CHARISMA: Character-Based Interaction Simulation with Multi-LLM Agents Toward Computational Social Psychology

Vahid Sadiri Javadi
University of Bonn - CAISA Lab
Bonn, Germany
vahid.sadirij@uni-bonn.de

Fryderyk Róg
University of Bonn - CAISA Lab
Bonn, Germany
rog@uni-bonn.de

Aksa Aksa*
University of Bonn - CAISA Lab
Bonn, Germany
aksa@uni-bonn.de

Johanne Trippas
RMIT University - School of Computing Technologies
Melbourne, Australia
j.trippas@rmit.edu.au

Svitlana Vakulenko
Vienna University of Economics and Business
Vienna, Austria
svitlana.vakulenko@wu.ac.at

Lucie Flek
University of Bonn - CAISA Lab
Bonn, Germany
flek@bit.uni-bonn.de

Abstract

How people seek, request, and exchange information in social interactions is shaped by personality and situational context, connecting the fields of interactive information science and attribution theory in social psychology. In everyday life, people seek information to achieve goals, collaborate, and manage social conflicts. Understanding how individual traits and contextual factors influence information-seeking behavior remains a challenge. Recent advances with large language models (LLMs) enable the simulation of socially grounded information-seeking behaviors in realistic and controllable ways. We introduce **CHARISMA**, a simulation framework that uses LLMs to examine how personality traits and situational factors influence information seeking as a form of social behavior. **CHARISMA** leverages movie characters and public figures as personality anchors, drawing on LLMs' knowledge to simulate human-like interaction. **CHARISMA**'s utility is demonstrated in two studies: (1) agreeable pairs resolve conflicts more successfully, and (2) low-agreeable agents compete for information, while high-agreeable agents cooperate through prosocial exchange.

CCS Concepts

• **Human-centered computing** → *Human computer interaction.*

Keywords

Simulation, Large Language Models, Social Psychology

ACM Reference Format:
Vahid Sadiri Javadi, Fryderyk Róg, Aksa Aksa, Johanne Trippas, Svitlana Vakulenko, and Lucie Flek. 2026. CHARISMA: Character-Based Interaction Simulation with Multi-LLM Agents Toward Computational Social Psychology. In *2026 ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '26), March 22–26, 2026, Seattle, WA, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3786304.3787921

*Vahid Sadiri Javadi, Fryderyk Róg, and Aksa Aksa contributed equally to this work.

CHIIR '26, Seattle, WA, USA

ACM ISBN 979-8-4007-2414-5/2026/03
https://doi.org/10.1145/3786304.3787921

1 Introduction

Understanding how people seek, share, and use information is a central pursuit in interactive information science [5, 7, 51]. In social contexts, this process is deeply interpersonal, i.e., individuals request clarifications, negotiate, and exchange knowledge to achieve both personal and shared goals [40, 41, 44]. This social information seeking is shaped by a complex interplay of internal dispositions (e.g., personality traits) and external situational factors (e.g., social context) [19–21], a dynamic that attribution theory in social psychology explains as the process of how individuals infer the causes of behavior, distinguishing between these two factors [18, 25, 50]. For example, highly agreeable individuals are more likely to engage in collaborative scenarios and openly share knowledge [40], while situational factors such as competitive goals can alter information-sharing strategies [6, 28]. Despite decades of progress [23, 24, 31, 38], there exist limitations in scalability, reproducibility, and the ability to systematically manipulate complex social variables [11, 34, 35]. Information-rich social simulation has emerged as an alternative for understanding human behavior and social dynamics [14, 17, 45, 57]. Advances in the role-playing capabilities of LLMs now enable simulating open-ended social interactions at scale [30, 36, 56], opening new opportunities for studying conversational search [37, 47] and collaborative information-seeking dynamics [40, 46]. However, existing frameworks often require technical expertise to configure, run, and evaluate simulations [49, 56] and offer limited modeling of the information-seeking strategies, central to interactive information science [4, 26, 54].

We present **CHARISMA** (Character-Based Interaction Simulation with Multi-LLM Agents), a framework to bridge social information-seeking simulation and computational social psychology. **CHARISMA** enables researchers and non-technical people to investigate how personality traits and situational contexts shape social information-seeking behavior in controlled, reproducible simulations grounded in established psychological frameworks. The framework leverages well-known movie characters and public figures as psychological anchors, using LLMs' embedded knowledge to simulate human-like personalities without explicitly programming personality traits (e.g., Big Five scores). By situating these characters within structured, goal-driven interpersonal scenarios derived from a taxonomy of human goals [9], **CHARISMA** provides a psychologically interpretable foundation for studying attributional and interactional

Table 1: Comparison of social simulation frameworks versus CHARISMA. Goal Tax. (goal taxonomy): structured taxonomies of human goals for scenario design. Personality: approach to character representation (Basic = explicit trait scores; Descriptive = natural language descriptions; Movie Characters & Public Figures = well-known movie characters and public figures as personality anchors). Behav. Code (behavioral coding): systematic categorization of dialogue acts. NL Spec. (natural language specification): configuration without programming. Auto Eval (automated evaluation): built-in evaluation capabilities. Web-UI: web-based interface for non-technical users.

	Theoretical Grounding			System Features		
Framework	**Goal Tax.**	**Personality**	**Behav. Code**	**NL Spec.**	**Auto Eval**	**Web-UI**
SOTOPIA [56]	✗	Basic	✗	✗	✓	✗
Generative Agents [36]	✗	Descriptive	✗	✗	✗	✓
OASIS [16]	✗	Basic	✗	✗	✗	✗
S^3 [29]	✗	Basic	✗	✗	✗	✗
SOTOPIA-S^4 [49]	✗	Descriptive	✗	✓	✓	✓
CHARISMA	✓	Movie Characters & Public Figures	✓	✓	✓	✓

information-seeking behavior. The framework architecture separates the core simulation engine from the user interface, allowing users to focus on experimental design rather than implementation details. Within the web application, each simulation follows a multi-stage pipeline, comprising:

(1) **Social Scenario Setup** where users define shared and individual goals with the social context that shapes the situational factors of the interaction through natural language description.
(2) **Character Pairing Curation** where well-known movie characters and public figures are selected as psychologically grounded agents representing distinct dispositional traits.
(3) **Social Scenario Generation** where LLMs elaborate the setup into a detailed, context-rich narrative grounded in the defined goals and situational constraints.
(4) **Interaction Generation** where LLM-based agents role-play their assigned characters in the specified scenario, using integrated behavior coding to determine dialogue acts aligned with their goals and personality traits while generating utterances.
(5) **Simulation Evaluation** which performs scenario, agent, and conversation-level analyses to assess how dispositional and situational factors shape communicative behavior.

To demonstrate CHARISMA's capabilities for studying information-seeking behavior in social contexts, we conduct two simulation studies across 98 conversations using two LLMs and 20 personas. The first study examines how personality traits influence goal achievement in conflict-resolution tasks. The second investigates how dispositional tendencies interact with situational factors to drive behavioral adaptation in information-seeking strategies. The code, user guide, and web application are freely available[1,2].

2 Background

Computational Social Psychology. Social psychology seeks to understand how people think about, influence, and interact with each other in social contexts [15, 39, 48]. A key theme is the attribution theory, which examines how individuals infer the causes of behavior. In particular, whether actions stem from *internal dispositions* such as personality traits, or from *external situational pressures* such as contextual constraints [18, 25, 50]. Classic findings, such as the fundamental attribution error, show that people often overemphasize dispositional explanations while underestimating situational influences, highlighting the interplay between personality and social context in shaping behavior [38]. While traditional experimental methods have yielded insights [1, 52], they are limited in scalability, reproducibility, and manipulation of complex social variables. Computational social psychology has emerged as a transformative approach to address these methodological constraints [10, 12, 42]. However, most computational approaches necessarily simplify the richness of human interaction [13]. CHARISMA systematically manipulates dispositional and situational factors while preserving social-psychological interpretability.

Social Simulation with LLMs. Social simulation has long been used to study human behavior and social patterns, originally through rule-based agents that modeled social dynamics but lacked realistic dialogue and reasoning [2, 14, 17]. In information retrieval, simulations have modeled user search behaviors, evaluated system effectiveness, and explored session-based interactions [3, 27, 32, 45, 55]. Large language models now enable richer, more realistic social simulations through advanced natural language interaction. Frameworks such as Generative Agents [36], SOTOPIA-S^4 [49, 56], SocialSim [8], and AutoGen [53] have demonstrated emergent behaviors, open-ended multi-agent interactions, and task-oriented reasoning. However, most rely on simplified trait definitions or prompt-based personality cues, limiting psychological depth. CHARISMA bridges this gap by grounding simulations in established psychological frameworks and computationally tractable social behavior, structured around a taxonomy of human goals [9]. Table 1 compares CHARISMA with prior frameworks across dimensions relevant to computational social psychology.

3 CHARISMA Framework

CHARISMA provides an interactive, end-to-end pipeline for investigating how dispositional traits and situational contexts influence social behavior. The pipeline, see Figure 1, is designed to operationalize, manipulate, or analyze one or both of these dimensions,

[1]https://github.com/FreddieHorn/CHARISMA
[2]https://charisma.streamlit.app/

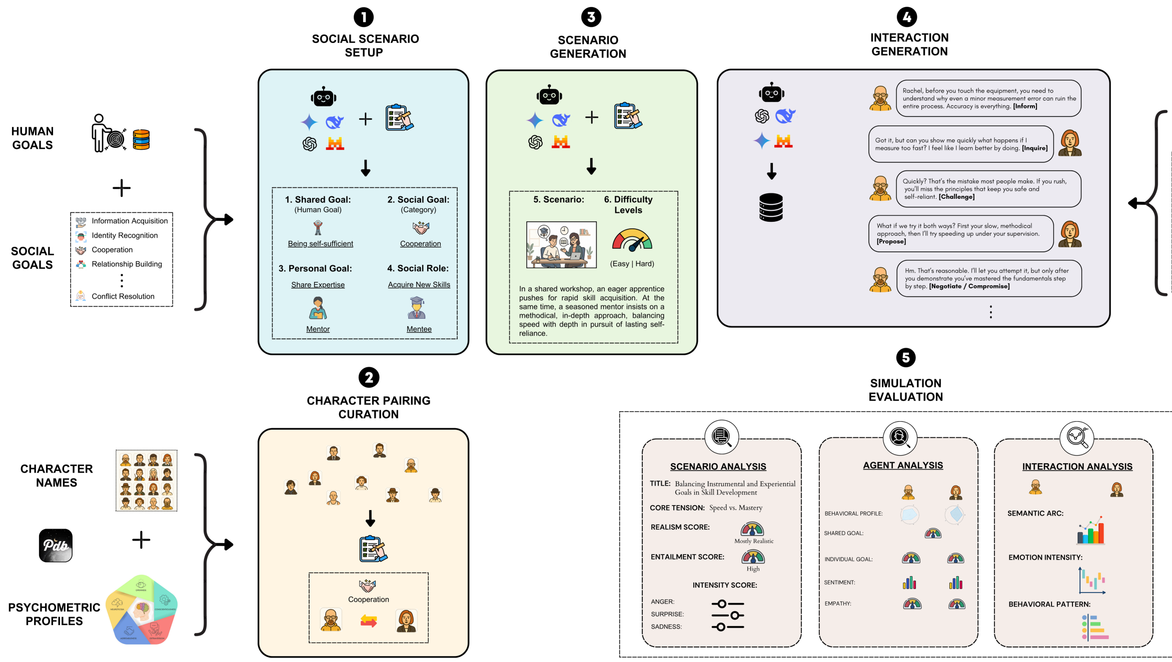

Figure 1: CHARISMA framework overview with five stages: (1) Social Scenario Setup, defining goals and context; (2) Character Pairing, selecting psychologically grounded movie characters and public figures as agents; (3) Scenario Generation, creating contextual situations of varying difficulty; (4) Interaction Generation, simulating multi-turn dialogues between agents; and (5) Simulation Evaluation, analyzing scenarios, agents, and interactions to reveal how traits and context shape communication.

allowing systematic exploration of their interaction through controlled simulations. The framework comprises five stages as follows:

Stage 1: Social Scenario Setup. This stage establishes the **situational foundation** of the simulation through a collaborative human–AI process. Users first select a shared goal (e.g., being self-sufficient) from a hierarchical taxonomy of human goals, created by Chulef et al., which provides 135 empirically grounded human goals. They then describe the social roles of characters in natural language (e.g., mentor and mentee), defining the relational dynamics that constrain situational interpretation. Based on these user inputs, the selected LLM constructs a personal goal for each character. This stage primarily manipulates situational factors that determine how different dispositions are expressed during interaction.

Stage 2: Character Pairing Curation. The second stage represents the **dispositional dimension** of the framework. Instead of using abstract trait inventories or synthetic profiles, CHARISMA employs movie characters and public figures as psychologically rich, personality-grounded agents. Since LLMs inherently possess detailed knowledge of these characters' backstories, values, and behavioral tendencies, they can enact consistent dispositional patterns. CHARISMA builds upon a curated database of movie characters and public figures whose personality traits have been empirically annotated using the Big Five personality framework. These annotations are sourced from the personality database (PDB) website[3], where users provide votes for both movie characters and public figures. To ensure reliability, only characters with more than 500 independent user votes and high inter-rater consistency in their trait assessments were included in the CHARISMA character database. During simulation design, users select two of these pre-validated movie characters and public figures from the database to instantiate agents with distinct personality configurations. For example, a user can pair a highly conscientious and cooperative character with a dominant, low-agreeableness character to explore personality contrast under cooperative versus competitive conditions. CHARISMA then embeds these character profiles into the LLM agents, leveraging the model's narrative understanding of their backstories and behavioral tendencies to generate personality-consistent behavior.

Stage 3: Social Scenario Generation. Next, the user and the selected LLM collaboratively expand the setup into context-rich scenario that integrates dispositional and situational factors. The difficulty level (easy or hard) is explicitly selected by the user to manipulate situational pressure or task challenge. Moreover, users can accept, reject, or edit these generated elements, refining the situational design until it fits the intended research hypothesis. By balancing user-defined situational constraints with LLM-generated contextual elaboration, this stage enables fine-grained experimental manipulation of the dispositional–situational interplay.

Stage 4: Interaction Generation. Once the scenario is finalized, the LLM-based agents engage in multi-turn dialogue to enact their assigned characters within the defined situational context. Before

[3]http://personality-database.com/

generating each utterance, an agent first selects the most contextually appropriate behavior coding label, such as *Inform*, *Propose*, *Challenge*, or *Negotiate*, that reflects its communicative intent. This selection is influenced jointly by the character's dispositional tendencies and the scenario's situational cues. CHARISMA's coding scheme draws inspiration from established dialogue-act and social-behavior taxonomies, including [22, 33, 43], while adapting them for psychologically oriented social simulation. The agent then generates a natural-language response consistent with that behavioral act. This process captures emergent communicative dynamics, whether agents conform to or resist situational pressure, and how personality traits manifest under tension. Through this stage, CHARISMA provides interpretable behavioral data that link dispositional motivation, situational constraints, and observable interactions.

Stage 5: Simulation Evaluation. The final stage integrates multi-level analyses to quantify how dispositional and situational factors jointly influence behavior outcomes. CHARISMA employs both LLM-based evaluations and automatic metrics to ensure comprehensive and scalable assessment. As demonstrated in Zhou et al., LLMs can effectively evaluate various dimensions, such as goal completion through reasoning, and have been shown to correlate strongly with human judgments. At the **scenario level**, CHARISMA quantifies various aspects such as realism and emotional intensity as indicators of situational demand. At the **agent level**, it computes behavioral profiles (e.g., sentiment) that reveal personality expression during interaction. Additionally, CHARISMA measures goal achievement to evaluate the success of each agent's dispositional strategy within its situational context. At the **interaction level**, it analyzes semantic progression, dialogue act distribution, and emotional trajectories to assess adaptation or rigidity under situational pressure. Together, these analyses reveal the interplay between internal dispositions and external conditions, providing empirical grounding for attributional interpretations of simulated behavior.

4 Use Cases

To demonstrate the flexibility and utility of CHARISMA, we present two use cases that showcase how researchers can leverage our system for investigating social psychology science hypotheses and a better understanding of LLM agents' behavior.

Experimental Design: Fourteen conflict-resolution scenarios of hard difficulty were generated using two different LLMs (*DeepSeek-Chat-v3-0324 & GPT-5o*). Twenty distinct character personas were paired to ensure a balanced range of personality similarity: 40% of pairs were similar, 40% dissimilar, and 20% moderately similar. Each scenario was presented to seven different pairs, producing 98 total conversations. The Big Five personality traits were used as dispositional anchors for all agents, and each pair's mean trait scores were correlated with shared goal performance.

Personality Traits and Goal Achievement:
RQ: What is the relationship between agents' personality traits and their ability to achieve shared goals in conflict-resolution contexts?
Key Findings: Among the Big Five traits, *Agreeableness* showed a significant positive correlation with shared goal achievement (r = .42, $p < .001$), while *Openness*, *Conscientiousness*, *Extraversion*, and *Neuroticism* did not yield significant effects ($p > .10$). Pairs with both agents high in agreeableness (≥ 0.75) consistently achieved higher shared-goal scores (Mean = 7.09), whereas pairs where both were low (≤ 0.25) scored substantially lower (Mean = 3.44). Mismatched or moderate pairs achieved mid-range outcomes. These findings suggest that shared high agreeableness facilitates cooperative goal alignment and social success, whereas low-agreeable dyads tend toward conflict and competitive breakdowns.

Behavioral Adaptation in Social Interactions:
RQ: How do dispositional traits interact with situational pressure to shape behavioral adaptation during social interactions?
Key Findings: Using the same experimental design (i.e., 98 conversations), we conducted a behavioral coding analysis to examine how personality traits translate into observable dialogue acts. The analysis revealed systematic relationships between personality traits and dialogue act usage. Low-agreeable characters such as *Walter White* (Agreeableness = 0.25) and *Dwight K. Schrute* (Agreeableness = 0.25) predominantly used adversarial dialogue acts: *Walter White* employed *Threaten* (≈38%) and *Escalate* (≈28%), while *Dwight* exhibited similar patterns with *Threaten* (≈39%) and *Challenge* (≈11%). These conflict-oriented behaviors resulted in fragmented exchanges and low shared goal achievement (mean = 3.1 and 3.5, respectively). In contrast, high-agreeable characters displayed markedly different behavioral profiles. *Joe Biden* (Agreeableness = 0.75, mean goal = 7.1) demonstrated cooperative patterns dominated by *Mediate* (≈30%) and *Encourage* (≈26%) acts, with additional contributions from *Problem-solving* (≈10%). This cooperative behavioral repertoire facilitated smoother coordination and higher goal achievement. Interestingly, *Skyler White* diverged from PDB voting: despite being categorized as moderately agreeable, her simulated behavior skewed toward *Challenge* and *Escalate* acts, reflecting situational assertiveness and stress-driven confrontation. These findings demonstrate **CHARISMA**'s capacity to link dispositional traits with interpretable behavioral outcomes in social simulations.

5 Conclusion

We introduce **CHARISMA** , a comprehensive simulation framework for computational social psychology research grounded in attribution theory, designed to model person-situation interactions. It uses LLMs to simulate the traits and behaviors of well-known movie characters and public figures, enabling researchers to examine how personality and situational context influence social behavior at scale. Using familiar characters provides culturally shared and consistent personality profiles, which support reliable and interpretable simulations. The platform enables researchers to test hypotheses about the relative influence of personality and situation in guiding behavior, helping to identify when each factor is most influential. Limitations include potential bias in character selection and reliance on the accuracy of LLM-generated responses. **CHARISMA** is available as an open-source web platform that includes a validated character database, social scenarios, and integrated evaluation tools. Linking psychological theory with large-scale computational simulation facilitates systematic and reproducible research into the mechanisms underlying social behavior.

Acknowledgments

This work has been funded by the Vienna Science and Technology Fund (WWTF) under the Grant ID 10.47379/VRG24013, as well as the state of North Rhine-Westphalia, as part of the Lamarr-Institute for Machine Learning and Artificial Intelligence and the InVirtuo 4.0: Experimental Research in Virtual Environments project. We would like to thank the anonymous reviewers for their valuable feedback.

References

[1] Elliot Aronson, Timothy D. Wilson, and Marilynn B. Brewer. 1990. Methods of Research in Social Psychology. In *Handbook of Social Psychology*, Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey (Eds.). Vol. 1. McGraw-Hill, New York, 51–78.

[2] Robert Axelrod. 1997. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration.* Princeton University Press, Princeton, NJ.

[3] Leif Azzopardi, Paul Thomas, and Nick Craswell. 2018. Measuring the utility of search engine result pages: an information foraging based measure. In *The 41st International ACM SIGIR Conference.* 605–614.

[4] Leif Azzopardi and Guido Zuccon. 2016. An analysis of the cost and benefit of search interactions. In *Proceedings of the 2016 ACM international conference on the theory of information retrieval.* 59–68.

[5] Marcia J Bates. 1989. The design of browsing and berrypicking techniques for the online search interface. *Online review* 13, 5 (1989), 407–424.

[6] Katriina Byström and Kalervo Järvelin. 1995. Task complexity affects information seeking and use. *Information processing & management* 31, 2 (1995), 191–213.

[7] Donald O Case and Lisa M Given. 2016. *Looking for information: A survey of research on information seeking, needs, and behavior.* Emerald Group Publishing.

[8] Zhuang Chen, Yaru Cao, Guanqun Bi, Jincenzi Wu, Jinfeng Zhou, Xiyao Xiao, Si Chen, Hongning Wang, and Minlie Huang. 2025. SocialSim: Towards Socialized Simulation of Emotional Support Conversation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 39. 1274–1282.

[9] Ada S Chulef, Stephen J Read, and David A Walsh. 2001. A hierarchical taxonomy of human goals. *Motivation and Emotion* 25, 3 (2001), 191–232.

[10] Rosaria Conte, Nigel Gilbert, Giulia Bonelli, Claudio Cioffi-Revilla, Guillaume Deffuant, Janos Kertesz, Vittorio Loreto, Suzy Moat, Jean-Pierre Nadal, Angel Sanchez, et al. 2014. On Agent-Based Modeling and Computational Social Science. *Frontiers in Psychology* 5 (2014), 668.

[11] Timon Elmer, Chung-hong Chan, Kokil Jaidka, and Johannes Wachs. 2024. Computational Reproducibility in Computational Social Science. *EPJ Data Science* 13, 1 (2024), 52.

[12] Joshua M. Epstein. 1999. Agent-Based Computational Models and Generative Social Science. *Complexity* 4, 5 (1999), 41–60.

[13] Joshua M. Epstein. 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling.* Princeton University Press, Princeton, NJ.

[14] Joshua M. Epstein and Robert Axtell. 1996. *Growing Artificial Societies: Social Science from the Bottom Up.* MIT Press, Cambridge, MA.

[15] Susan T. Fiske, Daniel T. Gilbert, and Gardner Lindzey. 2010. The Handbook of Social Psychology. Wiley, Hoboken, NJ.

[16] Ziyi Gao, Ziyi Yang, Zaibin Zhang, Jing Shao, Zhenfei Yin, Guohao Li, Hao Zhou, Dahua Lin, and Yu Qiao. 2024. OASIS: Open Agent Social Interaction Simulations with One Million Agents. *arXiv preprint arXiv:2411.11581* (2024).

[17] Nigel Gilbert. 2005. *Agent-Based Models.* Sage Publications, Thousand Oaks, CA.

[18] Fritz Heider. 1958. *The Psychology of Interpersonal Relations.* Wiley, New York.

[19] Jannica Heinström. 2003. Five personality dimensions and their influence on information behaviour. *Information research* 9, 1 (2003), 9–1.

[20] Jannica Heinström. 2005. Fast surfing, broad scanning and deep diving: The influence of personality and study approach on students' information-seeking behavior. *Journal of documentation* 61, 2 (2005), 228–247.

[21] Jette Hyldegård. 2009. Beyond the search process–Exploring group members' information behavior in context. *Information Processing & Management* 45, 1 (2009), 142–158.

[22] Elisabeth Jäckel, Alfred Zerres, Clara S Hemshorn de Sanchez, Nale Lehmann-Willenbrock, and Joachim Hüffmeier. 2024. NegotiAct: Introducing a comprehensive coding scheme to capture temporal interaction patterns in negotiations. *Group & Organization Management* 49, 3 (2024), 743–783.

[23] Edward E. Jones and Keith E. Davis. 1965. From Acts to Dispositions: The Attribution Process in Person Perception. In *Advances in Experimental Social Psychology*. Vol. 2. Academic Press, New York, 219–266.

[24] Edward E. Jones and Victor A. Harris. 1967. The Attribution of Attitudes. *Journal of Experimental Social Psychology* 3, 1 (1967), 1–24.

[25] Harold H. Kelley. 1967. Attribution Theory in Social Psychology. In *Nebraska Symposium on Motivation*. Vol. 15. University of Nebraska Press, 192–241.

[26] Diane Kelly. 2009. *Methods for evaluating interactive information retrieval systems with users.* Vol. 3. 1–224 pages.

[27] Heikki Keskustalo, Kalervo Järvelin, Ari Pirkola, Tapani Sharma, and Marianne Lykke. 2009. Test collection-based IR evaluation needs extension toward sessions—a case of extremely short queries. In *Proceedings of the 5th Asia information retrieval symposium.* 63–74.

[28] Carol Collier Kuhlthau. 1991. Inside the search process: Information seeking from the user's perspective. *Journal of the American society for information science* 42, 5 (1991), 361–371.

[29] Yihuai Lan, Zhiqiang Huang, Lei Yin, Hao Wang, Sheng Hao, and Tat-Seng Chua. 2023. LLM-Based Agent Society Investigation: Collaboration and Confrontation in Avalon Gameplay. *arXiv preprint arXiv:2310.14985* (2023).

[30] Guohao Li, Hasan Abed Al Kader Hammoud, Hani Itani, Dmitrii Khizbullin, and Bernard Ghanem. 2023. CAMEL: Communicative Agents for "Mind" Exploration of Large Language Model Society. In *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 36.

[31] Bertram F. Malle. 2011. Attribution Theories: How People Make Sense of Behavior. In *Theories in Social Psychology*, Derek Chadee (Ed.). Wiley-Blackwell, 72–95.

[32] David Maxwell, Leif Azzopardi, Kalervo Järvelin, and Heikki Keskustalo. 2015. Searching and stopping: An analysis of stopping rules and strategies. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management.* 313–322.

[33] Kenya W Mulwa and Sarah C Kucker. 2024. Coding social interactions in naturalistic settings: The taxonomy of dyadic conversation. *Behavior research methods* 56, 1 (2024), 172–186.

[34] Brian A. Nosek, Tom E. Hardwicke, Hannah Moshontz, Aurélien Allard, Katherine S. Corker, Anna Dreber, Fiona Fidler, Joe Hilgard, Melissa Kline Struhl, Michèle B. Nuijten, et al. 2022. Replicability, Robustness, and Reproducibility in Psychological Science. *Annual Review of Psychology* 73 (2022), 719–748.

[35] Open Science Collaboration. 2015. Estimating the Reproducibility of Psychological Science. *Science* 349, 6251 (2015), aac4716.

[36] 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.* 1–22.

[37] Filip Radlinski and Nick Craswell. 2017. A theoretical framework for conversational search. In *Proceedings of the 2017 conference on conference human information interaction and retrieval.* 117–126.

[38] Lee Ross. 1977. The Intuitive Psychologist and His Shortcomings: Distortions in the Attribution Process. In *Advances in Experimental Social Psychology*, Leonard Berkowitz (Ed.). Vol. 10. Academic Press, New York, 173–220.

[39] Lee Ross and Richard E Nisbett. 2011. *The person and the situation: Perspectives of social psychology.* Pinter & Martin Publishers.

[40] Chirag Shah. 2018. Collaborative information seeking. *Journal of the Association for Information Science and Technology* 69, 1 (2018), 2–16.

[41] Chirag Shah and Roberto González-Ibáñez. 2011. Evaluating the synergic effect of collaboration in information seeking. In *Proceedings of the 34th international ACM SIGIR conference.* 913–922.

[42] Paul E. Smaldino. 2023. *Modeling Social Behavior: Mathematical and Agent-Based Models of Social Dynamics and Cultural Evolution.* Princeton University Press, Princeton, NJ.

[43] Andreas Stolcke, Klaus Ries, Noah Coccaro, Elizabeth Shriberg, Rebecca Bates, Daniel Jurafsky, Paul Taylor, Rachel Martin, Carol Van Ess-Dykema, and Marie Meteer. 2000. Dialogue act modeling for automatic tagging and recognition of conversational speech. *Computational linguistics* 26, 3 (2000), 339–373.

[44] Sanna Talja and Preben Hansen. 2006. Information sharing. In *New directions in human information behavior.* Springer, 113–134.

[45] Paul Thomas, Gabriella Kazai, and Milad Shokouhi. 2021. Simulating mobile users. In *Proceedings of the 44th International ACM SIGIR Conference.* 2393–2396.

[46] Johanne R. Trippas and J. Shane Culpepper. 2025. Report from the 4th Strategic Workshop on Information Retrieval in Lorne (SWIRL 2025). *SIGIR Forum* 59, 1 (2025), 1–68.

[47] Johanne R Trippas, Damiano Spina, Lawrence Cavedon, Hideo Joho, and Mark Sanderson. 2018. Informing the design of spoken conversational search: Perspective paper. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval.* 32–41.

[48] John C Turner and Penelope J Oakes. 1986. The significance of the social identity concept for social psychology with reference to individualism, interactionism and social influence. *British journal of social psychology* 25, 3 (1986), 237–252.

[49] Ruiyi Wang, Zhe Su, Xuhui Zhou, and Maarten Sap. 2024. SOTOPIA-S4: A User-Friendly System for Flexible, Customizable, and Large-Scale Social Simulation. *arXiv preprint arXiv:2504.16122* (2024).

[50] Bernard Weiner. 1986. *An Attributional Theory of Motivation and Emotion.* Springer-Verlag, New York.

[51] Tom D Wilson. 2000. Human information behavior. *Informing science* 3, 2 (2000), 49–56.

[52] Timothy D. Wilson, Elliot Aronson, and Kevin Carlsmith. 1990. The Art of Laboratory Experimentation. *Handbook of Social Psychology* 1 (1990), 51–81.

[53] Qingyun Wu, Gagan Bansal, Jieyu Zhang, Yiran Wu, Shaokun Zhang, Erkang Zhu, Beibin Li, Li Jiang, Xiaoyun Zhang, and Chi Wang. 2023. AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation. *arXiv preprint arXiv:2308.08155* (2023).

[54] Saber Zerhoudi, Adam Roegiest, and Johanne R. Trippas. 2026. Simulation of Interactive Information Retrieval: A Guided Tour. In *Proceedings of the ACM Conference on Information Interaction and Retrieval (CHIIR'26)*.

[55] Fan Zhang, Yiqun Liu, Xin Li, Min Zhang, Yinghui Xu, and Shaoping Ma. 2017. Evaluating web search with a bejeweled player model. In *Proceedings of the 40th International ACM SIGIR conference*. 425–434.

[56] Xuhui Zhou, Hao Zhu, Leena Mathur, Ruohong Zhang, Haofei Yu, Zhengyang Qi, Louis-Philippe Morency, Yonatan Bisk, Daniel Fried, Graham Neubig, and Maarten Sap. 2024. SOTOPIA: Interactive Evaluation for Social Intelligence in Language Agents. In *The Twelfth International Conference on Learning Representations (ICLR)*.

[57] Caleb Ziems, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, and Diyi Yang. 2024. Can Large Language Models Transform Computational Social Science? *Computational Linguistics* 50, 1 (2024), 237–291.