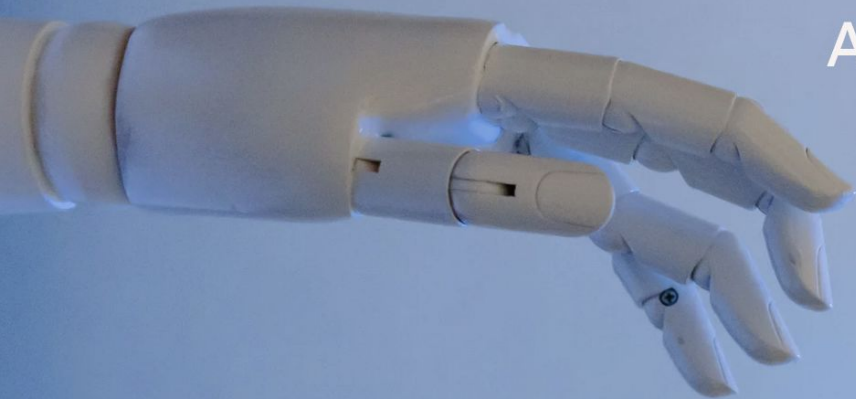# FOSTERING INTELLECTUAL HUMILITY THROUGH AI : AN AGENTIC APPROACH

TEAM 4

Members:
Dhanush Bhargav
Guruksha Gurnani

Guided by :
Prof. Ryan Watkins
Prof. Philip Moore
Prof. Amir Jafari

PROBLEM STATEMENT

01 **WHAT IS INTELLECTUAL HUMILITY ?**

Intellectual humility (IH) is the ability to recognize that one's knowledge may be incomplete and that personal beliefs could be wrong, encouraging openness to new evidence and perspectives.

02 **WHY IS INTELLECTUAL HUMILITY IMPORTANT?**

In a world increasingly polarized by rigid opinions, developing intellectual humility is crucial for promoting understanding, empathy, and meaningful dialogue.

03 **CHALLENGES:**

Traditional methods like debates or presenting evidence often fail to change entrenched views, driving the need for new techniques that can effectively nurture intellectual humility

PROJECT GOALS

01 **DESIGN INTERVENTION BY PROF. MOORE**
Designing the Intervention Strategy

02 **DEVELOP MULTI AGENT APPLICATION**
Creating the application for intervention

03 **TEST ON HUMAN SUBJECTS**
Conducting Tests on Participants

04 **REPORT FINDINGS OF PILOT STUDY**
Summarizing key outcomes and insights from the pilot study

TECHNIQUES

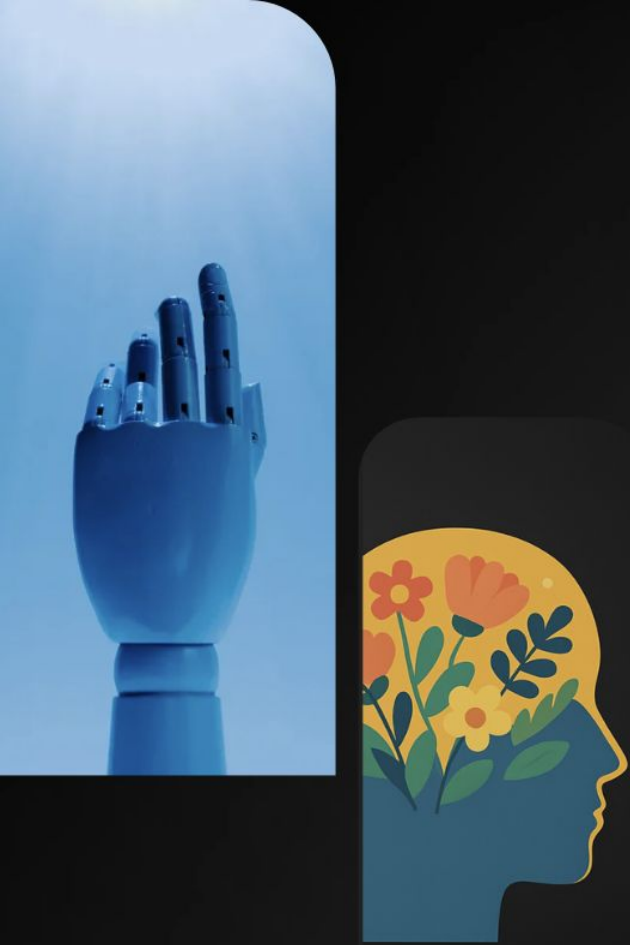| CHOOSE STANCE	PROVIDE ARGUMENTS	CATEGORIZE ARGUMENTS	MATCH ARGUMENTS	RATE IMPLICATION
SELECT YES/NO STANCE AND STRENGTH RATING FROM 1 -10	USER PROVIDES ARGUMENTS FOR AND AGAINST THE STANCE	DEFINE CATEGORIES FOR ARGUMENTS	MATCH ARGUMENTS T O CATEGORIES	RATE IMPLICATION OF EACH ARGUMENT WITHIN ITS CATEGORY

INTELLECTUAL HUMILITY (IH) QUESTIONNAIRE

01 **ASSESSMENT STRUCTURE:**

The questionnaire consists of 8 statements designed to evaluate different dimensions of intellectual humility

02 **RESPONSE SCALE:**

Participants rate each statement on a 5-point scale, ranging from 1 (Strongly Disagree) to 5 (Strongly Agree)

03 **SCORING METHOD:**

The final IH score is obtained by averaging the participant's responses across all 8 statements, resulting in a score between 1 and 5

SOCIAL DESIRABILITY QUESTIONNAIRE

01 **ASSESSMENT STRUCTURE:**

The questionnaire includes 11 statements that participants respond to with either "Yes" (scored as 1) or "No" (scored as 0)

02 **QUESTION TYPES:**

Some statements are moderately critical (e.g., "I like to gossip at times"), while others are extremely positive (e.g., "I have never intensely disliked someone").

03 **SCORING METHOD:**

Critical questions are reverse-coded, and the total social desirability score is calculated by summing the responses, resulting in a score between 1 and 11

RESULTS: CHANGE IN INTELLECTUAL HUMILITY

01 INCREASE IN INTELLECTUAL HUMILITY:

Change in IH (ΔIH) is computed by comparing participants' pre-intervention and post-intervention IH scores.

02 STATISTICAL TESTING:

A two-tailed t-test is conducted on ΔIH to assess whether the observed change is statistically significant.

03 SIGNIFICANCE OF RESULTS:

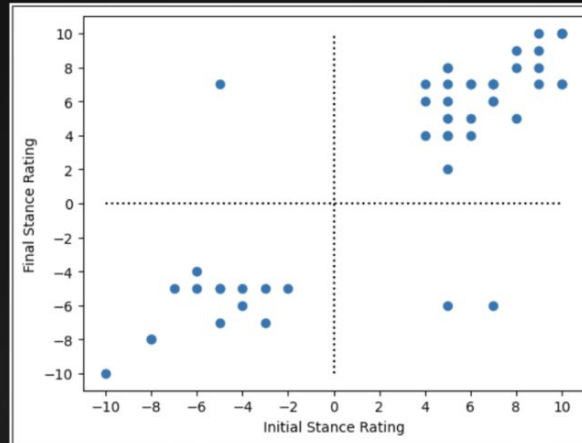The p-value indicates that the change is statistically significant, with a non-zero mean ($\mu\Delta$IH)

04 INTERPRETATION OF IMPROVEMENT:

A positive $\mu\Delta$IH suggests that participants' intellectual humility scores improved following the intervention.

| statistic	value
Mean ΔIH	0.191
T-statistic	4.557
p value	<.0001

RESULTS: STANCE CHANGES

STANCE CHANGES FOR THE QUESTION "SHOULD UNIVERSAL BASIC INCOME (UBI) BE ADOPTED?"
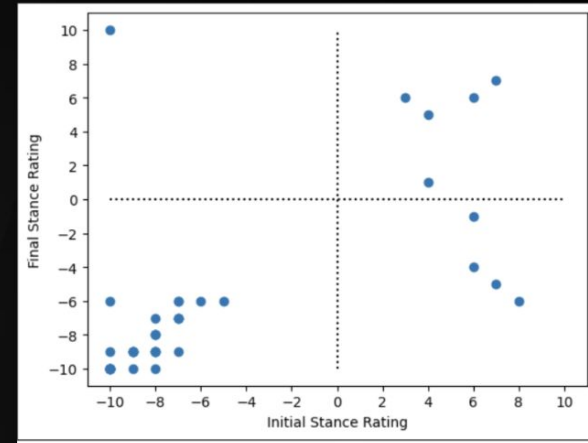
STANCE CHANGES FOR THE QUESTION "SHOULD TEACHERS BE ALLOWED TO CARRY FIREARMS IN SCHOOLS?"

RESULTS: STANCE CHANGE

01 STANCE RATINGS ASSOCIATED WITH A "NO" RESPONSE ARE ASSIGNED A NEGATIVE VALUE

02 STANCE RATINGS ASSOCIATED WITH A "NO" RESPONSE ARE ASSIGNED A NEGATIVE VALUE

03 DOTTED LINES ARE USED TO DIVIDE THE GRAPH INTO FOUR QUADRANTS FOR EASIER INTERPRETATION.

04 THE LEFT HALF OF THE GRAPH REPRESENTS INITIAL "NO" STANCES, WHILE THE RIGHT HALF REPRESENTS INITIAL "YES" STANCES

05 THE TOP HALF OF THE GRAPH CORRESPONDS TO FINAL "YES" STANCES, AND THE BOTTOM HALF CORRESPONDS TO FINAL "NO" STANCES
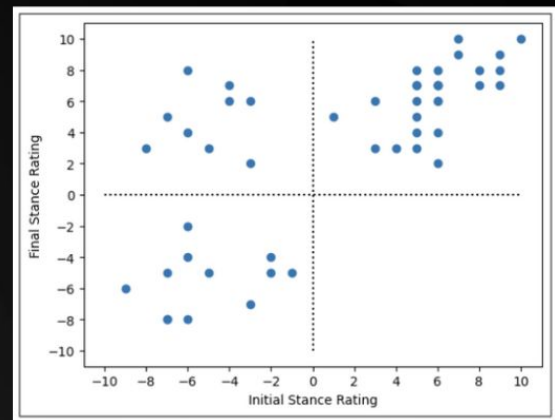
STANCE CHANGE FOR THE QUESTION "SHOULD NUCLEAR POWER BE CONSIDERED A KEY PART OF GREEN ENERGY SOLUTIONS?"

RESULTS: IMPACT OF SOCIAL DESIRABILITY

01 **TESTING METHOD:**

Two-sample t-tests were conducted to compare mean social desirability scores between participants with "yes" and "no" stances.
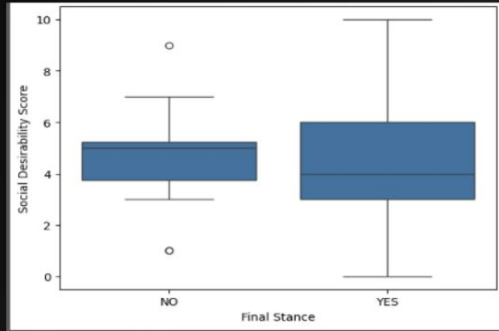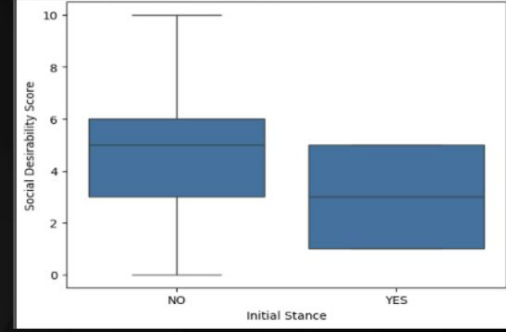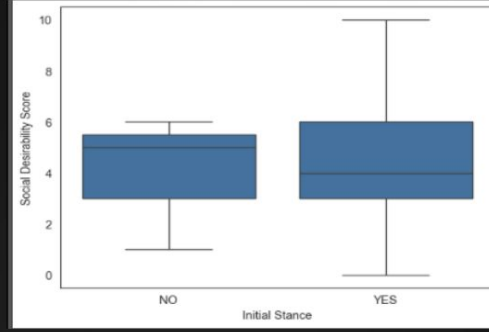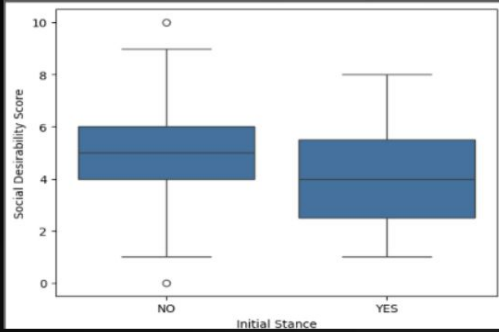
02 **FINDINGS:**

The t-tests showed no significant differences in mean scores across all three stance-related questions.

03 **CONCLUSION:**

There was no significant relationship between participants' initial or final stances and their social desirability scores.

RESULTS - IMPACT OF SOCIAL DESIRABILITY
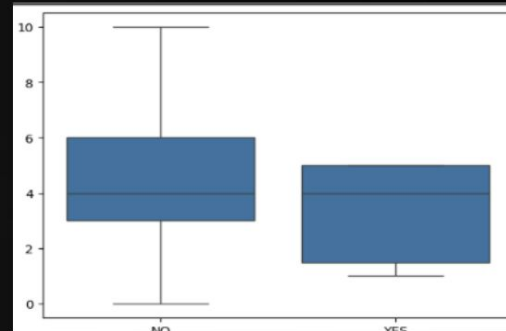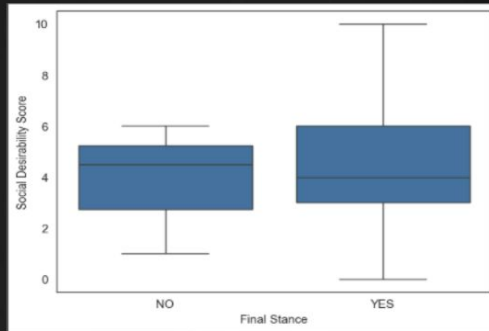
CONCLUSIONS

01 INCREASE IN INTELLECTUAL HUMILITY:

The intervention led to a statistically significant increase in IH scores among participants

02 POTENTIAL FOR FURTHER RESEARCH:

Results indicate that the intervention is promising and deserves deeper exploration in future studies.

03 STANCE CHANGES:

No clear pattern was observed in stance changes — although influencing stance was not the primary goal of the intervention.

04 STANCE CHANGES:

There was no correlation found between participants' stances and their social desirability scores.

05 INTERPRETATION :

The observed improvement in IH appears to stem from the intervention itself, rather than from participants' desire to present themselves positively

THANK YOU FOR YOUR TIME

ANY QUESTIONS?

Results - change in IH

- Change in IH (Δ_{IH}) is calculated from pre and post intervention IH scores.
- Two tailed t-test is performed on the Δ_{IH} to check for significance.
- P-value suggests significance and $\mu_{\Delta IH}$ is non-zero.
- Positive $\mu_{\Delta IH}$ suggests improvement in IH scores after intervention.

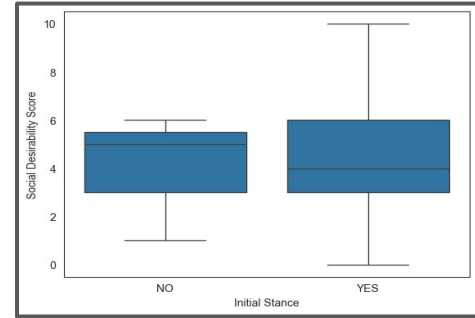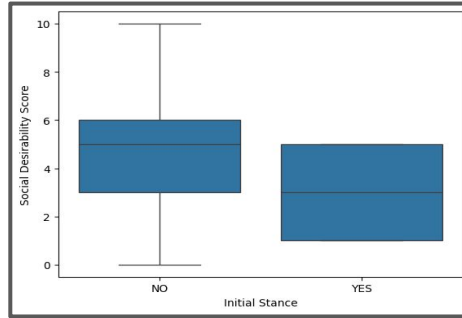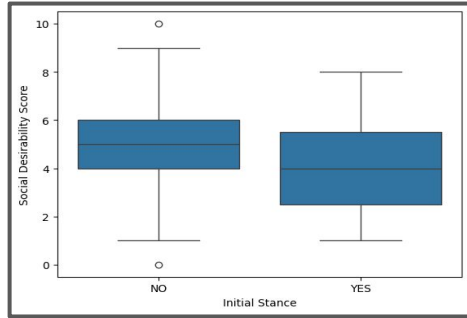| Statistic	Value
Mean Δ_{IH}	0.191
T-statistic	4.557
p-value	< .0001

Results - Stance Change

- Stance ratings for "no" stance are given negative valence.
- Position scale becomes -10 to 10.
- Dotted lines divide graph into 4 quadrants.
- Left and right halves represent initial "no" and "yes" stances respectively.
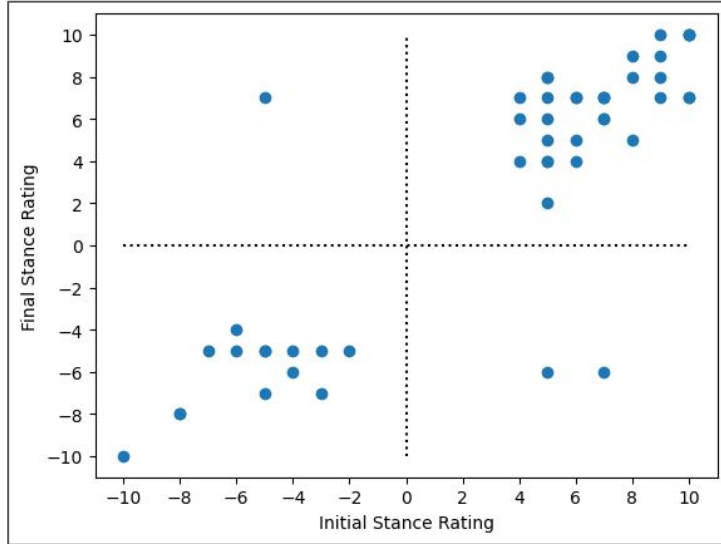- Top and bottom halves represent final "yes" and "no" stances respectively.
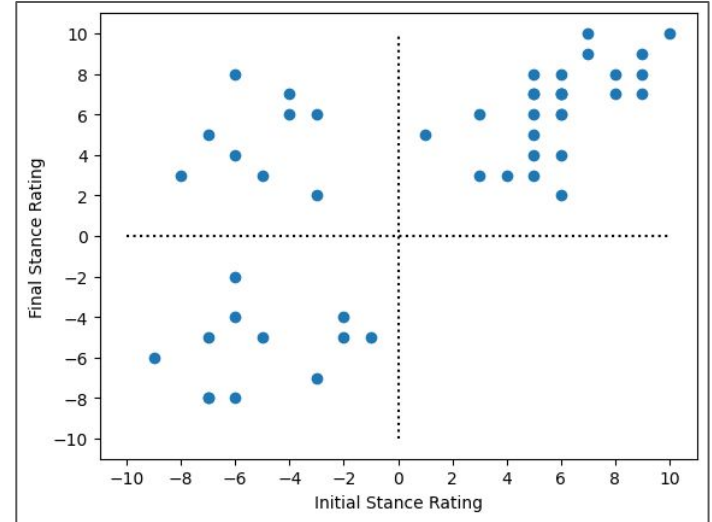
Stance change for the question "Should nuclear power be considered a key part of green energy solutions?"
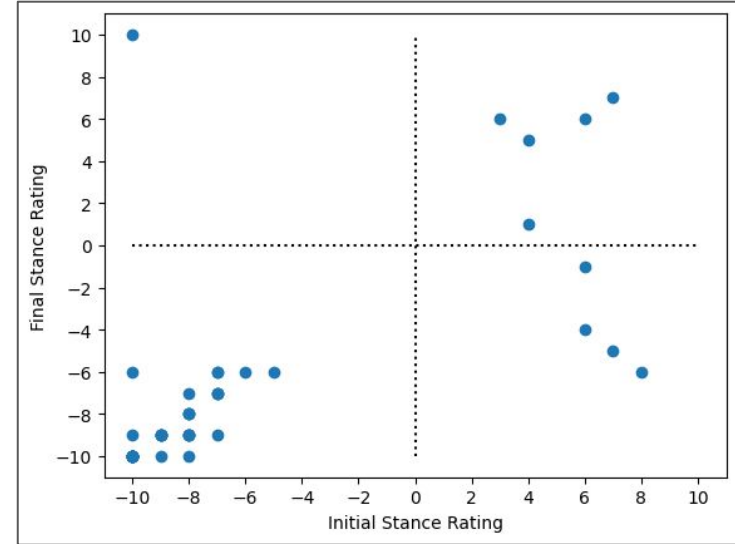
Results - impact of social desirability

Results - Stance Changes

Stance changes for the question "**Should universal basic income (UBI) be adopted?**"

Stance changes for the question **"Should teachers be allowed to carry firearms in schools?"**