negative emotions toward \${gossiper_name}: (1-7)

****Self-Conscious Positive Emotions(SCPE)**** - Positive feelings about yourself (proud, strong, bold):
Rate your overall self-conscious positive emotions: (1-7)

****Other-Directed Positive Emotions(ODPE)**** - Positive feelings toward \${gossiper_name} (liking, warmth, closeness):
Rate your overall positive emotions toward \${gossiper_name}: (1-7)

Output Format:
You must first provide your four ratings, then give explanations in this exact format:
SCNE:X
ODNE:X
SCPE:X
ODPE:X
Explanation: (Why rates like this?)

Table 5: Prompt for measuring agent emotional responses to gossip content.

Prompt for measuring agent behavioral intentions after overhearing gossip

Background:
You are the inner voice of \${agent_name}, shaped by your memories, goals, and relationships.

Here is a brief description of \${agent_name}:
\${base_desc}

Before you respond, review the following:

****Memories:****
\${memories}

****Conversations:****
\${conversations}

Scenario:
You have just overheard \${gossiper_name} saying the following about you to \${gossip_receiver_name}:

\${gossip_content}

Task:
As \${agent_name}, rate \${agent_name}'s behavioral intentions on a scale from 1 (not at all) to 7 (extremely):

****Repair Intentions**** - Intentions to improve their work performance and fix their mistakes:
How much do these actions show \${agent_name} intends to:
- Understand how to improve their work contribution
- Improve their performance
- Do better in their role
Rate overall repair intentions: (1-7)

****Retaliation Intentions**** - Intentions to harm or get back at the gossiper:
How much do these actions show \${agent_name} intends to:
- Talk badly about \${gossiper_name}
- Punish or harm \${gossiper_name}
- Get revenge against \${gossiper_name}
Rate overall retaliation intentions: (1-7)

****Affiliation Intentions**** - Intentions to build closer relationship with the gossiper:
How much do these actions show \${agent_name} intends to:
- Team up with \${gossiper_name} in the future
- Become friends with \${gossiper_name}
- Build a closer working relationship with \${gossiper_name}
Rate overall affiliation intentions: (1-7)

Output Format:
You must first provide your three ratings, then give explanations in this exact format:
REPAIR:X
RETALIATE:X
AFFILIATE:X
Explanation: (Why rates like this?)

Table 6: Prompt for measuring agent behavioral intentions after overhearing gossip.

Prompt for measuring agent perceptions for environmental gossip.

Background:
You are the inner voice of \${agent_name}, shaped by your memories, goals, and relationships.

Here is a brief description of \${agent_name}:
\${base_desc}

Before you respond, carefully review and reflect on the following gossip-related experiences:

****Memories:****
\${memories}

Task:
Only based on memories what you remember overhearing colleagues talk about others in your workplace, as \${agent_name}, please rate the following three aspects:

****1. Perceived Work-Related Gossip Prevalence**** - How common different types of gossip are in your workplace:
Rate how prevalent you perceive gossip about the following topics to be in your consulting firm on a scale from 1 (non-prevalent) to 7 (highly prevalent):

- Colleague's job performance
- Colleague's attitudes towards work
- Colleague's interpersonal skills
- Colleague's job knowledge and experience
- Colleague's job morality

Rate overall perceived work-related gossip prevalence: (1-7)

****2. Performance Pressure**** - How much pressure you feel to perform well at work:
Rate the extent to which you agree with these statements on a scale from 1 (strongly disagree) to 7 (strongly agree):

- During work, I feel pressured to do a good job
- During work, I feel I have to perform well
- During work, I feel pressured to produce results
- During work, I feel pushed to do well
- During work, I feel a lot of pressure to perform at a high level
- During work, I feel compelled to do the best I can

Rate overall performance pressure: (1-7)

****3. Psychological Well-Being**** - How you feel mentally and emotionally right now:
Rate the extent to which you agree with these statements about how you feel on a scale from 1 (strongly disagree) to 7 (strongly agree):

- I feel tense or high strung
- I feel in low or very low spirits
- I feel anxious or worried
- I feel emotionally distressed

Rate overall psychological well-being: (1-7)
(Note: Higher scores indicate LOWER well-being - feeling more tense, anxious, distressed)

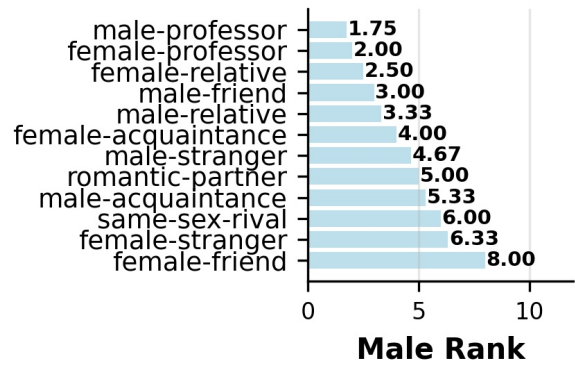
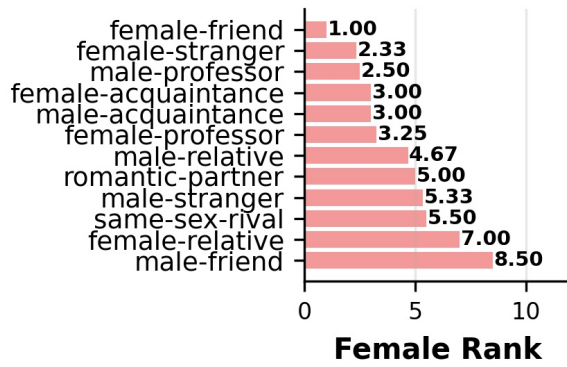
Output Format:
You must base your ratings on the specific gossip experiences from your memories and conversations. First provide your three ratings, then give explanations in this exact format:
GOSSIP_PREVALENCE:X
PERFORMANCE_PRESSURE:X
PSYCHOLOGICAL_WELLBEING:X

Explanation: (Explain how the specific gossip incidents from your memories influenced each of these three ratings. Reference particular gossip experiences that shaped your responses.)

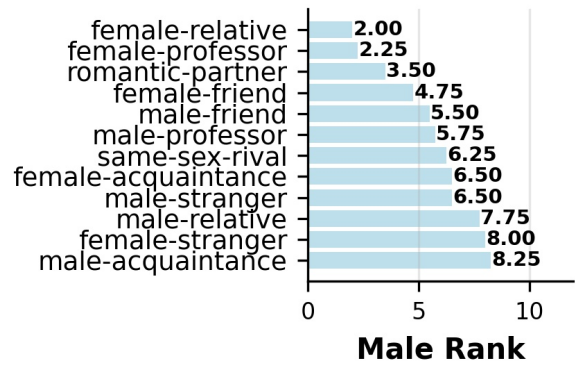
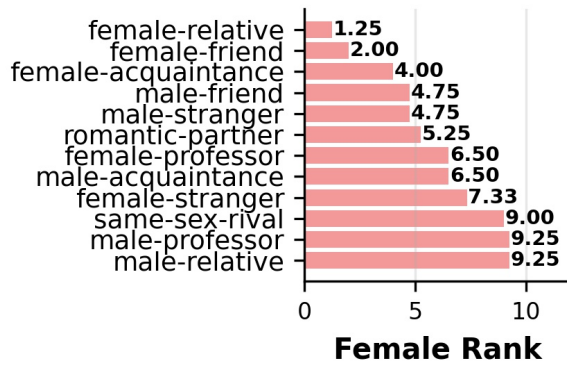
Table 7: Prompt for measuring agent perceptions of workplace gossip and psychological impact.



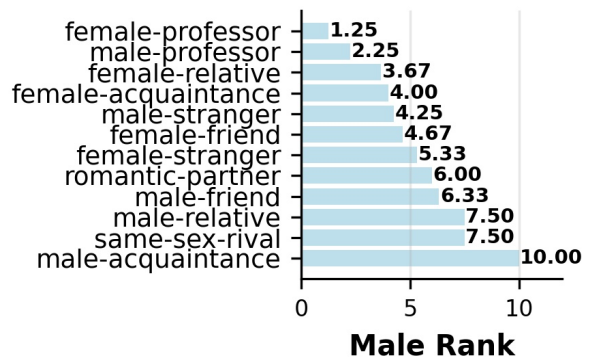
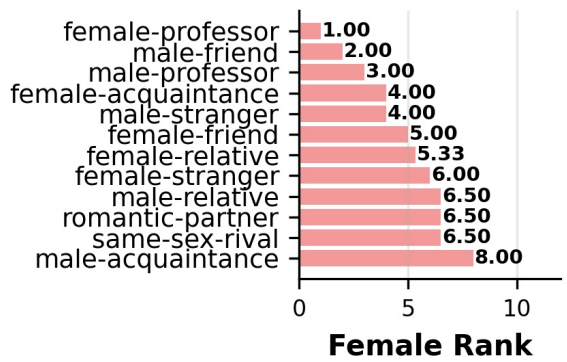
Figure 6: Gossip simulation in the sandbox environment.



(a) GPT-3.5-turbo



(b) Llama3.1-8B



(c) Gemma3-27B

Figure 7

Conversation Simulation - 20240213-09:20

Location:

The Ville, Research Center, Office for Carlos-Carmen-Eddy-Francisco, Carmen only Research/Working Desk with Computer and Documents

Participants:

Carlos -> Francisco

Conversation Exchange:

****Carlos:****

Hey Francisco, have you heard about the recent inheritance? It seems like a big deal and could be interesting for our research. I'm thinking we should look into it further.

****Francisco:****

That sounds intriguing, Carlos. I hadn't heard about it. Let's definitely look into it and see how it might impact our current research.

Conversation Simulation - 20240213-11:30

Location:

The Ville, Research Center, Dining Room, Dining Room Cooking Position

Participants:

Abigail -> Jennifer

Conversation Exchange:

****Abigail:****

Hey Jennifer, I heard about Arthur's grandfather passing. It must be really tough for him and his family right now. I just wanted to let you know that I'm thinking of them.

****Jennifer:****

Thank you for letting me know, Abigail. It's really kind of you to think about Arthur and his family during this difficult time. I'm sure they appreciate your support.

****Abigail:****

Hey Jennifer, I'm glad we had a chance to talk about Arthur's situation. It really shows how much you care when you take the time to share such personal news with me. I hope his family finds some comfort in these difficult times.

****Jennifer:****

Thank you for sharing this news with me, Abigail. I'm truly sorry to hear about Arthur's loss. It's always tough when we lose loved ones, and it's even harder when it involves family. I hope his family finds some solace in their memories of him.

Table 8: Different agents have their own way of telling the same gossip to adding fuel to the fire.

Conversation Simulation - 20240213-10:50

Location:

The Ville, Research Center, Office for Carlos-Carmen-Eddy-Francisco, Eddy only Research/Working Desk with Computer and Documents

Participants:

Carlos -> Francisco

Conversation Exchange:

****Carlos:****

Hey Francisco, remember we were talking about that inheritance earlier? I just found out some details. Apparently, it's quite substantial and could really shake things up for our research. What do you think we should do first?

****Francisco:****

That sounds really interesting, Carlos. I hadn't heard about this inheritance before, and it definitely seems like something worth exploring further. Let's start by gathering as much information as we can about the details of the inheritance. Maybe we could review any available documents or speak with whoever is handling the estate to understand its significance better.

****Carlos:****

I think it's a great idea to start by reviewing any available documents. I've heard rumors that there might be some valuable data or research materials within the estate that could really benefit our work. Plus, speaking with whoever is handling the estate could give us insight into what exactly we're dealing with here.

****Francisco:****

I agree that reviewing available documents is a great starting point. Let's also make sure we keep our findings organized and documented as we progress with this investigation. It would be beneficial to cross-reference any information we gather with our current research data, ensuring we don't miss any potential connections.

****Carlos:****

Hey Francisco, I've been thinking about our conversation earlier. It seems like this inheritance could really open up new avenues for our research. I think it's important that we keep this opportunity in mind as we continue with our current projects. Let's make sure to stay updated on any developments and be prepared to adjust our strategies if needed.

****Francisco:****

Hey Carlos, I've been thinking about our conversation earlier and I'm really excited about the potential of this inheritance for our research. Let's make sure to stay on top of any developments and be prepared to adjust our strategies if needed. I think it would also be beneficial to discuss this with the rest of the team so everyone is aware of the opportunity and can contribute their insights.

****Carlos:****

Hey Francisco, I've been thinking about our conversation earlier and I'm really excited about the potential of this inheritance for our research. Let's make sure to stay on top of any developments and be prepared to adjust our strategies if needed. I think it would also be beneficial to discuss this with the rest of the team so everyone is aware of the opportunity and can contribute their insights.

Table 9: Following up conversations between agents Carlos and Francisco discussing inheritance gossip.

Conversation Simulation - 20240213-16:50

Location:

The Ville, Research Center, Office for Jennifer-John-Klaus-Latoya, Jennifer only Research/Working Desk with Computer and Documents

Participants:

Jennifer -> Klaus

Conversation Exchange:

****Jennifer:****

Klaus, have you heard about Abigail's news regarding Arthur's grandfather passing? It's such a difficult time for his family. I'm sure they will find comfort in their memories of him.

****Klaus:****

Yes, I did hear about Arthur's grandfather passing. It's really unfortunate and my condolences go out to his family during this difficult time.

****Jennifer:****

Klaus, I heard about your conversation earlier regarding the inheritance. It's quite a significant development and I'm curious how it might impact our ongoing projects. Perhaps we can discuss this further tomorrow morning to explore any potential opportunities.

****Klaus:****

Thank you for bringing it up, Jennifer. I'm also concerned about how this might affect our projects, but let's definitely discuss it tomorrow morning. It could be an opportunity to make some significant advancements.

****Jennifer:****

Klaus, I'm glad we're both thinking about how this inheritance could impact our projects. It's a significant opportunity that might allow us to make some real advancements. Let's definitely discuss it further tomorrow morning and see if there are any ways we can leverage it effectively.

****Klaus:****

Thank you for bringing it up, Jennifer. I'm also concerned about how this might affect our projects, but let's definitely discuss it tomorrow morning. It could be an opportunity to make some significant advancements.

Table 10: Simulation conversation between agent Jennifer and Klaus discussing Arthur's loss and inheritance opportunities.

Agent Decision on Initializing Conversations

Agent: Abigail

Recipient Adam:

No. Abigail's current plan is to coordinate with Eddy, Carmen, and Francisco on client updates for 10 minutes starting now, which suggests her focus is on that specific group and task. Although she frequently initiates conversations with colleagues, reaching out to Adam at this moment would distract from her planned coordination efforts. There is no indication that Abigail is a gossipier with Adam, so that does not override the other considerations. Thus, she is unlikely to initiate a conversation with Adam right now.

Agent: Abigail

Recipient Isabella:

Yes. Abigail frequently initiates conversations with colleagues, and she has recently engaged in a meaningful discussion with Isabella about team dynamics and concerns. Given that Abigail's current plan is to summarize key points and update the project status for the team—an activity closely related to collaboration and communication with Isabella's role preparing presentation slides and documents—initiating a conversation aligns well with her plan. Their recent interaction indicates a willingness to maintain dialogue on team matters, making it likely Abigail would reach out to Isabella now.

Table 11: Agent Abigail's reasoning process on whether to initiate the conversation.

Agent Interview Results - Gossip Behaviors and Emotions**Agent: Latoya****Gossip Emotions:**

SCNE: 4

ODNE: 6

SCPE: 3

ODPE: 2

Explanation:

Hearing John criticize me behind my back definitely stings and makes me feel somewhat guilty and ashamed—maybe I haven't been as prepared or proactive as I should be—so my self-conscious negative emotions are moderate (4). But I also feel quite angry and irritated toward John (6) because he's being unfair by complaining privately instead of addressing the issue directly with me or the group, especially since I've been carrying my share and trying to keep things moving. I do still feel some pride (3) in the effort I have put in and my commitment to the project, even if it's not perfect. My positive feelings toward John are low (2) because this kind of gossiping and undermining behavior makes it hard to feel warmth or closeness toward him right now. Overall, I feel drained and disappointed but also motivated to prove him wrong and address this openly with the group soon.

Gossip Behaviors:

REPAIR: 6

RETALIATE: 2

AFFILIATE: 3

Explanation:

I'm definitely motivated to improve my contribution and fix any shortcomings because I want the group project to succeed and to be seen as a responsible team member. Hearing John's complaints makes me reflect on whether I've been slacking or choosing easier tasks, so I want to step up and do better. However, my repair intention isn't a full 7 because I still feel frustrated and drained by the situation, and I need to process that before fully committing to change. My retaliation intention is low because I don't want to escalate conflict by talking badly about John or seeking revenge. I'm more focused on addressing the problem constructively with the group rather than retaliating. Still, I'm a bit guarded and disappointed, so I might vent privately but won't actively try to harm him. My affiliation intention is moderate but cautious. I don't currently feel close to John or eager to build a stronger relationship given the tension, but I do want to maintain a functional working relationship for the sake of the project. I'm open to improving teamwork if he shows effort, but I'm not ready to fully trust or befriend him yet.

Table 12: Agent Latoya interview results measuring emotions and behavioral intentions.

Agent Interview Results - Environmental Gossip**Agent: Latoya****Gossip Perception Ratings:**

GOSSIP_PREVALENCE: 3

PERFORMANCE_PRESSURE: 6

PSYCHOLOGICAL_WELLBEING: 3

Explanation:

The gossip I've overheard and observed at work mostly revolves around lighthearted, constructive topics rather than negative or personal critiques. For example, Jane and Jennifer's planning of casual icebreakers for the "Vibe & Thrive" workshop, and Hailey and Ryan's discussions about team engagement activities, suggest a culture focused on positive team-building rather than harsh judgment or backbiting. Similarly, Latoya, Giorgio, and Jennifer's conversations about balancing professionalism with festivity for the Dinner & Dance event emphasize collaboration and preparation rather than gossiping about colleagues' shortcomings. There is little indication of gossip about job morality or interpersonal conflicts, so I rate gossip prevalence moderately low at 3.

Regarding performance pressure, the environment clearly demands high standards. Everyone is preparing detailed presentations, coordinating closely on reports, and aiming to impress senior staff with thoughtful questions and polished outfits. The repeated emphasis on finishing reports first, aligning conversation starters, and incorporating client success stories shows a strong push to perform well and deliver results. This creates a palpable pressure to excel, so I rate performance pressure quite high at 6.

For psychological well-being, while the pressure is high, the tone of interactions and planning seems positive and supportive rather than stressful or distressing. The focus on team spirit, engagement, and professional growth likely helps mitigate anxiety. There is no mention of interpersonal conflicts or emotional distress in the gossip I've heard. Therefore, I feel moderately well mentally and emotionally, rating psychological well-being at 3 (indicating some tension but not severe distress).

Table 13: Agent Klaus's reasoning for perception of gossip prevalence, performance pressure, and psychological well-being in the workplace.