

Understanding the Impact of Social Heuristics on Trust in Human-AI Teams' Decision-making

Huixin (Elsa) Zhong
Center for Doctoral Training in Accountable, Responsible, and Transparent AI,
Department of Computer Science,
University of Bath
E-mail: hz877@bath.ac.uk

Problem Statement

Decision-making in human-AI teams has achieved remarkable achievement in many fields. By utilizing the strength of both humans and AI, human-AI teams enable better than either party's performance (Bansal et al., 2019). However, the trust issue in AI has become a serious barrier that hinders the effectiveness of such collaborations. Trust issues can be categorized into two types: Distrust and over trust. Distrust in AI can lead to disuse, but over-trust in AI can cause misuse (Jacovi et al., 2021).

Research Gap

Current research on trust in human-AI teams focuses on either AI or human-related factors. However, human-AI teaming creates a unique social environment where the interaction between humans and AI may also impact trust (see figure 1).

When people interact with each other in human teams, individuals often rely upon a set of social heuristics (also known as intuition) to decide whom to trust. However, it is unknown how social heuristics impact trust in human-AI teams.

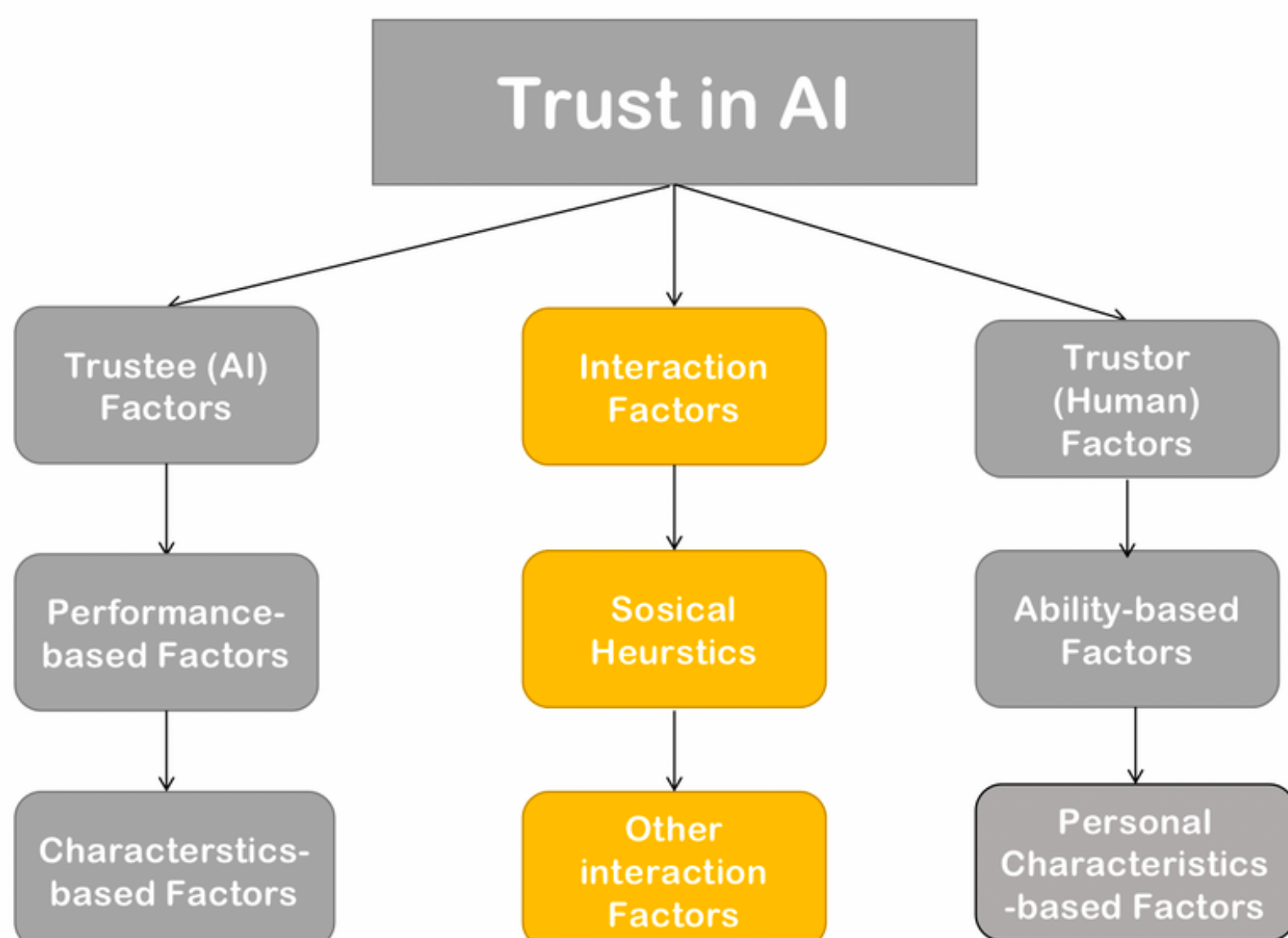
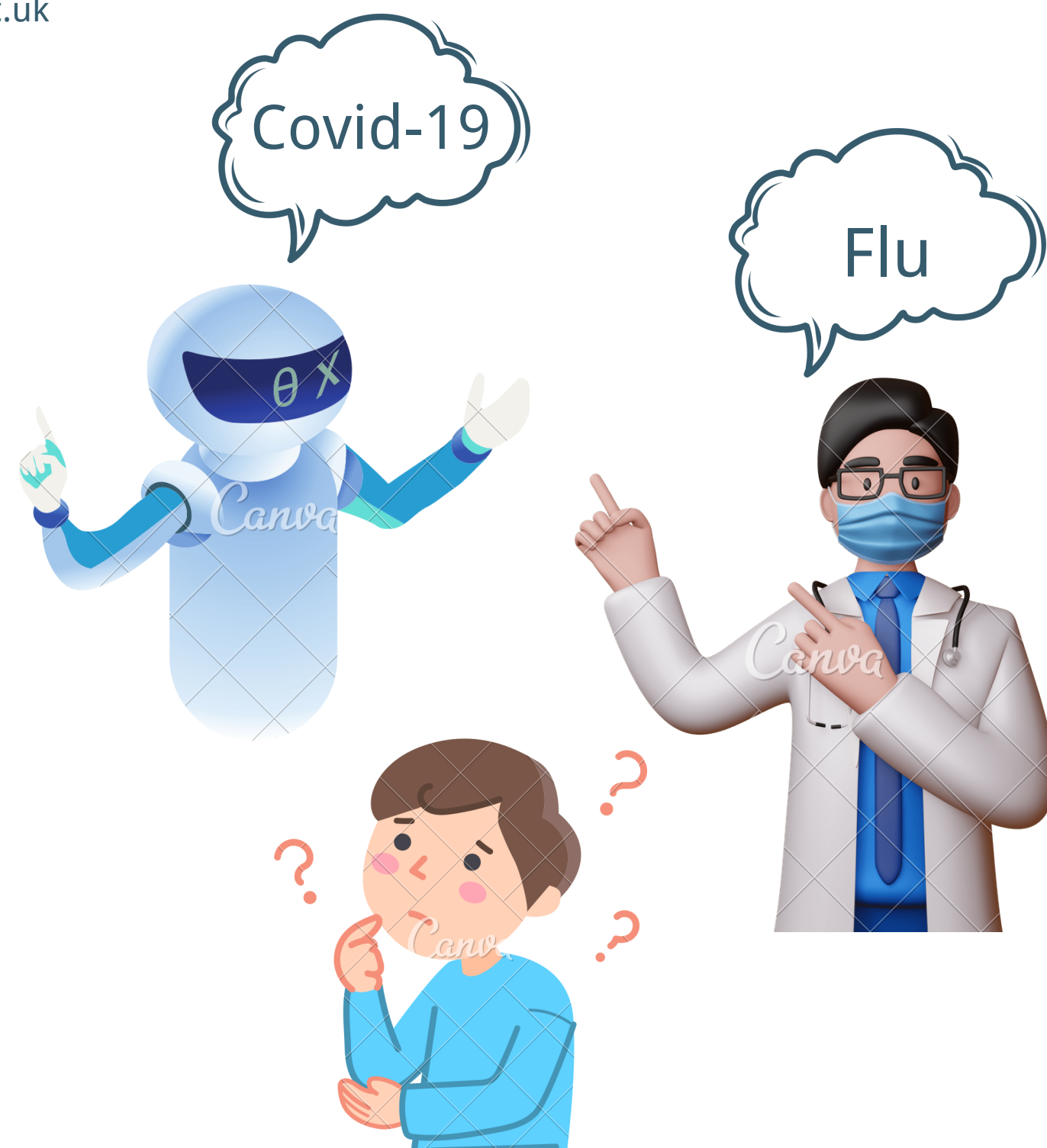


Fig 1. An overview of factors that influence trust in human-AI teaming.



Proposed Research Overview

The present Ph.D. project integrates theories and methodologies from psychology, behavioral economics, and Artificial Intelligence. We first selected two typical social heuristics: social conformity and authority influence to understand whether individuals also rely upon the two social heuristics to make decisions in human-AI teams. In the next step, we plan to integrate the two social heuristics in a theoretical trust model and design experiments to test the validation of our model.

Application

The overall results of the current research are expected to provide a set of novel trust calibration methods by utilizing human social heuristics (intuition). Besides, our results can provide references for policy-making to prevent algorithm manipulation and guide human-AI interaction for designers.