Ghost and the Machine: Attitudes and Attribution of Mind of Artificial Intelligence

Roger Young, M.A.
University of South Florida

Andrew Kuntz
University of South Florida

Introduction

- Attribution of mind has important implications for cognitive and emotional trust.
- This study explores factors that may influence people's cognitive and emotional trust of text-based artificial intelligence (AI).
  - **Cognitive trust**: How much users trust AI to make logical, evidence-based decisions (e.g., trusting an AI for financial advice).
  - **Emotional trust**: How much users trust AI to understand emotional context and display empathy (e.g., trusting AI to keep a personal secret).
- AI software has tremendous potential to influence the attitudes, behaviors, and decision-making of its users (Crawford et al., 2024).
- While past research has measured user's attitudes towards AI (Schepman & Rodway, 2023) and examined how AI characteristics relate to user's trust (Glikson, 2021), this study offers novel insight into how individual differences influence trust in AI.

Methods

- We recruited forty-four ($N = 44$) undergraduate students to complete an online survey in exchange for course credit.
- The survey randomly assigned participants into two kinds of conditions: assistant type (AI, human) and source type (mainstream, alternative, AI).
- The survey consisted of five sections that participants completed in random-order:

Trust Task
- Participants read a series of vignettes that described a situation in which they had to depend on an AI/another person. Participants used Likert-type scales to rate the extent to which they would trust the AI/other person in each vignette.

Personality Assessment
- Participants completed self-report measures in random order including attitudes towards AI (GAAIS), the extent to which they attribute theory of mind to AI, perceptions of AI objectivity, dependence on AI, conspiratorial attitudes towards AI, how differently participants view their conscious experience relative to others' (ISO), and the extent to which participants doubt their perceptions of reality (ESO).

Gist Perception Task
- Participants viewed an animation in which several shapes rapidly moved across the screen before estimating how many of a certain type they saw. After reporting their answer, they were provided an alternative score (+5 or -5) from an AI/another participant and an opportunity to revise their initial estimate.

Remote Association Task (RAT)
- Participants are shown three words and must identify a fourth word that is related to all three of the words. During the task, participants can ask for help from an AI/another person.

Fact-Checking Task
- Participants rated the credibility of five facts. Participants were randomly assigned to learn that the facts had been reported by a mainstream media source, alternative media source, or an AI. Credibility was assessed with a 5-point Likert-type scale (definitely false = 1, definitely true = 5).

- All analyses were completed using IBM SPSS Statistics version 29.0.1.0.
  - A two-tailed independent samples t-test assessed the main effect of assistant type (AI, human) on participants' scores on the trust (Figure 1), gist perception (Figure 3), and remote association tasks (Figure 4).
  - Pearson's product-moment correlation analyzed the association between participants' scores on measures within the personality assessment (Figure 2).
  - A univariate analysis of variance (ANOVA) assessed the main effect of source type (mainstream, alternative, AI) on credibility (Figure 5).

Results

- The preliminary correlations between personality measures and main effects for source and assistant type are reported below.
- The main effects for the *trust task* are highlighted in a separate section (Figure 1).

Personality Assessment (Figure 2)
- People who had more positive attitudes about AI were more likely to depend on AI, $r(42) = .49$, $p < .001$.
- People who suspected AI of conspiracy more had less positive attitudes towards AI, $r(42) = -.38$, $p = .011$
- There was a positive correlation between ESO and ISO, $r(42) = .45$, $p = .002$

Gist Perception Task (Figure 3)
- There was no significant effect for assistant type, $t(42) = -.632$, $p = .265$, despite people adjusting their estimates more to human ($M = 7.36$, $SD = 2.72$) than to AI scores ($M = 6.64$, $SD = 4.67$).

Remote Association Task (Figure 4)
- There was no significant effect for assistant type, $t(42) = .305$, $p = .381$, despite people spending more time completing the task with AI ($M = 30.15$, $SD = 23.12$) than with human assistants ($M = 28.06$, $SD = 22.44$).

Fact Checking Task (Figure 5)
- A One-way ANOVA was conducted with information source (Mainstream, Alternative, AI) as the between-subjects factor. There was insufficient evidence to conclude that the information source had an effect on people's perceptions of credibility, $F(2, 41) = .831$, $p = .443$.

> Participants trusted human assistants significantly more than AI assistants for financial decisions, medical advice, social skills feedback, and therapy (Figure 1).
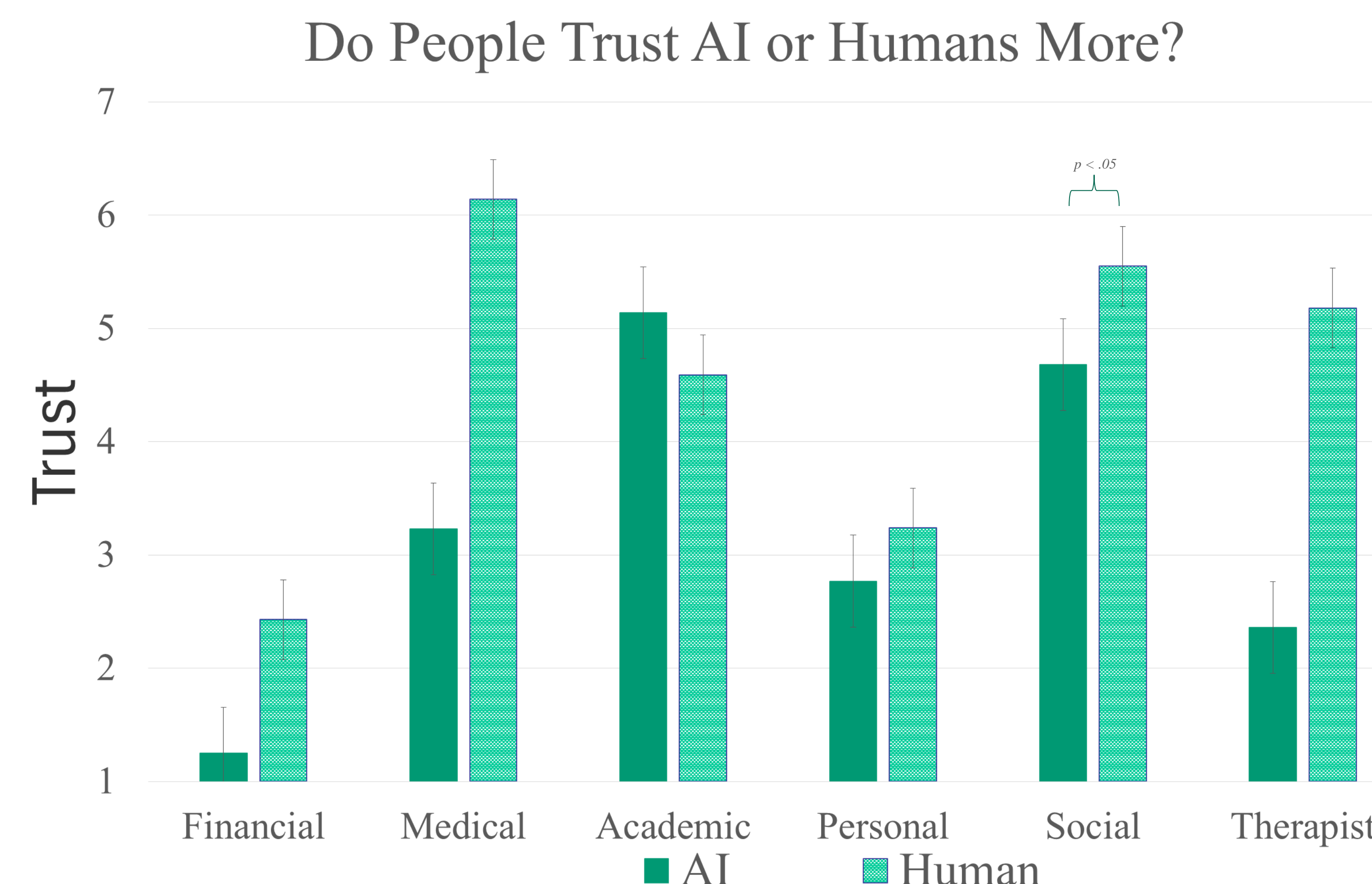
Figure 2

Personality Assessment Correlations (r)	M	SD	1	2	3	4	5	6
1. ESO	3.56	1.36						
2. ISO	5.26	1.10	**.45****					
2. Conspiracism	4.24	1.47	.27	-.11				
3. Objectivity	2.75	0.71	.11	-.03	.23			
4. TOM	3.31	0.92	.06	.03	.15	.18		
5. AID	1.89	1.07	.29	.08	-.07	.12	.25	
6. Positive GAAIS	2.90	0.64	.25	-.04	**-.38***	.25	.12	**.49****

ESO: Extrasubjective Opacity (the extent to which one doubts their perceptions of reality)
ISO: Intersubjective Opacity (how differently participants view their conscious experience compared to others)
TOM: Theory of Mind
AID: Artificial Intelligence Dependency
Positive GAAIS: A validated measure of positive attitudes towards AI (Schepman & Rodway, 2023)
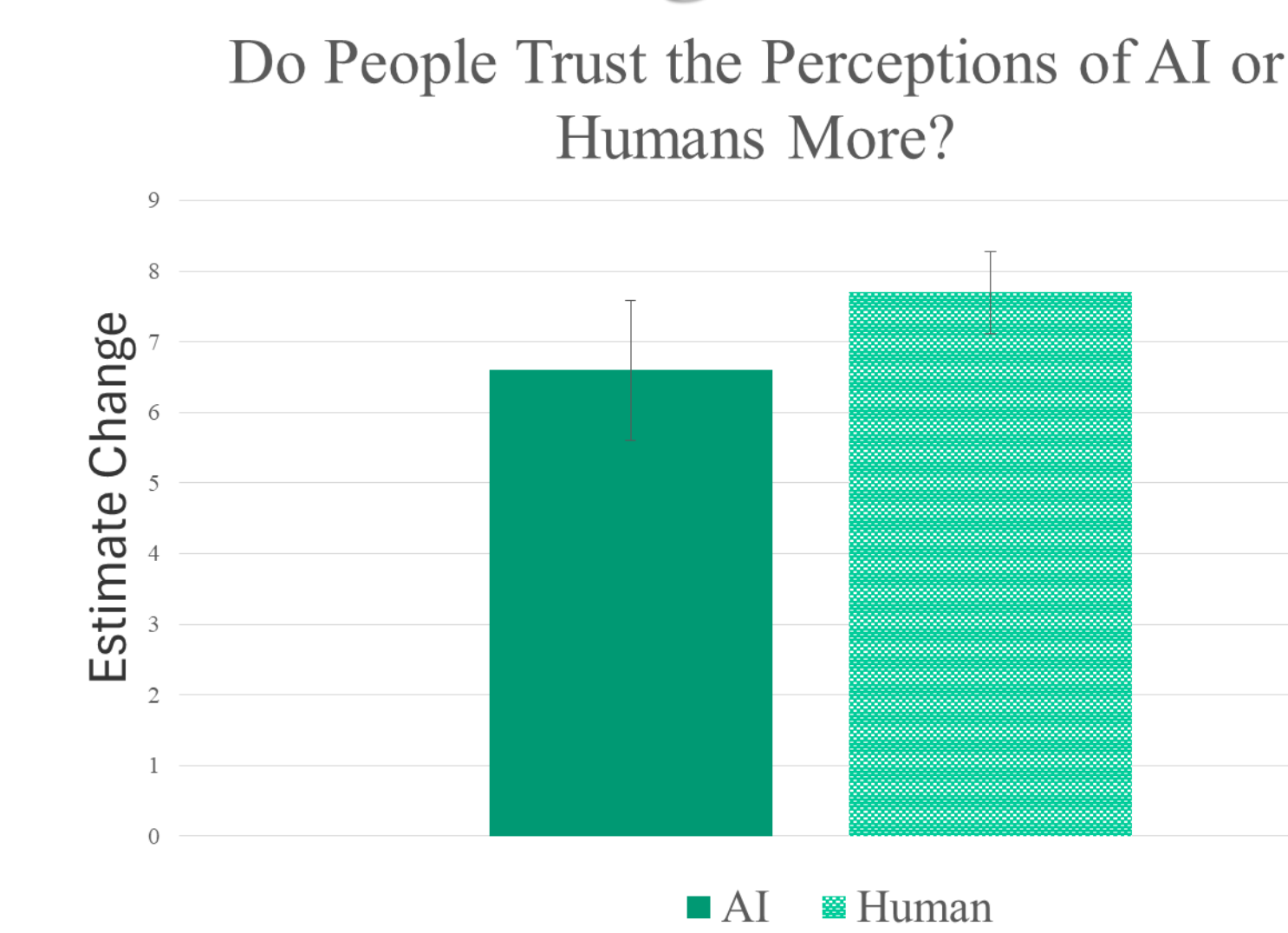
Figure 3

Do People Trust the Perceptions of AI or Humans More?

Figure 4

How Long Did it Take For Participants to Ask for Assistance?

Figure 5

Which Sources Do People Perceive as More Credible?

Figure 1.

Do People Trust AI or Humans More?
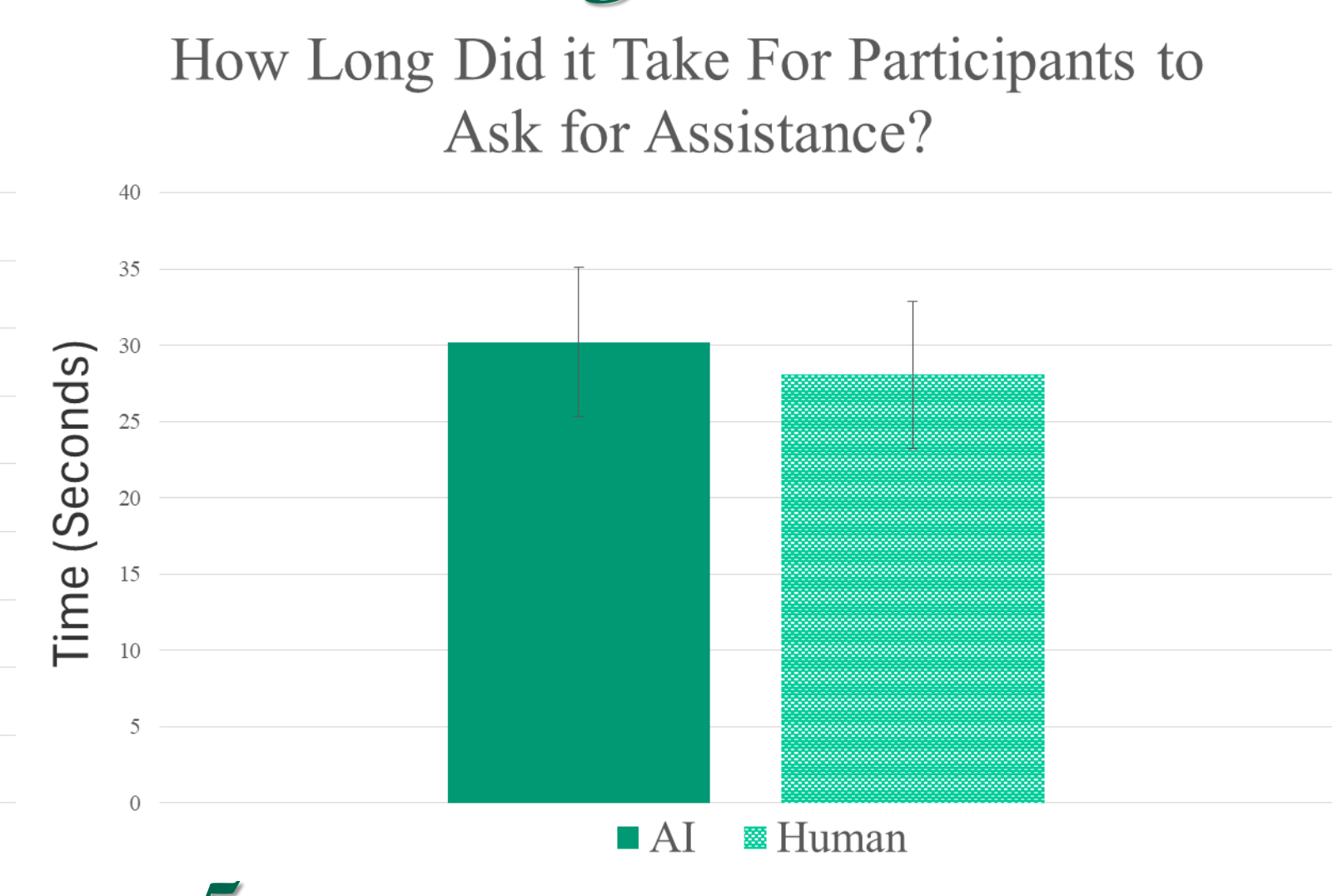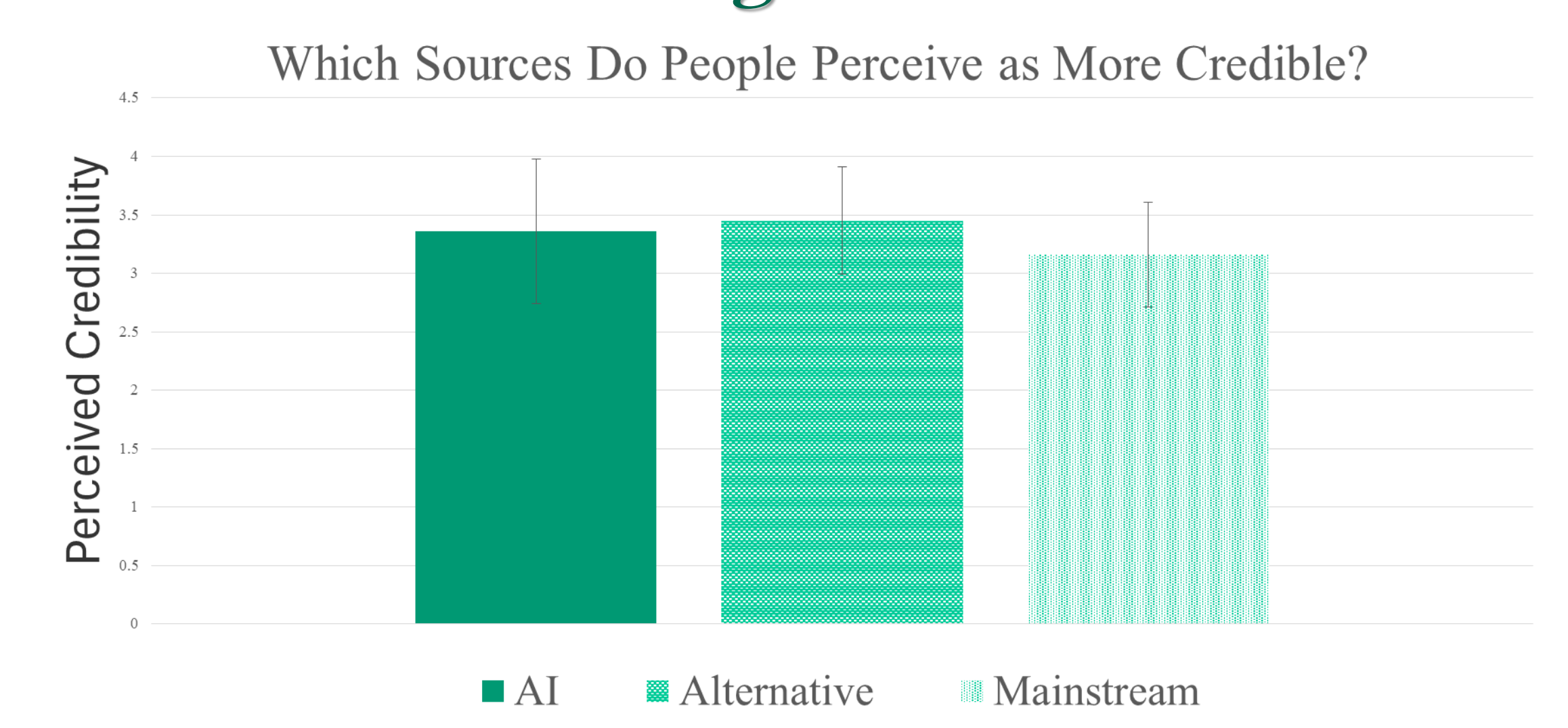
Discussion & Future Directions

- *Conclusions are limited*
  - For important and self-relevant decisions, participants generally trusted humans more than AI
  - Individual differences in users are related to their attitudes towards AI.
- *Data collection is still ongoing*
  - With a larger sample size and greater statistical power, we can examine the interaction between personality variables and assistant / source type.
  - The exploratory nature of this study reveals effects that can inform experimental studies examining the relationships between individual differences and trust in AI.

References

Crawford, J., Allen, K.-A., Pani, B., & Cowling, M. (2024). When artificial intelligence substitutes humans in higher education: the cost of loneliness, student success, and retention. *Studies in Higher Education, 49*(5), 883–897. https://doi.org/10.1080/03075079.2024.2326956

Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Empirical Research. *The Academy of Management Annals, 14*(2), 627–660. https://doi.org/10.5465/annals.2018.0057

Schepman, A., & Rodway, P. (2023). The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory Validation and Associations with Personality, Corporate Distrust, and General Trust. *International Journal of Human-Computer Interaction, 39*(13), 2724–2741. https://doi.org/10.1080/10447318.2022.2085400