My overarching goal is to study human-AI interaction for scaling human effort to solve complex societal problems. Many societal issues, e.g., climate change and health care, require decision-makers to study their stakeholders' behavior to design interventions or make a policy change. My research is motivated by the fundamental question of understanding humans at scale, i.e., how to conduct robust, generalizable, and engaging studies about human behavior? I study and create novel systems for scholars to expand our knowledge about ourselves. In my doctoral work, I have been examining this topic in the context of behavioral science research by integrating conversational AI into survey methods. As a long-established tool of querying human behaviors and preferences, survey informs countless decisions that impact our daily life. *Even a tiny improvement in survey quality could tremendously improve our life.* I design, build and examine conversational agents to conduct engaging surveys and collect high-quality information. My interdisciplinary background in computer science and social psychology allows me to couple explanatory social theories with novel algorithms and interactions. Through my research, I hope to create a more connected research community and democratize novel technologies to operationalize intuitions and curiosities about how we think and behave.

Improving survey research with conversational AI

In my thesis, I explored the role of conversational agents in conducting surveys through interviews, where an agent asks questions, interprets a participant's responses, and probes answers whenever needed. I addressed three challenges in today's survey methods: survey fatigue, inflexible survey structure, and poor personalization. I tackle this problem from both machine perspectives, where I build systems and natural language models for accessible and skillful interview chatbots [Xia+20a; Xia+a] and human perspectives where I studied how people interact with conversational agents with different designs [Xia+20b; XZF19; Xia+21].

Machine perspectives. How can we *build* engaging conversational agents to collect rich and truthful information through surveys? Building an effective interview conversational agent faces three major obstacles. First, the agent needs to effectively grasp and respond to participants' highly diverse and often complex natural language input. Second, the agent needs to handle complex conversation situations to deliver interview skills such as prompting and probing and to ensure survey completion. Third, it is difficult for researchers who don't have AI expertise or resources to take advantage of AI advances.

My first study examines the feasibility of using practical AI to build effective chatbots for survey interviews [Xia+20a]. I built a data-driven framework for rule-based chatbot platforms that enable an interview chatbot to interpret diverse respondent's inputs and generate proper responses with active listening skills. The framework contains a human-AI collaborative annotation module that prepares high-quality training data from small, unstructured real-world conversations and a classifier to trigger rules at run time to guide the generation of proper responses, enabling active listening. By focusing on publicly available techniques, models requiring little customization, and modular design, the framework supported reproducibility, practicality, and scalability. In the evaluation, the interview chatbot with active listening skills collected more informative and relevant responses while creating more engaging experiences. This work offers practical implementations to power chatbots with active listening skills, a hybrid chatbot design framework for developing progressive chatbot platforms, and design implications for building empathetic chatbots beyond interview tasks.

Having established the feasibility of using conversational AI to conduct surveys, the following step I focus on the challenge of static survey structure and lack of personalization, e.g., how could we enable conversational AI to generate follow-up questions to ask for more relevant and meaningful information [Xia+a]. Interview chatbots today are unable to parse the context of participants' responses and don't know what to ask next.

This results from ineffective follow-up questions like "could you say a bit more?". I wield a knowledge-driven approach with a deep neural language model (DialoGPT) to accomplish sophisticated follow-up question generation. I performed multi-task training to select salient knowledge entities and relations pairs in the context to teach the model what to ask. Subsequently, by identifying key contexts and knowledge, the model used prompt learning to generate follow-up questions that are coherent with the context and always aim for the most meaningful, deep, and interesting information. The model was deemed effective in both objective metrics and expert evaluations. With this work, future survey designers would create more dynamic and personalized survey without writing complicated rules and tons of question templates.

Human perspectives. Building a conversational agent to ask questions is merely the first step, what is less known is how respondents will react to such an agent. I conducted a series of studies to answer the following questions, 1) Could an agent drive an engaging survey while eliciting high-quality information? 2) How could we better design an agent to request information? 3) How does such an agent perform in the wild?

A conversational agent can frame survey questions in a more personalized manner, provide human-like social interactions, and encourage information exchange in a natural way while bearing the risk of inaccurately interpreting and improperly responding to humans' diverse free-text input due to is limited natural language capability. To see if the agent could conduct effective survey, I conducted a field experiment that compared the outcomes of an agent-driven survey and that of a typical form-based survey [Xia+20b]. To evaluate survey quality, I proposed a set of evaluation metrics based on Gricean Maxims, including the dimension of quantity, quality, relevance, and manner. Through an in-depth mixed-method analysis of over 5,200 free-text responses, I found the agent not only made the survey more engaging and but also collected significantly more relevant, specific, and clear responses. This study indicates a promising but currently under-utilized way of creating more effective and engaging survey through conversational agents.

In the second study, I answered the question of how we should design conversational agents to collect information [Xia+21]. In the context of using voice assistants to collect people's feedback for recommender systems in-situ. I examined three design dimensions of such an agent: who the agent is, how the agent should act and when the agent should act. This study has intriguing results: people are more likely to respond to the agent's feedback request and perceive it as less disruptive; comparing to direct question, giving user the instruction to provide feedback creates less experience friction; and when the agent becomes more proactive, people prefer the instruction more. With our results, we can now create more effective conversational agents to understand people's preference and intents in-situ.

How does such an agent perform in the wild? I built `Indigo`, a conversational agent that can chat with students to learn teaming preferences [XZF19]. `Indigo` was developed to replace traditional teaming surveys to gather initial information from team members, such as team preferences and expectations. `Indigo` interacted with 201 university students from an engineering class, formed 40 teams, and engaged in semester-long team projects. I found that students interacted with `Indigo` for an extensive period of time (e.g., 60-minute chat) and offered open and honest input. `Indigo` also elicited rich information that allowed instructors to better understand student team preferences and perceptions. This study indicates a viable path to create an AI teammate that can follow a team and interact with team members continuously during their team efforts.

In my thesis, I built and evaluated conversational agents to help decision-makers to better understand their stakeholders through survey. I also worked on other topics in human-AI interaction, including information-

seeking and behavioral nudging [Gao+18; Xia+b; Xia+19], social media [Vac+21], and educational technology [Xia+18; Wau+20]. My past research in human-AI interaction prepared myself with substantial skills and knowledge and shaped my future research direction.

Future Research

Looking ahead, I will continue my research on novel techniques to study human behavior. Three key aspects of human behaviors structure my future research agenda, *context*, *agency*, and *time span*. My goal is not only to enable scholars to study human behavior with *unprecedented granularity* but to deconstruct human-AI interactions with *explanatory theories*.

Sample social interactions. Asking the right question at the right time ensures data diagnosticity, enhances the external validity, and reduces redundant data. I want to explore sampling techniques beyond location, time, and activity in the near future, i.e., Could we sample social context and interactions? I am building a novel behavioral research platform, Petri, that combines mobile sensing and conversational agents to sense social contexts. I plan to develop sampling methods with few-shot learning to accurately query the participants while immersed in real-life situational contexts. The ability to probe into diverse and complex social interactions will empower behavioral scientists to dissect human behavior at a new level of granularity.

Rethinking informed consent in the age of AI. When AI agents interact with people to collect data, the experimenter has enormous power over the subject's data. However, more often than not the only power a participant could exert is to opt out. The power asymmetry exacerbates when more big corporations and governments deploy AI agents to work with their stakeholders. In my future research, I plan to look at two processes to empower participants, *informed consent*, how could AI help participants make informed decisions before joining the study, and *control over data*, how should we design system that grants participant's control over their own data with the consideration of academic rigor and unwanted data biases. Further, I will delve deeper into the power relationship and dynamic between an AI agent and people. What signals an AI agent's power status? How an AI agent's power status shapes human-AI collaboration? And How could we create AI agents that can manage power dynamics in human-AI collaboration? This line of work will foster a positive and trustworthy relationship among AI agents, experimenter, and study participants that ultimately benefits the society as a whole.

Conduct longitudinal study with AI. A longitudinal study reveals the rapid fluctuations in behaviors, thoughts, and emotions from moment to moment or day to day. However, it is expensive to keep participants engaged over a long period. The ubiquity of conversational agents opens opportunities to intervene and collect data through daily interaction. However, today's conversational agents are mostly designed for single-shot interactions. The conversational agent will face more challenges to conducting effective longitudinal studies: How can an agent adapt and react to unique individual experiences? How can an agent build and manage a long-term relationship with people? In the future, I will explore affective computing and knowledge-driven models to create agents capable of delivering engaging adaptive longitudinal studies. Such an AI agent will help researchers study individual behaviors across the lifespan, at scale, and low cost.

As AI becomes more and more embedded in our lives, there are big questions about effective interaction, power, and agency. While being mindful about privacy and personal information, my interest is to develop techniques that allow everyone to use AI to study human behavior.

References

[Gao+18] Mingkun Gao*, Ziang Xiao*, Karrie Karahalios, and Wai-Tat Fu. "To label or not to label: The effect of stance and credibility labels on readers' selection and perception of news articles". In: *Proceedings of the ACM on Human-Computer Interaction* 2.CSCW (2018), pp. 1–16.

[Vac+21] Kristen Vaccaro, Ziang Xiao, Kevin Hamilton, and Karrie Karahalios. "Contestability For Content Moderation". In: *Proc. ACM Hum.-Comput. Interact.* 5.CSCW2 (Oct. 2021). DOI: 10.1145/3476059. URL: https://doi.org/10.1145/3476059.

[Wau+20] Helen Wauck, Brian S Woodard, Ziang Xiao, Tiffany Wenting Li, and Brian P Bailey. "A Data-Driven, Player-Centric Approach to Evaluating Spatial Skill Training Games". In: *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. 2020, pp. 349–361.

[Xia+a] Ziang Xiao*, Yubin Ge*, Heng Ji, Jana Diesner, Karrie Karahalios, and Hari Sundaram. "What should I ask: A Knowledge-graph Based Approach for Follow-up Questions Generation in Conversational Surveys". In: *Submission*.

[Xia+b] Ziang Xiao, Tyrone Grandison, Q. Vera Liao, Michelle Zhou, and Yunyao Li. "Building AI-powered Chatbots with Expert Sourcing to Support Information Seeking during a Public Health Crisis: Lessons Learned and Design Implications". In: *Revised and Resubmit*.

[Xia+18] Ziang Xiao, Helen Wauck, Zeya Peng, Hanfei Ren, Lei Zhang, Shiliang Zuo, Yuqi Yao, and Wai-Tat Fu. "Cubicle: An adaptive educational gaming platform for training spatial visualization skills". In: *23rd International Conference on Intelligent User Interfaces*. 2018, pp. 91–101.

[Xia+19] Ziang Xiao, Po-Shiun Ho, Xinran Wang, Karrie Karahalios, and Hari Sundaram. "Should we use an abstract comic form to persuade? experiments with online charitable donation". In: *Proceedings of the ACM on Human-Computer Interaction* 3.CSCW (2019), pp. 1–28.

[Xia+20a] Ziang Xiao, Michelle X. Zhou, Wenxi Chen, Huahai Yang, and Changyan Chi. "If I Hear You Correctly: Building and Evaluating Interview Chatbots with Active Listening Skills". In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI '20. Honolulu, HI, USA: Association for Computing Machinery, 2020, 1–14. ISBN: 9781450367080. DOI: 10.1145/3313831.3376131. URL: https://doi.org/10.1145/3313831.3376131.

[Xia+20b] Ziang Xiao, Michelle X. Zhou, Q. Vera Liao, Gloria Mark, Changyan Chi, Wenxi Chen, and Huahai Yang. "Tell Me About Yourself: Using an AI-Powered Chatbot to Conduct Conversational Surveys with Open-Ended Questions". In: *ACM Trans. Comput.-Hum. Interact.* 27.3 (June 2020). ISSN: 1073-0516. DOI: 10.1145/3381804. URL: https://doi.org/10.1145/3381804.

[Xia+21] Ziang Xiao, Sarah Mennicken, Bernd Huber, Adam Shonkoff, and Jennifer Thom. "Let Me Ask You This: How Can a Voice Assistant Elicit Explicit User Feedback?" In: *Proc. ACM Hum.-Comput. Interact.* 5.CSCW2 (Oct. 2021). DOI: 10.1145/3479532. URL: https://doi.org/10.1145/3479532.

[XZF19] Ziang Xiao, Michelle X Zhou, and Wat-Tat Fu. "Who should be my teammates: Using a conversational agent to understand individuals and help teaming". In: *Proceedings of the 24th International Conference on Intelligent User Interfaces*. 2019, pp. 437–447.