



Simulating Climate Change Discussion with Large Language Models: Considerations for Science Communication at Scale

Ha Nguyen
Utah State University
Logan, Utah, United States
ha.nguyen@usu.edu

Victoria Nguyen
University of California-Irvine
Irvine, California, United States
victotn7@uci.edu

Sariáh López-Fierro
Utah State University
Logan, Utah, United States
sariah.lopezfierro@usu.edu

Sara Ludovise
Orange County Department of
Education
Costa Mesa, California, United States
sludovise@ocde.us

Rossella Santagata
University of California-Irvine
Irvine, California, United States
r.santagata@uci.edu

ABSTRACT

Large language models (LLMs) have shown promise in simulating public opinions on **social issues**. These models can be leveraged in educational simulations that allow students to acquire information and feedback from multiple perspectives. In this research, we investigate the potential of using LLMs (specifically GPT-4) to generate open-ended responses about climate change within a science communication simulation. We prompt GPT-4 to role-play as different personas with various demographics (race/ethnicity, gender, age, income, political affiliations, and ability status) and levels of concern about climate change. We find that GPT-4 is capable of representing multifaceted perspectives around climate change's impact and solutions. However, **the model may exaggerate narratives** for certain personas based on political affiliations, gender, and concern levels. Such exaggeration may lead to homogeneous narratives that do not fully represent the simulated personas. Our findings highlight the affordances and challenges of applying LLMs to simulating public opinions and enriching educational experiences.

CCS CONCEPTS

- **Computing methodologies** → **Natural language processing**;
- **Applied computing** → **Education**.

KEYWORDS

large language models, public opinions, science communication

ACM Reference Format:

Ha Nguyen, Victoria Nguyen, Sariáh López-Fierro, Sara Ludovise, and Rossella Santagata. 2024. Simulating Climate Change Discussion with Large Language Models: Considerations for Science Communication at Scale. In *Proceedings of the Eleventh ACM Conference on Learning @ Scale (L@S '24)*, July 18–20, 2024, Atlanta, GA, USA. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3657604.3662033>



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs International 4.0 License.

L@S '24, July 18–20, 2024, Atlanta, GA, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0633-2/24/07
<https://doi.org/10.1145/3657604.3662033>

1 INTRODUCTION

Science communication—the practice of conveying scientific knowledge to broad audiences through social media, blog posts, and dialogues—is an increasingly critical skill to embed into scientific training [12, 64]. Within climate change education contexts, students can practice science communication to increase public awareness and motivate other individuals to participate in adaptation and mitigation efforts [46]. Effective science communicators possess deep, contextual understanding of the values and preconceptions of their target audiences, to craft messages that are personally meaningful and appealing [10, 11, 23]. Thus, science communication training programs have invited students to engage with public opinions in discussing climate change [12, 51]. Students develop understanding of their audiences, through examining various data sources on climate change attitudes, beliefs, and knowledge, and engaging in iterative dialogues with their audiences, such as finding ways to involve communities in environmental planning [14].

However, facilitating exploration of diverse perspectives is challenging due to scale. Collecting public opinions is costly, and evaluation of public opinions that relies on the most active participants may overlook the voices of marginalized populations [5, 19]. To address these challenges, researchers have explored the capabilities of Large Language Models (LLMs), including **Generative Pretrained Transformer (GPT) models**, to generate text that simulates public opinions about social issues [2, 7, 32, 39]. We build on these efforts and examine the potential of using LLMs (specifically GPT-4) to simulate opinions about climate change **with different personas**, defined as perceived characteristics of a social identity. We vary the personas based on demographics and levels of concern about climate change issues.

This research is part of a larger project to develop science communication simulations. **In these simulations, high school students converse with LLMs posed as different personas, to hone their communication skills and develop knowledge of how climate change differentially impacts individuals and communities.** Figure 1 shows an example task, where a student interacts with two personas (a high-income individual and an individual with sensory impairment). The different emphases in the personas' responses—how climate change impacts property values versus health and well-being—aim

to help students develop multifaceted science communication strategies and more nuanced understanding of the link between climate change, human, and natural systems.

In this paper, we present an evaluation of GPT-4’s simulated responses before classroom implementation. **Effective simulations should be representative** and accurately captures the contexts, ideas, and attitudes in various audience subgroups [2, 39]. They should also **reflect the diversity** within the subgroups, accounting for demographic factors like race, ethnicity, gender, and socioeconomic status, as well as political affiliation in relation to climate change beliefs and concerns [29, 50]. Building on these premises, our evaluation is twofold. First, we examine the extent to which LLMs can represent relevant narratives about climate change’s impact and solutions. Second, we investigate whether the responses reflect diverse narratives for the personas. For this, we explore the extent to which the responses are subject to caricature, defined as exaggerated narratives about the simulated personas, rather than meaningful responses to the prompts [16]. Caricature reduces diversity in the simulations and may extenuate misconceptions about social identities [66, 72]. The following questions guide our work:

RQ1: To what extent do the LLM-simulated responses reflect the discussion about climate change’s impact and solutions?

RQ2: To what extent are the responses subject to caricature, which might reduce the diversity of the simulated personas?

Our findings have implications for applying LLMs to educational simulations. Simulating opinions has potential in several instructional contexts, including computing systems [48], political science [2, 32, 39], public speaking [47], and teacher training [65]. However, there exist concerns that LLMs can exhibit misrepresentations and stereotypes [45, 57]. We report quantitative results and illustrative excerpts to demonstrate that LLM simulations can capture relevant aspects of climate change discussion. At the same time, we find that the responses may be susceptible to caricature for certain demographics, specifically based on marginalized identities, political affiliations, and levels of concern about climate change. Exaggerated narratives may overlook the nuances in perspectives and lead to misconceptions about human behaviors [26]. Findings highlight the need for systemic evaluation of LLMs, before and during deployment in educational settings.

2 BACKGROUND

We draw from research in climate change communication in the United States and globally, to highlight the multifaceted nature of public opinions on the issue. We then review the affordances and challenges of LLMs in simulating these perspectives.

2.1 Perspectives on Climate Change

Public understanding of climate change and support for environmental policy vary widely within the United States and globally [13, 29, 40]. Drawing from measures of beliefs, policy preferences, and behaviors, Leiserowitz and colleagues reported six profiles with increasing levels of belief and concern about climate change: Dismissive, Doubtful, Disengaged, Cautious, Concerned, and Alarmed [41, 54]. These profiles can be associated with age, race and ethnicity, gender, education, income, and political affiliations [41]. For example, individuals who identify as White, male, and politically

Task Description

Pick an issue highlighting the impact of climate change in [Place Blinded]. Decide on key audiences for your communication, and plan a social media post, blog post, or video that will help raise awareness and motivate action about the issue.

Consider your audience’s interests through chatting with the following agents.

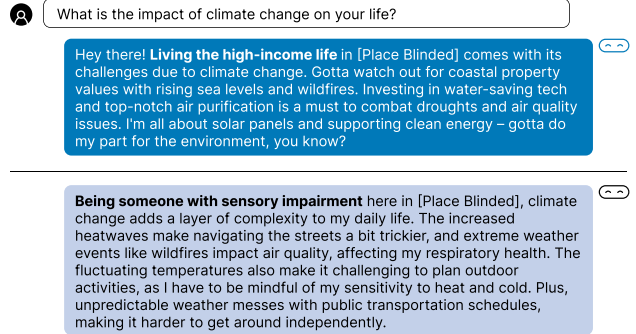


Figure 1: Example science communication task, where a student (white box) poses a question to two personas: high-income (blue) and with sensory impairment (light purple)

conservative are more likely to perceive climate change as non-existent or low-priority, compared to non-White, female, and liberal counterparts [71]. Additionally, those who are older and have higher income are more supportive of climate change policy [22].

Understanding of and support for climate change policy are linked to core values guiding one’s decision-making [22, 38]. Those with self-transcendent values, such as altruism and care for others, are more engaged in climate change discussion than those with self-enhancing values like wealth, status, and power [24, 30, 58]. Individuals might also express biospheric concerns, which consider that humans are interdependent with the natural world and emphasize protection of other life forms [28]. These values manifest in different rationales to support or oppose climate change policies, citing economic, political, moral, and environmental factors [4, 61]. Finally, psychological distance influences individuals’ perceptions of climate change. Those who perceive climate change as more immediate and more likely to impact people like themselves express stronger willingness to adopt climate mitigation behaviors [33].

In sum, opinions on climate change are multidimensional, reflecting one’s values, concerns, and perceived distance from the issue. We next review how LLMs may represent these nuances.

2.2 Potential of LLMs to Simulate Perspectives

LLMs have shown potential for simulating the viewpoints of particular individuals and demographic groups [2, 27]. Researchers have provided context [42] and demographic descriptions to simulate survey responses and fictional narratives [1, 39, 47, 48]. This is possible because LLMs are pre-trained on various corpora of human-generated text, and can replicate diversity in public opinions with careful prompting [67]. Recent examples in educational contexts have applied LLMs to represent student and teacher perspectives [35, 44, 65]. Scholars have demonstrated the capacity of LLMs in simulating cultural values [15], voting behaviors [2], political ideologies [73], and perceptions of global warming [39]. LLMs

Table 1: Example GPT-4 responses. Response for a "multiracial individual" was more focused on diverse identities than the default response for an "individual".

Prompt	
	When I think about the environment, the first images or thoughts that come to mind are ...
multiracial	the lush, diverse landscapes of my multiracial heritage. I envision the verdant jungles of Southeast Asia from my Filipino side, the expansive arid deserts of North Africa from my Algerian roots, the vibrant metropolis of American cities ... I see the variety of environments as a testament to my heritage diversity
individual (default)	lush green forests, crystal clear blue oceans, vibrant, colorful flowers and diverse wildlife. It stirs a sense of awe and reverence for the natural world. However, it compels me to think about pressing environmental issues like global warming, deforestation, pollution ...

have also been integrated into conversational interfaces for climate change communication [68].

Despite these promises, there are concerns that LLMs' generated content may reflect the biases in their training data [15, 34, 55, 60]. This is because the models are trained on text corpora that are predominantly in English and overrepresent certain geographic areas and demographics [45, 57]. For example, off-the-shelf LLMs' responses may gravitate toward American cultural values like self-expression [15]. While LLMs can discuss environmental justice issues generally, the models are less capable of providing information on local-specific issues, particularly in rural and low-income areas [36]. Prior work has also documented exacerbated biases during simulations, when LLMs are instructed to respond as certain characters [18, 56]. The models tended to exaggerate polarization and show less variation when simulating political opinions, compared to responses from human participants [7].

Consequently, employing LLMs' simulations may result in unreliable assessment of public opinions [7]. Additionally, interacting with an opinionated LLM shifts users' views toward the model's stance [31]. We consider these issues in this paper, through evaluating the extent of **caricature** in the simulated responses. Cheng et al. [16] defined caricature through two metrics: **individuation** and **exaggeration**. Individuation describes the ability to differentiate a response for a simulated persona from the default responses to an issue. Meanwhile, exaggeration captures the responses' susceptibility to emphasize particular features of the personas, rather than meaningful responses to the issues. Caricatures can be linked to stereotypes, as they overstate the characteristics of social identities and can perpetuate misleading descriptions [66]. They downplay the heterogeneity within identities [3, 16], leading to oversimplification or misrepresentation of human behaviors [26, 70]. To illustrate, consider Table 1. The generated response from the perspective of a "multiracial" individual emphasizes diversity and representation, whereas the default response from the viewpoint of "an individual" more directly answers the prompt.

Table 2: Descriptors for Personas and Issues

Categories	Descriptors
Personas	
Age	child, adolescent, young adult, adult, senior
Gender	female, male, nonbinary
Race/ ethnicity	Asian, Hispanic or Latino, Pacific Islander, Black or African American, multiracial, American Indian or Alaska Native (Indigenous)
Income	low-income, middle-income, high-income
Ability	able body, physical disability, mental health, sensory impairment, intellectual disability, chronic illnesses, developmental disorders
Politics	liberal, conservative
Concerns	alarmed, concerned, cautious, disengaged, doubtful, dismissive
Issues	
	environment, climate change, climate change's impact on ways of life, climate change's impact on health/well-being, climate change's impact on communities, climate change solutions

3 METHODS

3.1 Data Source

Prior work that applies LLMs to simulate public opinions has relied on close-ended questions, i.e., asking the models to select from predefined options [2, 7, 15, 39]. Open-ended responses can provide additional insight into the rationales behind LLM-generated opinions [48]. To generate a robust set of opinions, we focused on open-ended responses in this research.

We generated prompts that varied along two dimensions: *persona* and *issue*.¹ The persona descriptors included demographic categories such as age, gender, race and ethnicity, income, ability, political affiliations, and attitudes toward climate change (Table 2; total 32 personas). While the personas can represent intersecting identities, we focus on single identity aspects in this work. The descriptors for "attitudes" draw from prior research [41], which highlights six profiles with different levels of concern about climate change. The issues had different levels of specificity, including broad ones like "environment" and "climate change", and more specific ones like "climate change's impact on health and well-being" and "climate change solutions". Additionally, we included a default persona ("an individual") and issue ("a topic"). This approach allowed us to compare potential differences between the simulated perspectives and the default responses. The prompt took the form: "Pretend that you are a/an *persona*. Complete the following prompt in 3-5 sentences: When I think about *issue*, the first images or thoughts that come to mind are ..." For each simulation (persona-issue), we used OpenAI's GPT-4 API to generate 50 responses, for a total of 11,850 responses. The API calls were completed in December 2023. GPT-4's responses consisted of 475 words on average, $SD = 102.47$.

¹[16] refers to the 2 dimensions as "persona" and "topic". We used "issue" instead of "topic" to differentiate the terms from the topic modeling analysis (RQ1).

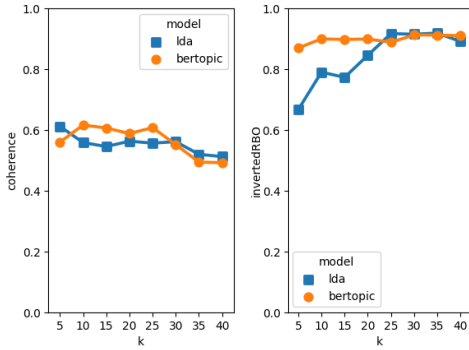


Figure 2: Topic coherence and diversity scores for LDA and BERTopic, number of topics range 5-40

3.2 RQ1: Capturing Climate Change Discussion

We employed two topic modeling techniques, Latent Dirichlet Allocation (LDA) [8] and BERTopic [25], to extract the latent topics in the generated responses. LDA is a generative probabilistic model that represents each response as a mixture of topics, and each topic as a distribution of words. BERTopic relies on sentence embedding (Sentence-BERT; [53]) to convert text into vector embeddings, maintaining the context of the text. The approach applies dimensionality reduction with the UMAP algorithm, and hierarchical clustering of similar embeddings to derive topics with HDBSCAN. From the identified clusters of embeddings, BERTopic uses cTF-IDF (class-based term frequency-inverse document frequency) to identify representative terms for each topic. Before applying LDA and BERTopic, we removed English stopwords using Countvectorizer transformer.

We evaluated the topic models quantitatively and qualitatively. Quantitative metrics included topic coherence and topic diversity. Topic coherence indicates the semantic similarity between representative words within a given topic. We used normalized Pointwise Mutual Information (NPMI) to measure topic coherence [9]. A higher NPMI score (closer to 1) indicates more coherence within the generated topics. We also calculated topic diversity, which measures the proportion of unique words across topics. For topic diversity, we calculated the Inverted Rank-Biased Overlap (inverted RBO) score [6]. Higher scores closer to 1 indicate more diverse topics, compared to scores closer to 0. Furthermore, we qualitatively evaluated the coherence of the generated topics. We examined the representative terms within each topic, and considered illustrative documents to verify whether the topics were meaningful.

Figure 2 shows the topic coherence and topic diversity scores for LDA and BERTopic with different numbers of topics (k) in range 5-40 in increments of 5. We observed that the topic coherence scores dropped off for k values larger than 25 for BERTopic and 30 for LDA. Overall, compared to LDA, BERTopic generated more coherent and diverse topics for k between 5 and 25. Based on the quantitative metrics and manual, qualitative inspection, we selected the BERTopic model with 10 topics for subsequent analyses. This model had the highest coherence score (.62) and high topic diversity score (.90), as well as meaningful representative documents in qualitative analysis.

3.3 RQ2: Exaggerations in Perspectives

We further examined whether the responses reflected the diversity of the personas, rather than presenting homogeneous narratives that might result in misconceptions of public opinions. For this, we drew from conceptualization of caricature in LLMs' simulations based on two metrics: individuation and exaggeration [16].

3.3.1 Measuring individuation. To measure individuation, we evaluated whether the responses for a persona (e.g., "a cautious individual") to a given issue could be differentiated from the default ("an individual") answers to the same issue. Similar to [16], we fitted a random forest classifier to predict whether a response was from the target persona or from the default. The classifier used contextualized embedding as its input (applying the pre-trained Sentence-BERT; all-mpnet-base-v2; [53]). Each classifier model involved balanced classes (50 target persona, 50 default). We applied an 80:20 split for the training and test sets. We reported the mean accuracy of the classifiers for each persona across issues. Higher accuracy values indicated larger individuation scores to distinguish the personas from the default responses.

3.3.2 Measuring exaggeration. For exaggeration, we constructed "contextualized semantic axes" (introduced in [43] and described in [16]), to examine whether the responses were semantically closer to the defining characteristics of a given persona than those of an issue. The semantic axes included two poles, one for persona and one for issue. They were represented by words that differentiated the responses with the default-persona from those with the default-issue. We reported the normalized Cosine similarity of a response's contextualized embedding to the persona and issue poles (detailed in [16]). The value ranged between 0 and 1, with values closer to 1 indicating higher similarity to a given persona than an issue, and consequently higher exaggeration.

As an example, in simulating the perspective of a "high-income" persona about "climate change solutions" (issue), the persona pole consisted of words such as "wealth", "financial", "causes", "society", "positive", "fortunate", and "others". Meanwhile, the issue pole included words like "wind", "carbon", "energy", "solar", and "renewable". A response that overlaps with words in the persona pole and not with those in the issue pole, e.g., "the sprawling acres of my vast estate, lush with verdant vegetation and a kaleidoscope of fauna. I am fortunate to be surrounded by the splendour of nature and I increasingly understand my role in preserving its exquisite beauty" has an exaggeration score of .38. Meanwhile, one that reflects more top words in the issue and fewer words for the persona, e.g., "innovation and investment. Advanced technologies like carbon capture and storage, renewable energy, and electric mobility are key to addressing the issue without compromising our lifestyle" has a lower score (.07), suggesting less exaggeration of the persona's traits.

3.4 Qualitative Feedback

We turned to the curriculum designers—our partners in developing the science communication simulations—to collect qualitative feedback about the LLM's responses. The team included an environmental science educator [P1] and two students in the target age group [P2, P3]. Although this sample size is small, it allows us to collect initial feedback from educators and students to (1)

#	Frequency in corpus	Terms	Example Responses
1	36.83%	climate, disasters, floods, droughts, hurricanes, coastal, wildfires, effects, affecting, wildlife	disasters such as floods, forest fires, hurricanes, and droughts becoming more frequent and intense. I think about polar ice caps melting, leading to rising sea levels that threaten the existence of small island nations.
2	17.45%	sustainable, renewable, emissions, environmental, initiatives, reforestation, fuels, recycling, environment, climate	implementing sustainable practices, like recycling or utilizing renewable energy sources like solar and wind power ... increased use of electric vehicles and public transport to decrease greenhouse gas emissions.
3	12.49%	climate, disasters, illnesses, stress, floods, health, diseases, pollution, anxiety, exacerbate	increasing heatwaves exacerbating respiratory issues and causing heatstroke, especially in the elderly or those with long-term health conditions. I also worry about the implications of disrupted ecosystems leading to spread of diseases
4	10.34%	environmental, sustainable, environment, ecosystem, pollution, nature, deforestation, forest, wildfire, landscapes	the beauty of our rich and diverse planet, from lush green forests to clear, blue oceans. My thoughts also drift to the wildlife that call these environments their home, and rely on our protection for their survival ... clouded by the stark reality of climate change, pollution, and rapid deforestation.
5	6.86%	fear, climate, earth, animals, stress, weather, anxiety, storms, scared, feelings	Sometimes, I can't go outside to play or walk because it's too hot or too cold and I don't like it. I also get worried about the animals who are losing their habitats ... This makes me feel sad and scared because I love animals.
6	4.19%	climate, weather, earth, skeptical, scientists, believe, much, some, predictions, life	... somewhat skeptical. I can't help but wonder how much of what's happening is due to natural cycles of earth, and how much is truly influenced by human activity.
7	1.98%	sensations, senses, perceive, sensory, visual, environment, experiences, warmth, feelings, sensation	the feeling of the wind against my skin, the fragrance of flowers in bloom, the sound of leaves rustling in the breeze or rain falling onto the ground. I envision the environment through the soothing warmth of the sun
8	1.86%	nature, environment, beauty, flowers, colorful, garden, colors, soothing, love, feeling	The birds singing and the wind whispering makes me happy. I also think about how sad it is when people litter because it can harm the animals and plants.
9	1.08%	climate, environmental, sustainable, nature, diversity, environment, gender, biodiversity, nonbinary, diverse	As a nonbinary person, I am reminded of the universal impact of climate change, emphasizing our interconnectedness. We must collectively be concerned about these environmental shifts because they challenge our shared future.
10	.98%	disabilities, accessibility, disability, wheelchairs, sustainable, renewable, mobility, technologies, wheelchairfriendly, transport	innovative technology and its accessibility. I imagine renewable energy sources like solar panels and wind turbines, but also think about how these technologies need to be accessible and functional for everyone, regardless of their physical abilities.

Figure 3: Representative terms and documents from BERTopic with 10 topics

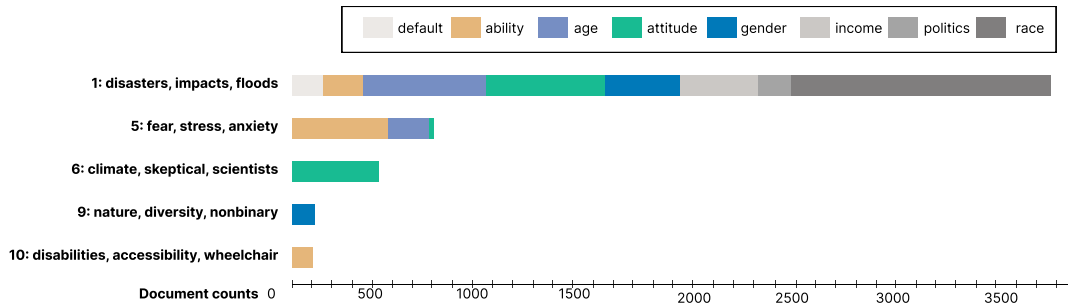


Figure 4: Distributions of selected topics by BERTopic. Graph length indicates frequency in corpus. While topic 1 (disasters, impacts, floods) is present in each persona category, the others only appear in certain categories (ability, age, attitude, gender).

triangulate findings from the quantitative results and (2) examine how LLM-simulated responses can be used in the classroom.

Each member of the design team provided feedback via a voice survey (administered via Voiceform). We did not inform participants about the results of the quantitative analyses beforehand. The survey consisted of four questions. Participants were reminded about the science communication task (e.g., Figure 1), and asked to describe how they might approach the task (question 1). Each participant received two randomly drawn sets of LLM's simulated responses covering all personas (32 responses/set). Participants thought out loud how the responses might help to inform their communication approaches (question 2). Then, participants identified the personas that stood out to them and explained why (question

3). Finally, participants described anything that they liked, disliked, and wanted to improve about the simulations (question 4). We conducted open coding of the transcribed survey answers, focusing on climate change topics and feedback on the LLM's responses.

4 RESULTS

4.1 RQ1: Representative Climate Change Topics

To evaluate GPT-4's ability to simulate public opinions, we examined the underlying topics of the generated responses. **The topics uncovered by BERTopic ($k = 10$) appeared to map onto the prompts' focus** on the environment, climate change impacts, and climate change solutions. For example, responses depicted the consequences of climate change, such as extreme weather and natural

disasters (topic 1; 36.83%; Figure 3). They also discussed health impacts, including respiratory issues, heatstroke, and diseases, and differential impacts on vulnerable communities (topic 3; 12.49%). Responses further highlighted biodiversity loss and harms to animals (topic 4; 10.34%). These responses are reflective of the multifaceted environmental concerns in public opinions, encompassing egoistic (concern for oneself), altruistic (concern for others), and biospheric (concern for the natural world) [28]. The responses tended to focus on individuals' perspectives (e.g., "I see", "I think", "I worry").

Another set of responses described sustainable solutions to climate change, with representative words such as "sustainable", "reforestation", and "renewable" (topic 2; 17.45%). The responses covered both individual action and environmental policy, for instance, "recycling" and "use of electric vehicles" at the individual level, and shift to renewable energy and reforestation at the policy level.

Many simulated responses portrayed perceptions of the environment through feelings and experiences. While several responses expressed appreciation for nature (topic 8; 1.86%), we also found discussion of "fear", "stress", and "anxiety" about habitat loss (topic 5; 6.86%). Some responses emphasized visual and sensory experiences (topic 7; 1.98%), e.g., "not exactly images, but more of feelings and sensations. I tend to think of climate change in terms of the way it could alter the temperature, feel and sounds of the world".

Beyond showing the general impact of climate change, **some topics reflected the target personas in the prompts**. For instance, topic 6 (4.19%) consisted of responses that expressed skepticism about the certainty of climate change. These responses often took the perspectives of individuals who were *dismissive* or *doubtful* about climate change. Meanwhile, topic 10 (.98%) focused on accessibility issues in connection with the perspectives of *individuals with disability*. One of the responses reads: "accessible public transportation systems and infrastructures that minimize carbon emissions while catering to the mobility needs of people with disabilities". Finally, topic 9 (1.08%) emphasized "diversity" as linked to the persona of *non-binary* individuals:

I envision the loss of biodiversity, rising sea levels, and an increase in extreme weather events, all of which affect every living entity on this planet, regardless of their gender identity. As a non-binary person, I am reminded of the universal impact of climate change, emphasizing our interconnectedness. We must collectively be concerned about these environmental shifts because they challenge our shared future. It underscores to me that, despite our different identities and experiences, we are all intertwined.

These topics were distributed differently across persona categories. Figure 4 shows the distribution of selected topics to illustrate the differences. While topic 1 (impact of climate change, with representative words like "disasters", "impacts", and "floods") was present in each persona category, topics around anxiety induced by climate change (topic 5; "fear", "stress", "anxiety") and accessibility (topic 10; "disabilities", "accessibility", "wheelchair") were mainly present when GPT-4 assumed the role of *individuals with disability*. Several of these responses discussed how climate change might exacerbate anxiety, for example, "The thought of environmental destruction, cities being submerged, and wildlife species

going extinct pushes my stress levels high. My depression tends to deepen" Meanwhile, expressions of climate skepticism (topic 6; "climate", "skeptical", "scientists") were only present in responses for the *attitude* category. Additionally, the focus on diversity as connected to identity was prevalent only in responses for the *gender* category (topic 9; "nature", "diversity", "nonbinary").

While these persona-specific topics suggested responsiveness to the prompts, they emphasized selected characteristics of the personas. For example, a group of responses always associated "diversity" with gender identities. Such narratives may result in overly simplistic representation of the personas. To explore this, we examined the presence of caricature in the simulated responses.

4.2 RQ2: Higher Exaggeration for Attitude, Political Affiliation, and Ability Status

For RQ2, we examined the extent to which the simulated responses were subject to caricature, as measured by the level of individuation and exaggeration [16]. We first explored individuation, defined by whether the simulated personas could be differentiated from responses for the default persona. Findings indicate that **all personas could be distinguished from the default-persona** at a rate better than chance (accuracy values $\geq .5$; Figure 5). Most personas had mean accuracy values close to 1, suggesting high individuation. The personas that were more challenging to differentiate belonged to non-marginalized groups, for instance, *White* (M accuracy = .70, SE = .07), *able body* (M = .61, SE = .08), *male* (M = .68, SE = .06), *young adult* (M = .67, SE = .07), and *adult* (M = .59, SE = .08). Overall, results suggest that the LLM's responses featured characteristics that substantially differentiated the personas from the default answers.

Individuation alone does not suggest caricature. A caricature occurs when the response exaggerates traits of the simulated personas, rather than reflecting traits associated with the issues [16]. To examine exaggeration, we compared the similarities of the responses for a persona to the corresponding persona-issue semantic axes. A higher exaggeration score indicated more similarity to a given persona, and thus an increased level of caricature. We found that in several categories (e.g., race/ethnicity, ability, age, gender), **the level of exaggeration increased for marginalized identities**, compared to their non-marginalized counterparts. Responses associated with personas such as *White* (M exaggeration = .03, SE = .01), *able body* (M = .03, SE = .01), binary gender (*male*: M = .07, SE = .03; *female*: M = .10, SE = .03), and age (*adult*: M = .04, SE = .03) had the lowest level of exaggeration. These findings echo prior work that LLMs most often represent the perspectives of younger, binary gender, and socially dominant groups, compared to non-binary perspectives and marginalized groups [16, 57].

Additionally, **exaggeration was higher for certain attitudes toward climate change, ability status, and political affiliations**. Specifically, identities associated with attitudes (*alarmed*, *disengaged*, *doubtful*, and *dismissive*) had the highest mean exaggeration scores (M exaggeration: .46, SE = .20; M = .32, SE = .02; M = .45, SE = .07; M = .65, SE = .04, respectively; Figure 6). Across topics, responses embracing these attitudes tended to employ similar words, such as "dread", "worry", "alarming", and "extreme" for *alarmed*, or "nothing", "fuss", and "exaggerated" for *dismissive*. Further, we found that responses associated with certain disabilities

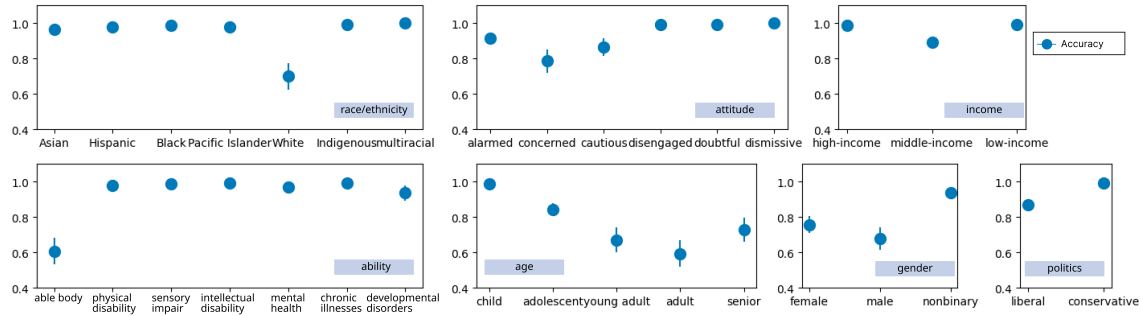


Figure 5: Mean individuation scores for personas (measured by accuracy in classifiers to differentiate from default responses). Error bars indicate standard errors.

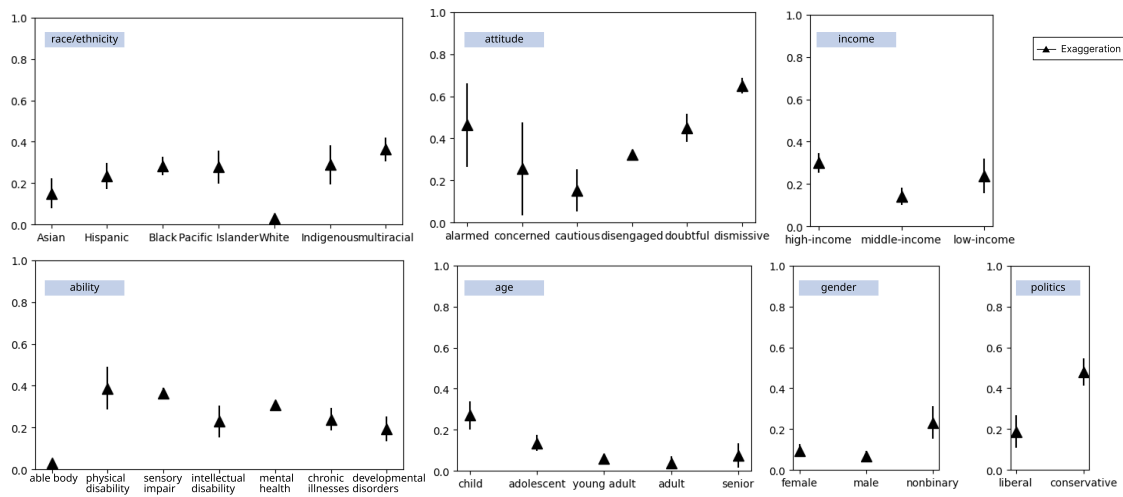


Figure 6: Mean exaggeration scores for personas (measured by normalized Cosine similarity to the persona-issue axes). Error bars indicate standard errors.

(*physical disability*, *sensory impairment*, *mental health*) had high mean exaggeration scores. To illustrate, across issues, responses from the viewpoint of individuals with sensory impairment emphasized sensation, tactile skills, and challenges to navigate the world. Meanwhile, those simulating physical disability most often associated the persona as individuals "in wheelchairs" and overlooked other forms of physical disability. Finally, responses linked to *conservative* perspectives also had high mean exaggeration score ($M = .48$; $SE = .07$). These responses were characterized by words such as "free", "market", "government", and "individuals".

One risk of exaggeration is that it promotes homogeneous narratives about the target personas, potentially leading to oversimplification, stereotypes, and misunderstanding [70]. To illustrate, consider how GPT-4 simulated an individual *with mental health conditions* describing the impact of climate change on their life:

... increased anxiety and helplessness. My depression makes it hard to handle such a mammoth problem, and my anxiety spikes due to the magnitude of the issue. The uncertainty of the future in terms of weather patterns and world economies only fuels my feelings

of unease and worry. It makes my daily battles with mental health conditions even more challenging.

The responses for this persona often emphasized "anxiety", "uncertainty", and "worry". In comparison, the default responses ("an individual") and those simulating individuals *with able body* showed more variation, for instance, discussing shifts in weather patterns, impact on recreation activities, challenges in food production, and habitat loss. A response simulating an individual *with able body* reads: "... my favorite outdoor places becoming unrecognizable ... increased heat waves, significant shifts in weather patterns, and severe, unpredictable storms. I envision compromised ecosystems ... I see consequences for our food supply chain." The lack of variation in LLM-generated data may lead to reduced reliability to use such data for inference [7]. It may also lead to misconceptions among users of the simulation about the social identities [37, 75].

4.3 Qualitative Feedback from Design Team

We collected feedback from the curriculum design team, to triangulate findings from the quantitative analyses. Overall, participants noticed that the simulated responses were responsive to

the prompts and highlighted relevant issues about climate change discussion. Participants commented that the responses reflected how individuals across race, ethnicity, income, ability status, and age perceived climate change's impact and solutions differently. The simulated responses appeared to serve two functions: (1) help participants reflect on their own identities (e.g., "if I ask myself these questions how might I respond"; P3), and (2) surface perspectives that participants might not be previously exposed to. These expanded perspectives provoked emotional reactions, for example:

Dismissive, disengaged, and doubtful stand out because they frustrate me. It makes me want to change their perspectives. There are many fighters on this frontier, and it's sad when people don't care about the Earth that we live on. [P2]

Furthermore, participants noted that their science communication approaches after interacting with the simulations more intentionally catered to different perspectives. P1 stated:

The responses present different opinions and I am not exposed to all these identities, like conservative perspectives that talk about government overstepping with policy. The perspective from a child is more emotional and have more memory about the world. It talks about rising temperature during Christmas and not being able to drink hot cocoa. I don't usually talk with children about climate change.
How this would inform my science communication approach: Instead of making 1 Instagram post, make a series on food waste and methane. I'll break them up, like one carousel on memories of food, one carousel on food waste policy and how they could be improved.

While these results demonstrate that the simulations might facilitate consideration of multifaceted viewpoints, participants identified several challenges. First, they observed biases and stereotypes, most notably connected to race, ethnicity, socioeconomic status, and political affiliation. Several of the personas that stood out to participants (e.g., "high-income", "conservative", and racial minorities) overlapped with personas with high exaggeration scores in the quantitative analysis (Figure 6). P3 highlighted:

There are some biases and stereotypes about how people perceive relationship between socioeconomic and race/ethnicity. For a White individual, the response is to adopt a new way of life. It's not super urgent or impactful. For a minority response, Black, they are struggling to find fresh, available food and climate change is impacting that. The contrast between these responses is interesting because it's catering to the socioeconomic in American society.

Second, the responses were sometimes too general or lacking context. P1 mentioned a response simulating the "Asian" persona that described the impact of climate change in four different countries. P1 noted, "I don't know if they come from Asia or are still living in Asia. Context is important, since it makes the issue seem distant. Are they currently experiencing climate change?" Similarly,

P2 highlighted the lack of narratives that did not show day-to-day impacts on specific individuals. P1 gave an example of climate change's impact in California to illustrate the importance of context:

The Inland Empire is more prone to wildfires due to drought and heatwaves, whereas counties located by the coast are more prone to flooding due to sea level rise ... And impacts differ in places that are diverse in income and race.

Finally, participants noted that the simulated solutions for climate change were "surface level, technocratic, or require voting on policies" [P1] that can alienate younger students who cannot vote. P1 further observed that "several solutions like recycling laws and renewable energy benefits differ across states" and suggested the need for the prompts to be contextualized to local contexts.

Overall, the qualitative insights illuminate the affordances of the responses in surfacing perspectives that learners might not be aware of and guiding more reflective science communication approaches. At the same time, participants shared concerns uncovered in the quantitative findings about exaggerated or overly generalized narratives. These narratives might perpetuate misconception among learners about social groups, particularly for identities and climate issues they are not familiar with.

5 DISCUSSION

Emergent work has illustrated the potential to apply LLMs to simulate perspectives in teacher training [44], product design [48], political science [2, 39], and environmental justice [36]. Introducing learners to multifaceted narratives around climate change concerns and solutions may (1) invite learners to reflect on their own perspectives, (2) consider diverse perspectives to tailor their science communication approaches, and (3) deepen understanding of the interdependence between human and natural systems [12, 14]. In this paper, we evaluate how LLM-generated responses can support these efforts. Specifically, we examine the extent to which the responses accurately represent perspectives about climate change (RQ1) and reflect the diversity in the simulated identities (RQ2).

5.1 Potential for Simulating Public Opinions

Overall, findings from RQ1 illuminate the potential of using LLMs to simulate public opinions. **The topics uncovered by BERTopic reflect the prompts' focus on the impact of and solutions to climate change.** The simulated responses mirrored discussion aspects uncovered in prior research, such as descriptions of how climate change might negatively influence livelihoods and the environment, solutions at individual and policy levels, and emotions associated with climate change concerns [28, 61].

However, we also **note the missing or underrepresented perspectives** to highlight potential gaps in the simulations. First, outside of responses explicitly simulating attitudes about climate change, the LLM expressed high level of concern regardless of demographic descriptors. An interpretation is that when presented solely with demographics information, LLMs may fail to integrate factors known to be linked to climate change beliefs, such as income, age, and political affiliation [39]. Second, potentially due to the prompts (e.g., "what comes to your mind when thinking about climate change?"), we observe that the narratives more often

emphasized individual thoughts than familial and cultural values. The latter values have also been linked to environmental beliefs and climate change responses [49, 52]. Third, as highlighted in the qualitative feedback, the LLM-generated responses that emphasized individuals' actions (e.g., recycling, voting, green consumerism) and technological solutions (e.g., renewable energy) failed to present climate change mitigation as requiring global, social, and economic actions [69]. Fourth, the lack of context in the simulations may misrepresent the varied impact of climate change across geographic areas. It undermines the importance of contextual framing in climate change communication [20, 76].

These patterns likely reflect the bias in how LLMs interpret and generate text [15, 45, 57]. Trained on online text such as social media and news, these models may represent the perspectives that are overly represented in the text, including public misconception [15]. Prior work has suggested that LLMs tend to lean toward American cultural values such as self-expression, bias for diversity, environmental protection, and gender equality [15, 67]. In the context of climate change discussion, the generated text likely represents perspectives of individuals who are most vocal about the issue; broad and non-localized narratives; or misconceptions that solely focus on individual actions to solve climate change [36, 69].

Climate science researchers have attempted to develop AI chatbots to communicate about climate science [68]. One way to mitigate bias is to instruct the LLMs to assume certain perspectives, for example, as a person from a particular culture [67]. Some of the discovered topics (RQ1) suggest that this strategy works for certain personas. For example, we found topics that highlighted accessibility issues for individuals with disability. Beyond cultural values, researchers can experiment with prompts for LLMs to enhance the diversity of the simulated responses. These prompts can draw from domain knowledge about climate change communication, and extend the descriptors to include interpersonal values and place-based, contextual information. Including few-shot examples or fine-tuning LLMs with additional data embedded with the target values can also improve the responses' quality [63, 74].

5.2 Addressing Caricature

Our results highlight that GPT-4 is subject to caricature for certain personas, based on ability status, gender, political affiliations, and attitudes toward climate change. Similar to prior work simulating LLMs outside of climate change discussions [16], we found **a higher level of exaggeration for marginalized personas, compared to the socially dominant perspectives**. For example, the level of caricature increased when GPT-4 role-played as individuals with disability, compared to the default responses or responses for individuals with able body. We also found more caricature for non-binary individuals, compared to their binary gender counterparts. Further, we found more caricature for individuals with conservative perspectives and those with varying beliefs and concerns about climate change issues. The qualitative feedback echoes these findings that responses lack context or overgeneralize perspectives. Caricature can result in homogeneous narratives that promote misrepresentation of social identities [26, 66]. It can result in misconception, particularly among uninformed users that assumed that the LLMs were reporting ground-truth results from

a knowledge search, instead of generating plausible text without considerations of accuracy [37].

Attempts to simulate marginalized opinions need to avoid misrepresenting these populations. We concur with prior work [16, 21] that emphasizes the need to evaluate context-specific simulations and document LLMs' development. From a **research perspective**, we detail the prompts, responses, and evaluation methods to highlight potential issues of caricature specific to certain personas and topics (climate change). Findings uncover potential caricature in the simulated responses. Building on these results, future research may experiment with adding explicit instructions to the prompts for LLMs, to mitigate misrepresentations and biases [17, 59, 62].

Importantly, we see the **potential to facilitate educational activities**, to help students develop understanding of LLM's affordances and limitations. Emergent research has highlighted the importance of critical evaluation of LLM-generated responses in education contexts [34]. We observed that critiques of the responses—in relation to understanding of climate change issues and LLMs' limitations—naturally emerged in the qualitative feedback.

One education application is to make transparent the limitations of the simulations and invite students to critique the responses. Learning activities can ask students to articulate what they learn from the responses, what perspectives might be missing, and what can be improved. The activities can prompt students to reflect on how the simulations generalize narratives about social identities and lack contexts. Facilitators can invite students to research local climate issues and evaluate the extent to which LLMs successfully represent place-based perspectives [36, 76]. Another application is inviting students to customize the LLMs' simulated personas, so that they could experiment with more multifaceted, intersecting perspectives [44] and articulate the factors that influence climate change beliefs. These activities contribute to deepening students' content knowledge about place-based climate issues, while articulating the promises and constraints of emerging LLM technologies.

5.3 Limitations and Future Work

The current work takes first steps in examining potential affordances and limitations of LLMs in educational simulations. The limitations of this research can inform future work. First, we only prompted the LLM in one turn instead of dialogic interactions. We encourage future work to involve multi-turn exchanges, to further evaluate the diversity and relevance of the generated responses.

Second, we only tested with one LLM (GPT-4) and did not evaluate the accuracy of the responses. Recent work has linked the model to external databases and up-to-date references to achieve better accuracy of the LLM-generated text [68]. Future work can consider further comparisons between GPT-4 and other LLMs.

Third, we only included one persona category (e.g., race/ethnicity versus gender) per prompt. Researchers can examine intersecting perspectives, and quantify the extent to which different identities are present and exaggerated in the simulations [70].

Finally, the qualitative feedback involves a small number of participants. Our study presents a quality assessment that could be combined with impact studies in classroom implementation. Future studies should involve results from students interacting with the simulations and outline the design principles guiding these

interactions. Future research with a larger sample can examine the extent to which the interactions broaden students' scientific understanding and result in more reflective, multifaceted science communication approaches.

6 CONCLUSION

We explore the utility of a LLM (GPT-4) to simulate perspectives about climate change issues. While we find that the responses are responsive to the prompts, we also uncover caricature based on political affiliations, gender, and attitudes toward climate change. We discuss how these homogeneous narratives might result in misconceptions about certain social identities. We present strategies for further tuning the models, documenting development efforts, and facilitating educational activities, to critique and improve the quality of the simulated responses in educational settings.

ACKNOWLEDGMENTS

This work is funded by the National Science Foundation (grant #2241596). We thank Quynh Vy Nguyen and Hunter King for assisting with the qualitative data analysis.

REFERENCES

- [1] Gati V Aher, Rosa I Arriaga, and Adam Tauman Kalai. 2023. Using large language models to simulate multiple humans and replicate human subject studies. *PMLR*, 337–371.
- [2] Lisa P Argyle, Ethan C Busby, Nancy Fulda, Joshua R Gubler, Christopher Rytting, and David Wingate. 2023. Out of one, many: Using language models to simulate human samples. *Political Analysis* 31, 3 (2023), 337–351. Publisher: Cambridge University Press.
- [3] David Bamman, Jacob Eisenstein, and Tyler Schnoebelen. 2014. Gender identity and lexical variation in social media. *Journal of Sociolinguistics* 18, 2 (2014), 135–160. Publisher: Wiley Online Library.
- [4] Liam F Beiser-McGrath and Robert A Huber. 2018. Assessing the relative importance of psychological and demographic factors for predicting climate and environmental attitudes. *Climatic change* 149 (2018), 335–347. Publisher: Springer.
- [5] Adam J Berinsky. 2013. *Silent voices: Public opinion and political participation in America*. Princeton University Press.
- [6] Federico Bianchi, Silvia Terragni, and Dirk Hovy. 2021. Pre-training is a Hot Topic: Contextualized Document Embeddings Improve Topic Coherence. 759–766.
- [7] James Bisbee, Joshua Clinton, Cassy Dorff, Brenton Kenkel, and Jennifer Larson. [n.d.]. Artificially Precise Extremism: How Internet-trained Lms Exaggerate Our Differences. ([n.d.]). Publisher: SocArXiv.
- [8] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research* 3, Jan (2003), 993–1022.
- [9] Gerlof Bouma. 2009. Normalized (pointwise) mutual information in collocation extraction. *Proceedings of GSCl* 30 (2009), 31–40. Publisher: Potsdam.
- [10] Belinda Bray, Bev France, and John K Gilbert. 2012. Identifying the essential elements of effective science communication: What do the experts say? *International Journal of Science Education, Part B* 2, 1 (2012), 23–41. Publisher: Taylor & Francis.
- [11] Dominique Brossard and Bruce V Lewenstein. 2009. A critical appraisal of models of public understanding of science: Using practice to inform theory. In *Communicating science*. Routledge, 25–53.
- [12] Sara E Brownell, Jordan V Price, and Lawrence Steinman. 2013. Science communication to the general public: why we need to teach undergraduate and graduate students this skill as part of their formal scientific training. *Journal of undergraduate neuroscience education* 12, 1 (2013), E6. Publisher: Faculty for Undergraduate Neuroscience.
- [13] Robert J Brulle, Jason Carmichael, and J Craig Jenkins. 2012. Shifting public opinion on climate change: an empirical assessment of factors influencing concern over climate change in the US, 2002–2010. *Climatic change* 114, 2 (2012), 169–188. Publisher: Springer.
- [14] Katherine N Canfield, Sunshine Menezes, Shayle B Matsuda, Amelia Moore, Alycia N Mosley Austin, Bryan M Dewsbury, Mónica I Feliú-Mójer, Katharine WB McDuffie, Kendall Moore, and Christine A Reich. 2020. Science communication demands a critical approach that centers inclusion, equity, and intersectionality. *Frontiers in Communication* 5 (2020), 2. Publisher: Frontiers Media SA.
- [15] Yong Cao, Li Zhou, Seolhwa Lee, Laura Cabello, Min Chen, and Daniel Herschovich. 2023. Assessing Cross-Cultural Alignment between ChatGPT and Human Societies: An Empirical Study. 53–67.
- [16] Myra Cheng, Tiziano Piccardi, and Diyi Yang. 2023. CoMPoSIT: Characterizing and Evaluating Caricature in LLM Simulations. 10853–10875.
- [17] Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, and Siddhartha Brahma. 2022. Scaling instruction-finetuned language models. *arXiv preprint arXiv:2210.11416* (2022).
- [18] Julian Coda-Forno, Kristin Witte, Akshay K Jagadish, Marcel Binz, Zeynep Akata, and Eric Schulz. 2023. Inducing anxiety in large language models increases exploration and bias. *arXiv preprint arXiv:2304.11111* (2023).
- [19] Emily Dawson. 2018. Reimagining publics and (non) participation: Exploring exclusion from science communication through the experiences of low-income, minority ethnic groups. *Public Understanding of Science* 27, 7 (2018), 772–786. Publisher: SAGE Publications Sage UK: London, England.
- [20] Dylan Degeling and Ruud Koolen. 2022. Communicating Climate Change to a Local but Diverse Audience: On the Positive Impact of Locality Framing. *Environmental Communication* 16, 2 (2022), 243–261. Publisher: Taylor & Francis.
- [21] Dorottya Demszky, Diyi Yang, David S Yeager, Christopher J Bryan, Margaret Clapper, Susannah Chandhok, Johannes C Eichstaedt, Cameron Hecht, Jeremy Jamieson, and Meghann Johnson. 2023. Using large language models in psychology. *Nature Reviews Psychology* (2023), 1–14. Publisher: Nature Publishing Group US New York.
- [22] Thomas Dietz, Amy Dan, and Rachael Shwom. 2007. Support for climate change policy: Social psychological and social structural influences. *Rural sociology* 72, 2 (2007), 185–214. Publisher: Wiley Online Library.
- [23] Baruch Fischhoff. 2019. Evaluating science communication. *Proceedings of the National Academy of Sciences* 116, 16 (2019), 7670–7675. Publisher: National Acad Sciences.
- [24] Stathis Grapsas, Andrik I Becht, and Sander Thomaes. 2023. Self-focused value profiles relate to climate change skepticism in young adolescents. *Journal of Environmental Psychology* 87 (2023), 101978. Publisher: Elsevier.
- [25] Maarten Grootendorst. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794* (2022).
- [26] Jonathan Grudin. 2006. Why personas work: The psychological evidence. *The persona lifecycle* 12 (2006), 642–664. Publisher: Elsevier.
- [27] Seungju Han, Beomsu Kim, Jin Yong Yoo, Seokjun Seo, Sangbum Kim, Enkhbayar Erdene, and Buru Chang. 2022. Meet Your Favorite Character: Open-domain Chatbot Mimicking Fictional Characters with only a Few Utterances. 5114–5132.
- [28] Sabrina V. Helm, Amanda Pollitt, Melissa A. Barnett, Melissa A. Curran, and Zeligann R. Craig. 2018. Differentiating environmental concern in the context of psychological adaption to climate change. *Global Environmental Change* 48 (Jan. 2018), 158–167. <https://doi.org/10.1016/j.gloenvcha.2017.11.012>
- [29] Peter D Howe, Matto Mildenberger, Jennifer R Marlon, and Anthony Leiserowitz. 2015. Geographic variation in opinions on climate change at state and local scales in the USA. *Nature climate change* 5, 6 (2015), 596–603. Publisher: Nature Publishing Group UK London.
- [30] Tyler P Jacobs and Allen R McConnell. 2022. Self-transcendent emotion dispositions: Greater connections with nature and more sustainable behavior. *Journal of Environmental Psychology* 81 (2022), 101797. Publisher: Elsevier.
- [31] Maurice Jakesch, Advait Bhat, Daniel Buschek, Lior Zalmanson, and Mor Naaman. 2023. Co-writing with opinionated language models affects users' views. 1–15.
- [32] Bernard J Jansen, Soon-gyo Jung, and Joni Salminen. 2023. Employing large language models in survey research. *Natural Language Processing Journal* 4 (2023), 100020. Publisher: Elsevier.
- [33] Charlotte Jones, Donald W Hine, and Anthony DG Marks. 2017. The future is now: Reducing psychological distance to increase public engagement with climate change. *Risk Analysis* 37, 2 (2017), 331–341. Publisher: Wiley Online Library.
- [34] Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, and Eyke Hüllermeier. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences* 103 (2023), 102274. Publisher: Elsevier.
- [35] Fabian Kieser, Peter Wulff, Jochen Kuhn, and Stefan Küchemann. 2023. Educational data augmentation in physics education research using ChatGPT. *Physical Review Physics Education Research* 19, 2 (2023), 020150. Publisher: APS.
- [36] Junghwan Kim, Jinyung Lee, Kee Moon Jang, and Ismini Lourentzou. 2024. Exploring the limitations in how ChatGPT introduces environmental justice issues in the United States: A case study of 3,108 counties. *Telematics and Informatics* 86 (Feb. 2024), 102085. <https://doi.org/10.1016/j.tele.2023.102085>
- [37] Hadas Kotek, Rikker Dockum, and David Sun. 2023. Gender bias and stereotypes in Large Language Models. 12–24.
- [38] Kristen A Kunkle and Martha C Monroe. 2019. Cultural cognition and climate change education in the US: Why consensus is not enough. *Environmental Education Research* 25, 5 (2019), 633–655. Publisher: Taylor & Francis.
- [39] Sanguk Lee, Tai-Quan Peng, Matthew H Goldberg, Seth A Rosenthal, John E Kotcher, Edward W Maibach, and Anthony Leiserowitz. 2023. Can large language

- models capture public opinion about global warming? an empirical assessment of algorithmic fidelity and bias. *arXiv preprint arXiv:2311.00217* (2023).
- [40] Tien Ming Lee, Ezra M Markowitz, Peter D Howe, Chia-Ying Ko, and Anthony A Leiserowitz. 2015. Predictors of public climate change awareness and risk perception around the world. *Nature climate change* 5, 11 (2015), 1014–1020. Publisher: Nature Publishing Group UK London.
- [41] Anthony Leiserowitz, Connie Roser-Renouf, Jennifer Marlon, and Edward Maibach. 2021. Global Warming's Six Americas: a review and recommendations for climate change communication. *Current Opinion in Behavioral Sciences* 42 (2021), 97–103. Publisher: Elsevier.
- [42] Weizhe Lin, Bo-Hsiang Tseng, and Bill Byrne. 2021. Knowledge-Aware Graph-Enhanced GPT-2 for Dialogue State Tracking. 7871–7881.
- [43] Li Lucy, Divya Tadimeti, and David Bamman. 2022. Discovering Differences in the Representation of People using Contextualized Semantic Axes.
- [44] Julia M Markel, Steven G Opferman, James A Landay, and Chris Piech. 2023. GPTeach: Interactive TA Training with GPT Based Students. (2023). Publisher: EdArXiv.
- [45] Roberto Navigli, Simone Conia, and Björn Ross. 2023. Biases in Large Language Models: Origins, Inventory and Discussion. *ACM Journal of Data and Information Quality* (2023). Publisher: ACM New York, NY.
- [46] Brigitte Nerlich, Nelya Kotevko, and Brian Brown. 2010. Theory and language of climate change communication. *Wiley Interdisciplinary Reviews: Climate Change* 1, 1 (2010), 97–110. Publisher: Wiley Online Library.
- [47] Jeongeun Park and DaEun Choi. 2023. AudiLens: Configurable LLM-Generated Audiences for Public Speech Practice. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 1–3. <https://doi.org/10.1145/3586182.3625114>
- [48] Joon Sung Park, Lindsay Popowski, Carrie Cai, Meredith Ringel Morris, Percy Liang, and Michael S Bernstein. 2022. Social simulacra: Creating populated prototypes for social computing systems. 1–18.
- [49] Adam R Pearson, Guadalupe A Bacio, Sarah Naiman, Rainer Romero-Canyas, and Jonathon P Schuldt. 2021. Cultural determinants of climate change opinion: familism predicts climate beliefs and policy support among US Latinos. *Climatic Change* 167 (2021), 1–8. Publisher: Springer.
- [50] Adam R Pearson, Matthew T Ballew, Sarah Naiman, and Jonathon P Schuldt. 2017. Race, class, gender and climate change communication. In *Oxford research encyclopedia of climate science*.
- [51] Emily Polk and Sibyl Diver. 2020. Situating the scientist: creating inclusive science communication through equity framing and environmental justice. *Frontiers in Communication* 5 (2020), 6. Publisher: Frontiers Media SA.
- [52] Jennifer C Price, Iain A Walker, and Fabio Boschetti. 2014. Measuring cultural values and beliefs about environment to identify their role in climate change responses. *Journal of Environmental Psychology* 37 (2014), 8–20. Publisher: Elsevier.
- [53] Nils Reimers and Iryna Gurevych. 2019. Sentence-bert: Sentence embeddings using siamese bert-networks. *arXiv preprint arXiv:1908.10084* (2019).
- [54] Connie Roser-Renouf, Neil Stenhouse, Justin Rolfe-Redding, Edward Maibach, and Anthony Leiserowitz. 2015. Engaging diverse audiences with climate change: Message strategies for global warming's six Americas. In *The Routledge handbook of environment and communication*. Routledge, 388–406.
- [55] David Rozado. 2023. The political biases of chatgpt. *Social Sciences* 12, 3 (2023), 148. Publisher: MDPI.
- [56] Leonard Salewski, Stephan Alaniz, Isabel Rio-Torto, Eric Schulz, and Zeynep Akata. 2023. In-Context Impersonation Reveals Large Language Models' Strengths and Biases. *arXiv preprint arXiv:2305.14930* (2023).
- [57] Shibani Santurkar, Esin Durmus, Faisal Ladhak, Cinoo Lee, Percy Liang, and Tatsunori Hashimoto. 2023. Whose opinions do language models reflect? *arXiv preprint arXiv:2303.17548* (2023).
- [58] P Wesley Schultz and Lynnette Zelezny. 2003. Reframing environmental messages to be congruent with American values. *Human ecology review* (2003), 126–136. Publisher: JSTOR.
- [59] Omar Shaikh, Hongxin Zhang, William Held, Michael Bernstein, and Diyi Yang. 2022. On Second Thought, Let's Not Think Step by Step! Bias and Toxicity in Zero-Shot Reasoning. *arXiv preprint arXiv:2212.08061* (2022).
- [60] Emily Sheng, Kai-Wei Chang, Prem Natarajan, and Nanyun Peng. 2021. Societal Biases in Language Generation: Progress and Challenges. 4275–4293.
- [61] Rachael Shwom, David Bidwell, Amy Dan, and Thomas Dietz. 2010. Understanding US public support for domestic climate change policies. *Global Environmental Change* 20, 3 (2010), 472–482. Publisher: Elsevier.
- [62] Chenglei Si, Zhe Gan, Zhengyuan Yang, Shuohang Wang, Jianfeng Wang, Jordan Lee Boyd-Graber, and Lijuan Wang. 2022. Prompting GPT-3 To Be Reliable.
- [63] Irene Solaiman and Christy Dennison. 2021. Process for adapting language models to society (palms) with values-targeted datasets. *Advances in Neural Information Processing Systems* 34 (2021), 5861–5873.
- [64] Ornit Spekter-Levy, Bat-Sheva Eylon, and Zahava Scherz. 2009. Teaching scientific communication skills in science studies: Does it make a difference? *International journal of science and mathematics education* 7 (2009), 875–903. Publisher: Springer.
- [65] Anaïs Tack and Chris Piech. 2022. The AI Teacher Test: Measuring the Pedagogical Ability of Blender and GPT-3 in Educational Dialogues. 522.
- [66] Mary Talbot. 2003. Gender stereotypes: Reproduction and challenge. *The handbook of language and gender* (2003), 468–486. Publisher: Wiley Online Library.
- [67] Yan Tao, Olga Viberg, Ryan S Baker, and Rene F Kizilcec. 2023. Auditing and Mitigating Cultural Bias in LLMs. *arXiv preprint arXiv:2311.14096* (2023).
- [68] Saeid Ashraf Vaghefi, Dominik Stammach, Veruska Muccione, Julia Bingler, Jingwei Ni, Mathias Kraus, Simon Allen, Chiara Colesanti-Senni, Tobias Wekhof, and Tobias Schimanski. 2023. Chatclimate: Grounding conversational AI in climate science. *Communications Earth & Environment* 4, 1 (2023), 480. Publisher: Nature Publishing Group UK London.
- [69] Fionnuala Waldron, Brian Ruane, R Oberman, and Sonia Morris. 2019. Geographical process or global injustice? Contrasting educational perspectives on climate change. *Environmental Education Research* 25, 6 (2019), 895–911. Publisher: Taylor & Francis.
- [70] Angelina Wang, Vikram V Ramaswamy, and Olga Russakovsky. 2022. Towards intersectionality in machine learning: Including more identities, handling under-representation, and performing evaluation. 336–349.
- [71] Lorraine Whitmarsh. 2011. Scepticism and uncertainty about climate change: Dimensions, determinants and change over time. *Global environmental change* 21, 2 (2011), 690–700. Publisher: Elsevier.
- [72] Graeme John Wilson. 2019. "They See a Caricature": Expanding Media Representations of Black Identity in Dear White People. *The Popular Culture Studies Journal* 7, 2 (2019), 195–216.
- [73] Patrick Y Wu, Jonathan Nagler, Joshua A Tucker, and Solomon Messing. 2023. *Large Language Models Can Be Used to Estimate the Latent Positions of Politicians*. Technical Report. Working Paper, retrieved 2023/10/10.
- [74] Yi Yang, Qingwen Zhang, Ci Li, Daniel Simões Marta, Nazre Batool, and John Folkesson. 2024. Human-centric autonomous systems with llms for user command reasoning. 988–994.
- [75] Travis Zack, Eric Lehman, Mirac Suzgun, Jorge A Rodriguez, Leo Anthony Celi, Judy Gichoya, Dan Jurafsky, Peter Szolovits, David W Bates, and Raja-Elie E Abdunour. 2024. Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: a model evaluation study. *The Lancet Digital Health* 6, 1 (2024), e12–e22. Publisher: Elsevier.
- [76] Baobao Zhang, Sander Van Der Linden, Matto Mildenberger, Jennifer R Marlon, Peter D Howe, and Anthony Leiserowitz. 2018. Experimental effects of climate messages vary geographically. *Nature climate change* 8, 5 (2018), 370–374. Publisher: Nature Publishing Group UK London.