Data Science & Machine Learning for Engineering Applications

Prof. Daniele Quercia

CLIM8s

Authors:

Khairy, Sherif	343780	Hamdan, Dana	335265
Shedid, Mohamed	343782	El Sahmarani, Mohannad	343383
Mehri, Ali	342414	Mansour, Anthony	339397
Petrarulo, Antonio	332907	Bayat, Amir	336941

Date: 09/06/2025

Beyond the Thermometer: Deconstructing Climate Change Perceptions and Pro-Environmental Behavior

A Survey Study

Hamdan, Dana
Georesources and Geoenergy
Engineering
Politecnico di Torino
s335265@studenti.polito.it

Mehri, Ali
Georesources and Geoenergy
Engineering
Politecnico di Torino
s342414@studenti.polito.it

Khairy, Sherif
Georesources and Geoenergy
Engineering
Politecnico di Torino
s343780@Studenti.polito.it

Shedid, Mohamed
Georesources and Geoenergy
Engineering
Politecnico di Torino
s343782@studenti.polito.it

Mansour, Anthony
Georesources and Geoenergy
Engineering
Politecnico di Torino
s339397@studenti.polito.it

Petrarulo, Antonio
Chemical and Sustainable
Processes Engineering
Politecnico di Torino
s332907@studenti.polito.it

El Sahmarani, Mohannad
Georesources and Geoenergy
Engineering
Politecnico di Torino
s343383@studenti.polito.it

Bayat, Amir
Georesources and Geoenergy
Engineering
Politecnico di Torino
s336941@studenti.polito.it

ABSTRACT

This study addresses the crucial issue of understanding how the public responds to climate change, specifically its multi-faceted nature that is influenced by psychological, social, and demographic factors. By conducting a representative survey of 178 participants, we examined how assessed constructs, specifically, the Big Five personality factors, researchers' previous discoveries with climate change conspiracy theories, and demographic influences affect individuals' perceptions of climate and their willingness to act on it. After analyzing responses utilizing multiple different analysis types, correlation matrices to consider relationships across the responses overall, stratification so we could determine group-level differences, and predictive modeling via logistic regression and advanced classification, we assessed variables which predicted attitudes toward climate, and we also used clustering approaches to identify sub-groups according to those characteristics. In summary, we provided some details to better shape the terrain of climate beliefs and behavior, providing substantial evidence to tailor climate communication more effectively.

DATA COLLECTION

For the purposes of this project, survey data were collected from a total of 178 respondents. The strategy we used was to obtain participants through personal and social networks and to persuade them to take part in the survey - we used an online survey to make it convenient for participants, as the full questionnaire took some time to complete. The surveys we administered included Climate Survey, Big Five Personality Test, and an Additional Survey. We also included brief, customized questions, as a means of gathering further unique findings. Since we also had access to people from Italy and outside Italy, we then needed to make different versions (Italian, Arabic, Farsi, French) of the survey, in order for the people we reached to complete the survey.

To boost participation, we offered a digital prize lottery. At the end of the survey, each respondent received a unique code that they could use to check if they won. The winner was randomly selected after the data collection phase, and the announcement and code verification processes were hosted through our Instagram page. To reward participants beyond just digital incentives, we also promised real-world impact: for every 100 respondents, we committed to planting one tree as part of our contribution to climate action.

Data was gathered from participants systematically regarding several personal characteristics. The data collected for each respondent encompassed their gender, country of origin, and residence. Other characteristics collected were their job status, job role, job level, significant topic(s) of primary interest, and social media habits.

For the consent form, each respondent was prompted before filling the survey. It was stated that by completing the survey, they consent to the use of their data anonymously for the purpose of the study. The data we collected was compiled into a single CSV file with each completed survey responses organized in a single row, distinguished by their anonymous identifier only.

Our goal throughout was not only to gather data, but to inspire awareness and engagement, turning participants into part of the solution.

RESEARCH QUESTIONS

The current research explored a variety of aspects that could affect perceptions and engagement of climate change. Utilizing the data collected in their survey, our study attempted to address the following questions:

1. **Is there a relationship between conspiracy mindset scores and the tendency to deny climate change?** This question asked whether greater scores on different measures of conspiracy belief are associated with greater likelihood of rejecting or disbelieving the scientific consensus on climate change.

2. **Do people who claim to care about climate change actually engage in sustainable behaviors?** This question examines the possible dissonance or congruence between self-reported concern with climate change and participation in sustainable environmental practices or behaviors.

3. **Are personality traits (e.g., openness) associated with environmental awareness?** This question considers the relationship between specific Big Five personality traits, including Openness to Experience, and an individual's degree of knowledge or awareness of environmental issues and climate change.

4. **Which interests are more associated with a higher level of environmental awareness?** This question explores the link between an individual's topics of interest and their level of exposure to climate change news, as well as their level of climate change awareness.

METHODS

Data cleanup

The downloaded responses file had some errors. A few responses were not stored in their respective questions. For example, a response to a climate perception question (e.g. "strongly agree"), was stored as a response to the Gender question. This cleanup process was conducted manually on Excel for maximum accuracy.

Using the Filter method on Excel, we could efficiently filter column by column. By clicking on the drop-down menu, all the unique responses could be seen, and it would be easy and fast to find the incorrectly stored responses. To ensure accuracy of results, these erroneous rows were removed from our cleaned dataset. In total, 74 rows of misaligned responses were removed.

Data Preparation

Next, in our data preparation process, we sought to reduce the number of columns based on their usefulness in analysis. This was conducted through a filtration on remaining missing values, as well as their respective variance.

The additional questions that were asked based on each responder's interest had a very limited dataset size, and so, they were excluded from the analysis.

The question "Do you believe in Climate Change?" had extremely limited variation. Around 90% of responders answered "Yes", with the remaining 10% distributed evenly across the two remaining answers: "No", and "I Don't Know". Such questions would also be excluded from the analysis.

Our next step in Data Preparation is classification of questions according to Predictors and Outcomes. It was also noted that some questions could be used as both, such as the responder's perception of climate change as well as their perception of scientists' view on climate change. Those can be used as outcomes with the predictors being age, gender, interest, or any other predictor. On the other hand, they can also be used as predictors themselves, with the outcome being a question more related to action towards climate change.

Next, we classified the Predictors themselves. We had General Predictors, Big Five Personality Questions, and Conspiracy Questions. Each of those would be treated differently, as their numerical transformation and scoring methods are not the same. That will be discussed further in the Data Analysis part of the Methods section.

The Climate Perception questions were also classified by semantically similar questions. This will come in handy when we apply a correlation across all the perception questions in order to reduce their number. Another filtration process will be the variance as mentioned above. Using both the correlation and the variance filtration, we can reach a smaller number of perception questions.

Data Analysis

Data Processing and Transformation

The first phase consisted of cleaning and transforming the raw survey data into a quantitative analysis-ready format.

Response Encoding and Likert Scale Standardization: All of the categorical and textual responses were translated into numerical values using replacer dictionary. This involved converting Likert scale responses (e.g., 5-point Likert Scale from 'strongly disagree' to 'strongly agree') to numbers from 1-5, and any yes/no responses turned into binary 1/0s.

Dummy Variables and Multi-Hot Encoding: The categorical variables Gender, Job Status, Job role, and Region of Residence were converted to numerical dummy variables. Based on the topics of interest, a multi-hot encoding process was used to create a binary column for each topic.

Handling Missing Values and Data Type Conversion: Missing values were treated in the context of the analysis to mean replacing those responses with zeros, and all columns would be explicitly converted to numerical data types.

Feature Engineering and Scoring

Conspiracy Mindset Score Calculation: A custom scoring method was used for conspiracy questions based on the standardized test that was utilized. Each response was ranked 0-4 according to a custom Likert Scale. The result of the questions were summed for each category (aliens, concealed experiments, global conspiracy, government coverup, and information control), then divided by the maximum sum (12) and converted to a percentage to obtain the category score. Then the category deviation was calculated as the difference between the category score and the population average for each category as reported by the standardized test. Then an overall deviation score was calculated by averaging the five deviations.

Big Five Personality Trait Scoring: Big Five Personality Test responses were mapped to a numerical score (1-5). The 10 questions were grouped according to five traits, with each trait having a question which is normally scaled, and another one with a reverse scaling. To obtain each trait score, the results of its two respective questions would be averaged. The main five traits are: Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness.

Exploratory Data Analysis (EDA) and Feature Selection

Z-score Standardization: This process scaled numerical and Likert scale questions to a common scale by standardizing scores to mean of 0 and standard deviation of 1 (Gelman & Hill, 2007).

Correlation Matrix Analysis: Correlation matrices for all questions, and some particular subsets (predictors, perceptions, actions) were generated. Heatmaps displayed the Pearson correlation coefficients to visualize relationships and detect multicollinearity.

Dropping Highly Correlated Features: Columns that showed absolute correlation above a certain threshold (typically 0.6 or 0.85) were flagged and dropped. A custom function on Python was created to make this process more efficient.

Variance Thresholding: In addition, Columns with low variability (below a threshold of 0.75) within the perception questions were identified and dropped to improve predictive power.

Predictive Modeling

Linear Regression: The intended purpose for the statistical analysis was to answer Research Questions 1, 2, 3, and 4 to assess linear relationships between predictor variables (conspiracy scores, perception questions, personality traits, interests) and continuous outcomes (e.g. harm perception, willingness to pay) (Gelman & Hill, 2007). The model performance was evaluated using R-squared (R^2) and Mean Squared Error (MSE).

Logistic Regression: For the decision tree model, the analysis was for binary outcomes (e.g. climate belief), which was also dependent on the scores for conspiracies. Model performance was looked at in terms of accuracy, confusion matrices, and classification reports (precision, recall, and F1-score) (Gelman & Hill, 2007).

General Statistics and Visualization: Data was summarized using sums and means. The plots were created using bar plots (e.g. beta coefficients and average interest scores) and heatmaps (e.g. confusion matrices). This was mainly used to analyze the effect of the respondents' Topics of Interest as will be further described in the Results section.

RESULTS

Correlation Matrix of Survey Variables

To explore potential relationships between different variables in our dataset, we generated a full correlation matrix covering all survey questions. This matrix visually highlights how variables such as personal beliefs, demographic traits, media exposure, personality factors, and climate-related attitudes relate to one another.

In the matrix:

- Red and dark orange areas **represent** stronger positive correlations**, where two responses tend to increase together.**
- Blue areas **show** weaker correlations**,** indicating that no consistent pattern exists.
- Diagonal values (perfect correlation with self) are shown in **dark red (r = 1.0)**.
- For the sake of clarity, only correlations with a value higher than **0.6** had written annotations. The rest were considered NaNs (shown in blue).

This matrix served as an initial diagnostic tool to:

- Identify **clusters** of strongly related items (e.g., questions within the same thematic area, such as climate concern or conspiracy beliefs).
- Guide our selection of the most relevant predictors and targets for each research question.
- Discover **unexpected or weak correlations**, which were then explored in detail in the analysis of each research question.

The matrix confirms the complexity of public perception: while some groups of questions are closely related, others show more scattered or nuanced relationships.
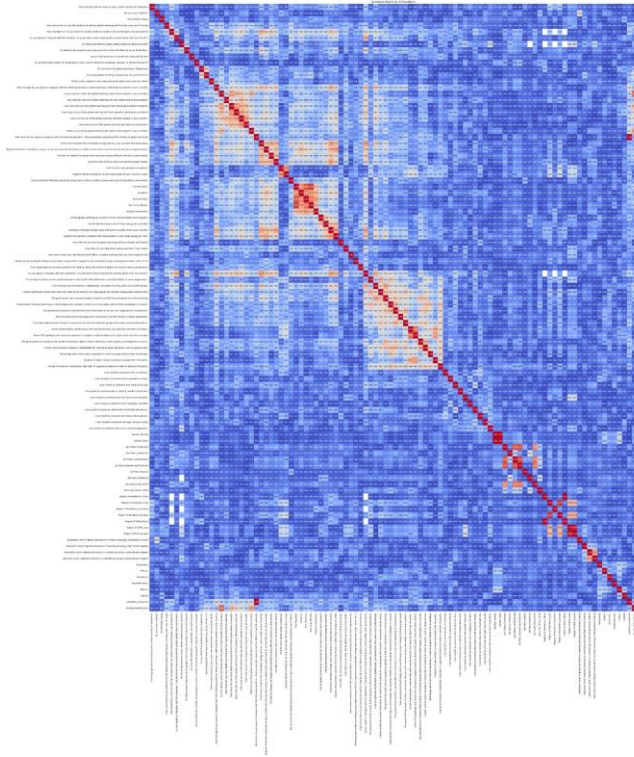
Figure 1: Correlation Matrix for all the Questions

Research Question 1: Are conspiracy perceptions related to climate perceptions?

To investigate the link between conspiracy beliefs and climate change perception, we used a standardized 15-item conspiracy belief questionnaire (Imhoff & Bruder, 2014; Tam & Chan, 2023). This included topics such as alien contact, government cover-ups, information suppression, scientific manipulation, and global conspiracies. Each response was rated on a five-point Likert scale, and we calculated both individual category scores and an overall conspiracy score for each respondent.

We selected the below target question to assess climate perception:

- "Which comes closest to your own view about what most scientists think about global warming?"

Our analysis revealed no strong correlations between overall conspiracy scores and these climate perception questions. However, several **weak to moderate correlations** provided interesting insights.

We observed a **weak positive correlation (~0.2)** between the respondent's **alien conspiracy beliefs** and their view of the general scientific perception of climate change. We observed higher correlations in other categories such as **concealed experiments**

(**~0.28**) and **government coverups (~0.30)**. While the highest correlations were found in **global conspiracies (~0.36)** and **information control (~0.35)** which are highly correlated with the **overall conspiracy score (~0.35)**. This is a surprising insight since it was expected that there would be a negative correlation, not a moderately positive one.

Results were different when for the linear regression model to predict climate change perception with conspiracy thinking. Both positive and negative beta coefficients were observed as shown below. However, the R^2 score was only **(0.149)**.

- Alien conspiracies ($\approx$ -0.0577)
- Concealed experiments ($\approx$ -0.2293)
- Government cover-ups ($\approx$ 0.0502)
- Information control ($\approx$ 0.1989)
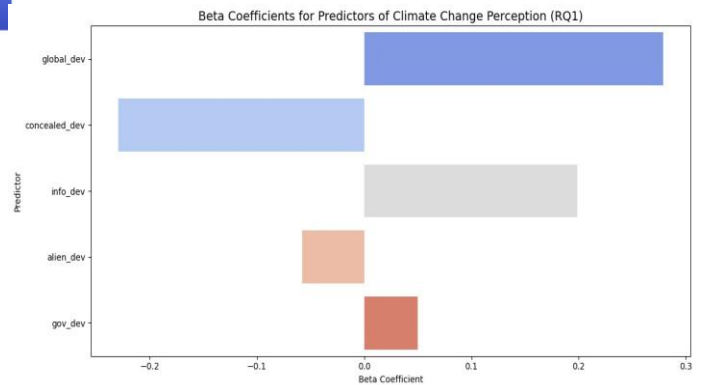- Global conspiracies ($\approx$ 0.2788)

Figure 2: Beta Coefficients for Conspiracy vs Climate Change Perception

While belief in global conspiracies and information control had a positive correlation with trust in scientific studies related to climate change, a higher belief in the existence of concealed experiments had a strong negative correlation. This is a more expected result as it means that people who believe scientists are hiding experiments, do not conform with the common scientific perception of climate change.

However, the weakness of correlation coefficients and beta coefficients suggest that belief systems may be **compartmentalized**, with some respondents able to endorse both conspiracy thinking and confidence in climate science, depending on context. This means that conspiracy cannot be accurately correlated to climate change perception. Moreover, these results may be due to inaccuracies in the responses.

Research Question 2: Do people who claim to care about climate change actually engage in sustainable behaviors?

This question explored whether belief in climate change translates into willingness to act. We analyzed correlations between climate belief indicators and actions or behavioral intent.

We selected the following target questions related to sustainable behavior:

- Q1: "How concerned are you that policies to address global warming will increase your cost of living?"
- Q2: "How often do you discuss global warming with your friends and family?"
- Q3: "Would you personally be willing to pay higher energy bills to support renewable energy development?"
- Q4: "How responsible do you feel personally for helping reduce the effects of global warming on future generations?"

While we have not found a very strong correlation between the four selected questions, we opted for Q1 for its strong representation of the respondents' willingness to endure lifestyle changes for the sake of the environment. It was also the most strongly correlated question with the others.
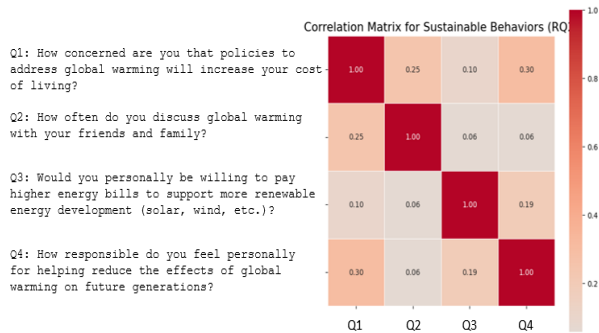

Figure 3: Correlation Matrix for the Climate Action Questions

Based on the beta coefficients, the most strongly related predictor was, "In your opinion, when will global warming start to harm people in your country?" **(~0.40)**. Interestingly, this shows that those who believe the effects are already happening or imminent are more concerned with having a higher cost of living. This suggests that their belief is a cause for concern, which may be due to a sense of fear or a sense of responsibility (van der Linden, 2015).

We also found a negative beta coefficient (~ -0.21) with those who have a more perceived severity of the damages of climate change. This suggests that they could be less concerned with higher living costs for the sake of the environment.
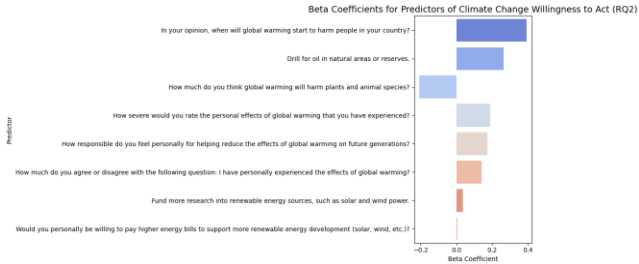

Figure 4: Beta Coefficients for the Predictors of Climate Change Willingness to Act

To further investigate this link, the same model was applied to Q3 to explore the respondent's willingness to pay for climate action. The model returned smaller beta values, with the main contribution related to:

- How much do you think global warming will harm plants and animal species? (~ 0.1313)
- How responsible do you feel personally for helping reduce the effects of global warming on future generations? (~ 0.1448)

Other predictors made minimal contribution. From this it can be observed that the more people perceive a larger damage and feel a higher sense of responsibility, the more they are willing to pay for climate action (Cologna et al., 2024).
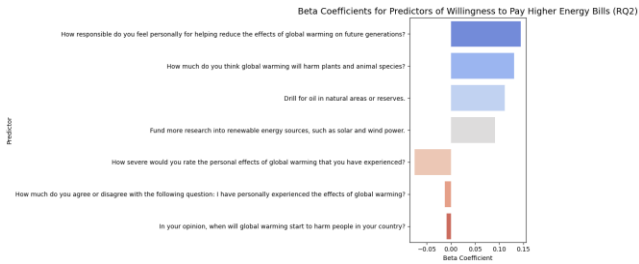

Figure 5: Beta Coefficients for the Predictors of Climate Change Willingness to Pay Higher Energy Bills

Research Question 3: Which personality traits are associated with higher environmental awareness?

To explore psychological predictors of climate concern, we analyzed the relationship between the Big Five personality traits and the perception of climate threat.

We used the question:

- "In your opinion, when will global warming start to harm people in your country?" as the dependent variable, reflecting how immediate and serious respondents believe the climate threat to be.

Each participant's personality was scored using the Big Five model (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism), calculated through averaging responses after appropriate scale reversal for specific items.

A linear regression showed:

- Openness had the strongest positive correlation ($\beta \approx$ +0.224), consistent with prior research linking openness to acceptance of scientific consensus and global thinking (van der Linden, 2015; Cologna et al., 2024).
- Extraversion ($\beta \approx -0.220$) and Neuroticism ($\beta \approx -0.126$) showed moderate negative correlations. While surprising, these may reflect disengagement or emotional avoidance in the face of complex global issues.
- Conscientiousness ($\beta \approx -0.036$) and Agreeableness ($\beta \approx$ +0.001) had negligible effects.
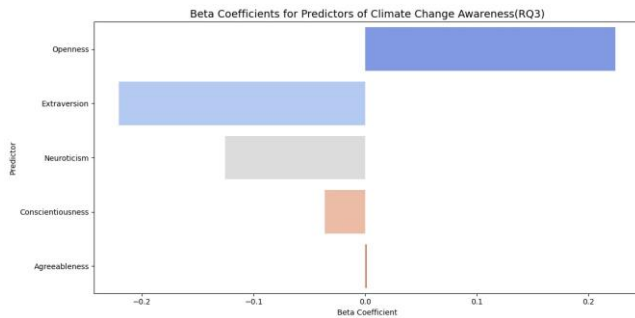

Figure 6: Beta Coefficients for Big Five Traits and Climate Perception

These findings reinforce existing psychological literature: Openness is a reliable predictor of environmental concern, while other traits play minimal or context-dependent roles.

Research Question 4: Which interests are more associated with a higher level of environmental awareness?

To investigate whether respondents' personal interests are related to their awareness of climate change, we analyzed the relationship between selected interest topics and climate perception. The target question used as a measure of environmental awareness was:

- "In your opinion, when will global warming start to harm people in your country?"

We applied a "Hot Topic Method", transforming the multi-selection interest field into individual binary columns for each interest (e.g., 1 if the respondent selected Politics, 0 otherwise). We then calculated the average awareness score for each topic by correlating it with the selected climate awareness question.
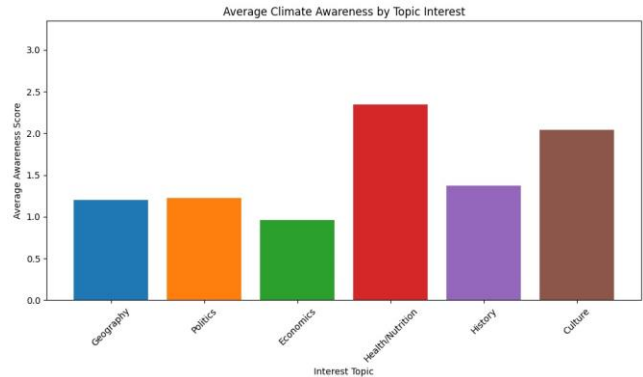

Figure 7: Average Climate Awareness by Topics of Interest

As shown in the figure, Health & Nutrition and Culture stood out with notably higher climate awareness scores compared to other interests like Geography, Politics, or Economics.

- Those interested in Health & Nutrition may already be inclined toward sustainability and lifestyle choices that align with environmental values.
- Respondents interested in Culture may be more exposed to trending social topics, including climate action, possibly increasing their awareness through cultural and media channels.

To deepen the analysis, we performed the same calculation using a different target:

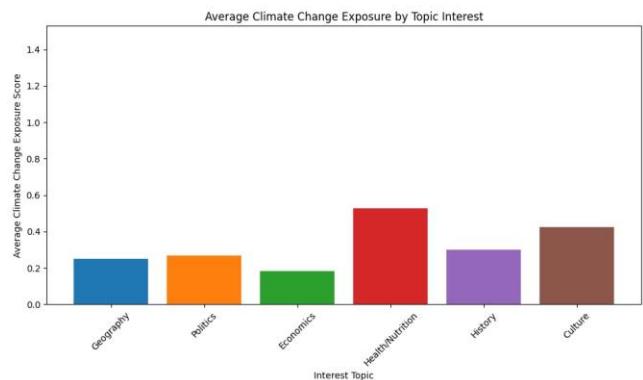- "How often do you hear about global warming in the media?"

Figure 8: Average Climate Change Exposure by Topic Interest

The second chart revealed that exposure to climate change news was highest among the same two groups, those interested in Health & Nutrition and Culture. This suggests that these interests not only correlate with awareness, but also with increased exposure to climate discourse (Duffy et al., 2022; Global Witness, 2024) further reinforcing the relationship.

These findings suggest two opportunities for climate communication:

1. Leverage existing awareness: Individuals already interested in Health & Nutrition or Culture are potential advocates. They are more climate-aware and may be receptive to messaging about action or behavior change.
2. Target less aware groups: Interests like Economics, Geography, and Politics showed lower average awareness and exposure. Tailoring climate messaging to align with these topics (e.g., economic benefits of climate action) could help bridge the perception gap.

CONCLUSION

Our research provided insight into the complexities of climate change beliefs. The relationship between beliefs in certain conspiracy theories negatively related to perceived climate threat and scientists' credibility, while some still had compartmentalized those beliefs. Second, there is a huge gap between climate concern and climate action; the more immediate the harm to a membership group the more likely action will happen, perceived personal cost will often tempt the individual into a "tradeoff mindset". Openness had a positive relationship to environmental awareness, while Extraversion and Neuroticism had unexpected negative relationships with climate change. Interest in Health & Nutrition and Culture aligned with increased climate awareness and media awareness. These findings suggest we could mediate existing awareness in targeted groups or enhance climate change engagement methods to mitigate the perceived barriers to action.

REFERENCES

[1] Minto, B. (1996). *The Pyramid Principle: Logic in Writing and Thinking*. Pearson Education Limited.
[2] *Imhoff, R., & Bruder, M. (2014). Speaking (un-)truth to power: Conspiracy mentality as a generalised political attitude. Frontiers in Psychology, 5, 279.*
[3] Gelman, A., & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
[4] Global Witness. (2024, December 12). *Climate of confusion: Bot-like accounts sharing climate disinformation and conspiracies*.
[5] Tam, K.-P., & Chan, H.-W. (2023). Conspiracy theories and climate change: A systematic review. Journal of Environmental Psychology, 91, 102129.
[6] Cologna, V., Kotcher, J., Mede, N. G., Besley, J. C., Maibach, E. W., & Oreskes, N. (2024). Trust in climate science and climate scientists: A narrative review. *PLOS Climate*, *3*(5), e0000400.
[7] Duffy, B., Malcolm, F., May, G., Hewlett, K., & Haggar, T. (2022). *Public perceptions on climate change* (PERITIA Project Report). King's College London, Policy Institute.
[8] van der Linden, S. (2015). *The social-psychological determinants of climate change risk perceptions: Towards a comprehensive model*. *Journal of Environmental Psychology*, 41, 112–124.

APPENDIX

Declaration section

This report used ChatGPT and Gemini to assist in drafting, editing, and structuring some report sections. All content was reviewed and finalized by the student team.

We also utilized AI tools such as ChatGPT and Gemini for faster and more efficient debugging of the data analysis code. These tools also proved useful in assisting with the analysis of the results, as well as researching authoritative sources to corroborate our findings.

Team Diversity

Academically, we come from petroleum, chemical, and mining engineering, which gave us a solid technical foundation.

Culturally, our team represents various Mediterranean, North African, and Middle Eastern regions. This offers diverse perspectives that shaped both our analysis and survey design, allowing us to approach challenges from multiple angles and find innovative solutions. The richness of our backgrounds fosters a uniquely collaborative spirit, and we sometimes even share a "dab" of playful camaraderie, reinforcing our unity and lightheartedness.

Division of Labor

When it comes to brain storming, reviewing work, and planning the next steps, our efforts were reflective of a true team. We met several times from before the Warmup Assignment till the last day of the deadline. As with every group project, we had disagreements, but we handled them with elegance and emerged from the project having formed deeper connections with our colleagues.

By the end of this project, each and every one of us will remember fondly – and painfully – the countless hours we spent with each other around Torino at PoliTO and other places trying to get this project together. At the end, we can say this is a reflection of our collective work and ideas. When it comes to individual specialized responsibilities, they can be summarized as follows:

Name	Tasks
Hamdan, Dana	Coding, Readme, & Voice Over
Shedid, Mohamed	Report & Presentation
El Sahmarani, Mohannad	Report, Presentation, & Script
Mehri, Ali	Report & Script
Mansour, Anthony	Report, Presentation & Voice Over
Bayat, Amir	Report & Script
Khairy, Sherif	Coding & Report
Petrarulo, Antonio	Report & Script

Supplementary Files & Resources

- **Video Presentation**:

  YouTube Link:
  https://youtu.be/Tz5E9_hGzq8

  Downloadable Link:
  https://drive.google.com/file/d/1n8LHjNxICH4SKn7Oe GPExlqGF--_c5RQ/view?usp=sharing

- **Code & Data Archive (ZIP)**:

  https://drive.google.com/drive/folders/19i_M59MUuuL YouqkQaeq4nKe3VAVZ4TE?usp=sharing

- **GitHub Repository**:

  https://github.com/Clim8s/Beyond-the-Thermometer-Deconstructing-Climate-Change-Perceptions-and-Pro-Environmental-Behavior